



**HAL**  
open science

# Économie informelle en Haïti, marché du travail et pauvreté : analyses quantitatives

Roseman Aspilaire

► **To cite this version:**

Roseman Aspilaire. Économie informelle en Haïti, marché du travail et pauvreté : analyses quantitatives. Economies et finances. Université Paris-Est, 2017. Français. NNT : 2017PESC0122 . tel-01938993

**HAL Id: tel-01938993**

**<https://theses.hal.science/tel-01938993v1>**

Submitted on 29 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# UNIVERSITÉ PARIS-EST CRETEIL

ÉCOLE DOCTORALE ORGANISATIONS, MARCHÉS &  
INSTITUTIONS

Laboratoire ÉRUDITE

## THÈSE DE DOCTORAT

---

# Economie Informelle en Haïti, Marché du Travail et Pauvreté: Analyses Quantitatives

---

*Auteur:*

**Roseman ASPILAIRE**

*Directeur:*

**Philippe ADAIR**

*En vue de l'obtention du grade de  
DOCTEUR DE L'UNIVERSITÉ PARIS-EST*

*Discipline : Sciences Économiques*

Présentée et soutenue publiquement le 03 Novembre 2017

### Composition du Jury:

*Examineur:* M. François **LEGENDRE**

*Examineur:* Mme. Melika **BEN SALEM**

*Examineur:* Mr. Michel **ROCCA**

*Rapporteur:* M. Jacques **CHARMES**

*Rapporteur:* M. Zaka **RATSIMALAHELO**

*Directeur:* Mr. Philippe **ADAIR**

Professeur, Université Paris-Est Créteil

Professeur, Université Paris-Est Marne-la-Vallée

Professeur, Université Grenoble-Alpes

Professeur Emerite, Institut de Recherche  
sur le Développement

Maître de Conférences HDR, Université de  
Franche-Comté

Maître de Conférences HDR, Université  
Paris-Est Créteil



UNIVERSITY PARIS EST CRETEIL

ORGANISATIONS, MARKETS & INSTITUTIONS DOCTORAL  
SCHOOL

ERUDITE Research Lab

DOCTORAL THESIS

---

# Informal Economy in Haiti, Labour Market and Poverty: A Quantitative Analysis

---

*Author:*

Roseman ASPILAIRE

*Supervisor:*

Philippe ADAIR

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in*

*Economics*

Presented and defended on November 3, 2017

Committee members:

*Examiner:* Mr. François LEGENDRE

*Examiner:* Mrs. Melika BEN SALEM

*Examiner:* Mr. Michel ROCCA

*Reviewer:* Mr. Jacques CHARMES

*Reviewer:* Mr. Zaka RATSIMALAHELO

*Supervisor:* Mr. Philippe ADAIR

Professor, Université Paris-Est Créteil

Professor, Université Paris-Est Marne-la-Vallée

Professor, Université Grenoble-Alpes

Professor Emeritus, Institut de Recherche  
sur le Développement

Associate Professor, Université de Franche-Comté

Associate Professor, Université Paris-Est Créteil



## **Declaration of Authorship**

L'Université Paris-Est n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.



## Résumé

La prédominance de l'informel dans l'économie d'Haïti, où plus de 80% de la population vit en dessous du seuil de la pauvreté et plus de 35% au chômage, laisse entrevoir des liens étroits entre l'économie informelle, la pauvreté et le marché du travail. Faire ressortir ces interrelations, exige une évaluation de cette économie informelle qui fait l'objet des quatre chapitres de notre thèse traitant successivement l'évolution de la situation macroéconomique, le capital humain, les gains des travailleurs informels, et la segmentation du marché du travail.

Le premier chapitre fait un diagnostic du phénomène selon l'état des lieux des théories élaborées et l'évolution du cadre macroéconomique d'Haïti de 1980 à 2010 et propose une évaluation macroéconomique de l'informel à partir d'un modèle PLS (Partial Least Squares) en pourcentage du PIB. Le chapitre deux établit les relations entre l'évolution de l'économie informelle, dérégulation et politiques néolibérales grâce à un modèle LISREL (Linear Structural Relations). Nous examinons les incidences des politiques fiscales, budgétaires et monétaires des 30 dernières années sur l'économie informelle. Nous réévaluons aussi les causes de l'évolution de l'informel généralement évoquées par les études empiriques (taxes, sécurité sociale).

Au chapitre trois, nous analysons la dimension micro-réelle de l'informel grâce à un modèle des gains à la Mincer estimé par les équations logit à partir des données d'une enquête nationale sur l'emploi et l'économie informelle (EEEI) de 2007. Nous analysons les déterminants des gains informels au regard de la position des travailleurs sur le marché (saliés, entrepreneurs et indépendants) ; et les revenus (formels et informels) et les caractéristiques socioéconomiques des travailleurs pauvres et non-pauvres par rapport au seuil de pauvreté. Au chapitre quatre, nous testons d'abord la compétitivité et la segmentation du marché de l'emploi en faisant usage de modèle de Roy et du modèle de Roy élargi à travers une estimation d'un modèle Tobit. Nous utilisons un modèle de Processus de Dirichlet : d'abord analyser la segmentation et la compétitivité éventuelle du marché du travail informel ainsi que ses déterminants, selon les données de l'EEEI-2007 ; ensuite, pour distinguer les caractéristiques fondamentales des informels involontaires (exclus du marché du travail formel) de celles des informels volontaires qui en retirent des avantages comparatifs.

**Mots-clés** : économie informelle, Haïti, LISREL, marché du travail, pauvreté, PLS, processus de Dirichlet.

**JEL**: O170, C51



## Abstract

The predominance of the informal sector in the economy of Haiti, where more than 80% of the population lives below the threshold of poverty and more than 35% unemployed, suggests links between the informal economy, poverty and the labour market. Highlight these interrelationships, requires an assessment of the informal economy, which is the subject of the four chapters of this thesis, dealing successively with the evolution of the macroeconomic situation, human capital, the informal earnings of workers, and the segmentation of the labour market.

The first chapter made a diagnosis of the phenomenon according to the State of affairs of the developed theories and the evolution of the macroeconomic framework of Haiti from 1980 to 2010. And then offers a macroeconomic assessment of the informal sector as a percentage of GDP from a PLS (Partial Least Squares). Chapter two sets out the relationship between the evolution of the informal economy, deregulation and neo-liberal policies through a LISREL (Linear Structural Relations) model. We look at the impact of the budgetary, fiscal and monetary policies of the past 30 years on the informal economy. We also reassess the causes of the evolution of the informal economy generally evoked by the empirical studies (taxes, social security). In the chapter three, we analyse the micro-real dimension of the informal economy through a model of the Mincer earnings estimated by the equations logit from data in a national survey on employment and the informal economy (EEEI) in 2007. We analyse the determinants of informal gains in terms of the position of the market workers (employees, entrepreneurs and self-employed); and revenues (formal and informal) and the socio-economic characteristics of the working poor and non-poor compared to the poverty line.

In chapter four, we first test the competitiveness and the segmentation of the labour market by making use of model of Roy and the expanded Roy model through an estimate a model Tobit. We use a model of Dirichlet process: first analyse the segmentation and possible informal work and market competitiveness as its determinants, according to data from the EEEI 2007; then, to distinguish the fundamental characteristics of the involuntary informal (excluded from the formal labour market) than the voluntary informal who gain comparative advantages.

**Keywords:** Dirichlet process, Haiti, Informal economy, Labour market, LISREL, PLS, Poverty.

**JEL:** O170, C51

## *Acknowledgements*

My first and greatest thanks go to my advisor Philippe ADAIR for his seamless availability, his precious advices, his kindness and the way he coached me. Without his help and his encouragement, this dissertation would not be completed. He taught me how to seek rigorous answers to important questions.

I am very thankful to the members of my defence committee, particularly the Emeritus Professor Jacques CHARMES, and Professor Zaka RATSIMALAHELO for reviewing this dissertation. I also want to thank the examiners of the defence committee, Professor François LEGENDRE, Professor Melika BEN SALEM, and Professor Michel ROCCA for agreeing to be members of my examination committee.

I am especially grateful to my courageous mother Roselaine CHARLENE, who made me what I am, and my father Truman ASPILAIRE, the baobab that protects me from the sun rays.

I would never have achieved doctoral studies if I had not met the inspiring Professor Lesly DEJEAN during my early university curriculum. I would also like to thank Bérard CENATUS, director of Ecole Normale Supérieure of Haiti. I would like to address special thanks to my high school teacher in mathematics Eugene CARTTIGHT, who was the first to inspire the love for science in me, probably my career would have been different if I had not met him.

Thanks to my wife Blondine RODNEY who has bring many sunbeams in my life during the difficulties that I have met during this study.

Special thanks to Walguen OSCAR for his help in producing the latex codes for the tables particularly.

Words are not sufficient to thank Adélita ROMAIN, who transformed into tables all the R results that I found when implementing the statistical models.

Big thanks to Nadège LAFONTANT, Septimus W. PIERRE, and Obest LOUISSAINT for their encouragement and their support.

My last year at Creteil would have been less interesting, productive and fun without the conferences and seminars as well as all my peers of the laboratory ERUDITE, especially Oscar Manco LOPEZ.



# Contents

<b>Résumé</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>Introduction</b>	<b>1</b>
<b>1 The shadow Economy in Haiti: a counter-cyclical impact on GDP</b>	<b>9</b>
1.1 Introduction	9
1.2 Literature Review and Evolution of the Macroeconomic Framework of Haiti (1980-2010)	10
1.2.1 The informal economy: operational definitions, institutional and criterion-referenced	11
1.2.2 Theoretical approaches: Dualism, institutionalism and structuralism: towards synthesis?	12
Dualistic Approach	13
Theory of Institutionalism	15
The Unifying Theory	16
1.3 Macroeconomic situation and informality growth in Haiti 1980-2010	17
1.3.1 Macroeconomic outlook	18
Liberalization of the economy	18
Growth of self-employment and wage decline	19
1.3.2 Informality and GDP growth: a countercyclical relationship	19
Macroeconomic factors of informal expansion	19
Links between informality and GDP growth	20
1.4 Modelling and measurement methods of the informal economy	21
1.4.1 Measurement methods	21
Direct methods	21
Indirect Methods	22
1.4.2 Modelling methodology	24
Theoretical structural equation model with latent variables: the PLS approach	25
Assumptions and estimation method	26
1.4.3 Theoretical justification of the economic model and variable use	26
Macroeconomic and Microeconomic foundations	27

	Justification of econometric modelling . . . . .	28
1.5	Design and Interpretation of the Model: Variables, Equations, Estimates and Results . . . . .	29
1.5.1	Data, variables and modelling . . . . .	29
	Data . . . . .	29
	Modelling . . . . .	30
	Estimates . . . . .	31
1.5.2	Model validation . . . . .	32
	Block's Unidimensionality and Reliability . . . . .	33
	Global Model and sub-models Quality . . . . .	34
1.5.3	Results analysis . . . . .	36
1.6	Conclusion . . . . .	38
<b>2</b>	<b>Informal Economy Growth in Haiti: A Critical Evaluation of the Neoliberal and Deregulatory Policies</b>	<b>41</b>
2.1	Introduction . . . . .	41
2.2	Fiscal, Budgetary and Monetary Policies : Some Styl- ized Facts . . . . .	43
2.2.1	Budgetary and Fiscal Policies . . . . .	43
2.2.2	Monetary policy . . . . .	46
2.3	Econometric Modelling . . . . .	50
2.3.1	Theoretical model, assumptions and estimation method . . . . .	50
	Theoretical justification of the Econometric model variables . . . . .	55
	Theory on automatic stabilization of Fiscal Policy	56
	Monetary Policy: theory of liquidity preference	57
2.3.2	Structural Equation Modelling: Theoretically and practical considerations . . . . .	58
	Usual dynamic models: inadequate to analyse the shadow economy . . . . .	58
	Unobserved phenomenon: Advantages of Struc- tural Equation Modelling . . . . .	59
2.4	Model Designing . . . . .	60
2.4.1	Data description and Model . . . . .	60
	Data and Software . . . . .	60
	Conditions on Multicollinearity and Multivari- ate normality . . . . .	61
2.4.2	Model Design . . . . .	63
2.4.3	Strategy of Analysis . . . . .	66
2.5	Estimates, Model Validation and Analysis Issues . . . . .	67
2.5.1	Model Validation . . . . .	67
2.5.2	Estimates and Model choice . . . . .	71
2.5.3	Analysis Issues . . . . .	74
2.6	Conclusion . . . . .	77

<b>3</b>	<b>Employment, Poverty and Human Capital: Is informal employment in Haiti a result of exclusion?</b>	<b>81</b>
3.1	Introduction . . . . .	81
3.2	Stylized facts of poverty and the informal employment in Haiti . . . . .	84
3.2.1	Evolution of poverty in Haiti . . . . .	84
3.2.2	Public policies against poverty . . . . .	86
3.2.3	Linkage between informality and poverty . . . . .	88
3.3	Literature review of human capital theory . . . . .	89
3.3.1	The main theoretical models . . . . .	90
	Becker's Model . . . . .	90
	Ben-Porath's Model . . . . .	91
	The separation theorem . . . . .	91
3.3.2	The Mincer Wage Model . . . . .	92
3.4	Data and Econometric Approach . . . . .	95
3.4.1	Multinomial Logit Model . . . . .	96
3.4.2	Data . . . . .	99
3.4.3	Model designing . . . . .	102
3.5	Analysis issues . . . . .	103
3.5.1	Determinants of revenues in the informal employment market . . . . .	104
3.5.2	Human capital and socio-professional characteristics on the informal earnings . . . . .	106
3.5.3	Market exclusion and Informal employment . . . . .	109
	Poverty in the employment market . . . . .	109
	Comparing informal and formal non-poor workers' characteristics . . . . .	111
3.5.4	Conclusion . . . . .	114
<b>4</b>	<b>Investigating labour market segmentation with Dirichlet Process model: the informal labour market in Haiti</b>	<b>117</b>
4.1	Introduction . . . . .	117
4.2	Literature Review on labour market segmentation and mobility in Haiti . . . . .	120
4.2.1	The segmented Labour market . . . . .	121
4.2.2	The integrated informal market . . . . .	122
4.3	The Competitive and Segmented Models . . . . .	123
4.3.1	The Segmented Model . . . . .	123
4.3.2	The competitive Model . . . . .	125
4.4	Bayesian Model: Dirichlet Process . . . . .	126
4.4.1	The Model . . . . .	126
4.4.2	Implementing a Dirichlet Process: Theoretical foundation and Practical Reasons . . . . .	129
4.5	Estimating the Dirichlet Process . . . . .	130
4.5.1	Estimating the number of clusters by the Cutoff Method . . . . .	131

4.5.2	Estimating the model parameters by the Expectation- Maximization Algorithm . . . . .	132
4.6	Empirical Analysis . . . . .	133
4.6.1	Situation of the Labour Market and Data de- scription . . . . .	133
4.6.2	Testing Labour Markets Differences . . . . .	136
4.6.3	Can a unique wage function represent the global labour market? . . . . .	138
4.6.4	Structure of the labour market: Segmentation or Competition? . . . . .	142
4.7	Conclusion . . . . .	149
<b>Conclusion</b>		<b>151</b>
<b>A Calculation of the Estimates of the PLS Model</b>		<b>155</b>
A.1	Estimates of the coefficients . . . . .	155
A.2	Tables of PLS models . . . . .	158
<b>B Mathematical Validity of Bootstrapping</b>		<b>159</b>
<b>C Mincer model and the theorem of separation</b>		<b>163</b>
C.1	Theoretical Aspect . . . . .	163
C.2	Ordinal multinomial logit estimates of the Mincer wages model by category of workers . . . . .	165
<b>D Markov Chain Monte Carlo in the Dirichlet Process</b>		<b>169</b>
D.1	Describing of the Markov Chain process into the clusters	169
D.2	Describing of the Markov Chain Monte Carlo in the Dirichlet Process . . . . .	170

# List of Figures

1.1	The PLS Model . . . . .	25
2.1	Evolution of Gross Domestic Product and Net Domestic Credit 1991-2011 . . . . .	45
2.2	Evolution of the Consumer Price Index and the Minimum Legal wage 1987-2010 . . . . .	46
2.3	Evolution of the exchange rate, loans to private sector and deposit 1991-2010 . . . . .	48
2.4	Evolution of the inflation and the circulating currency (M2) 1989-2010 . . . . .	48
3.1	Figure 3.1: Densities of monthly log-wages for workers according to the gender . . . . .	100
4.1	Multinomial sampling with a Dirichlet prior, noisy observations and the auxiliary variable . . . . .	128
4.2	Densities of monthly log-wages for formal and informal workers . . . . .	139
D.1	Earnings distribution according to educational level . . . . .	171





# List of Tables

1.1	Comparison of GDP growth and increase in the informal economy (in %) . . . . .	21
1.2	Estimation results . . . . .	31
1.3	Coefficient of the latent variables driving the shadow economy . . . . .	32
1.4	Composite reliability . . . . .	33
1.5	Relationship between manifest variables and latent variables by block . . . . .	34
1.6	Coefficient of communalities and redundancies of the model . . . . .	35
1.7	Fit model quality . . . . .	35
2.1	Multicollinearity test for indicators variables given by the correlation matrix . . . . .	61
2.2	Jarque-Bera normality test of the variables . . . . .	62
2.3	Shapiro-Wilk multivariate normality test . . . . .	62
2.4	Indices of the model goodness-fit of data . . . . .	68
2.5	Estimates of the six a priori models . . . . .	72
2.6	Indices of the model goodness-fit of data . . . . .	73
2.7	Reduced-form $R^2$ for all the variables in all of the six a priori estimated models . . . . .	74
3.1	Employment situation and consumption expenditures of the households in gourdes . . . . .	85
3.2	General characteristics of the informal workers . . . . .	101
3.3	Poverty extent among the formal and informal workers gender . . . . .	102
3.4	Multinomial logit estimates of the extended Mincer wage model of informal workers . . . . .	105
3.5	Multinomial logit estimates of the extended Mincer wage model of informal workers - Global fitting model . . . . .	106
3.6	Ordinal multinomial logit estimates of the Mincer wages model by category of workers . . . . .	107
3.7	Ordinal multinomial logit estimates of the Mincer wages model of poor workers . . . . .	110
3.8	Ordinal multinomial logit estimates of the Mincer wages model of non-poor workers . . . . .	112
4.1	Summary statistics on the labour market . . . . .	135
4.2	Labour Market Participation and Segments Comparison . . . . .	137
4.3	Wage functions for the sectors: OLS models . . . . .	139

4.4	Tobit (two-step Heckman) Model for a Unique Labour Market . . . . .	141
4.5	Selectivity of the model . . . . .	141
4.6	Component and mixture proportion for the partitioned labour markets . . . . .	143
4.7	Criterion indexes for the number of segments of the global market . . . . .	143
4.8	Semi-parametric mixture model determining wage functions for the global labour market . . . . .	144
4.9	Component and mixture proportion for the two-segments informal labour market . . . . .	145
4.10	Criterion indexes for the number of the two-segment informal market . . . . .	146
4.11	Semi-parametric mixture model determining wage functions for the global labour market . . . . .	147
4.12	Probability to belong to one segment for global labour market and informal labour market . . . . .	148
A.1	Correlations between variables blocks and components factors . . . . .	158
A.2	Correlations between variables blocks and components factors . . . . .	158
A.3	Correlations between variables blocks and components factors . . . . .	158
C.1	Estimates of the Mincer wages model by category of workers for other variables . . . . .	165
C.2	Multinomial logit estimates of the extended Mincer wage model by category of workers - Global fitting model . . . . .	166
C.3	Estimates of the Mincer wages model of poor workers for other variables . . . . .	166
C.4	Estimates of the Mincer wages model of non-poor workers for other variables . . . . .	167
C.5	Multinomial logit estimates of the extended Mincer wage model of non-poor workers - Global fitting model . . . . .	167

# List of Abbreviations

<b>AIC</b>	Akaike Information Criterion
<b>BIC</b>	Bayesian Information Criterion
<b>BRH</b>	Banque de la République d'Haïti
<b>CAIC</b>	Consistent Akaike Information Criterion
<b>CELADE</b>	Comisión Económica para América Latina y el Caribe
<b>CEPAL</b>	Commission Economique pour l'Amérique Latine
<b>CFI</b>	Comparative Fit Index Abbreviated
<b>CIA</b>	Central Intelligence Agency
<b>CNSA</b>	Coordination Nationale de la Sécurité Alimentaire
<b>CNUCED</b>	Conférence des Nations Unies sur le Commerce et le Développement
<b>CPI</b>	Consumer Prices Index
<b>CTPEA</b>	Centre de Techniques de Planification et d'Economie Appliquée
<b>DSNCRP</b>	Document de Stratégie Nationale pour la Croissance et la Réduction de la Pauvreté
<b>DYMIMIC</b>	Dynamic Multiple Indicators Multiple Causes
<b>ECBM</b>	Enquête Budget Consommation des Ménages
<b>EEEI</b>	Enquête Sur l'Emploi Et l'Economie Informelle
<b>EM</b>	Expectation-Maximization Algorithm
<b>FY</b>	Fiscal Year
<b>GDP</b>	Gross Domestic Product
<b>GNP</b>	Gross National Product
<b>HIV/AIDS</b>	Human Immunodeficiency Virus/ Acquired Immune Deficiency Syndrome
<b>ICL</b>	Integrated Completed Likelihood
<b>ICLS</b>	International Conference of Labour Statisticians
<b>IDB</b>	International Development Bank
<b>IHSI</b>	Institut Haïtien de Statistique et d'Informatique
<b>ILO</b>	International Labour Organization
<b>IMF</b>	International Monetary Fund
<b>LISREL</b>	Linear Structural Relation
<b>MCMC</b>	Markov Chain Monte Carlo
<b>MEF</b>	Ministère et l'Economie et des Finances
<b>MIMIC</b>	Multiple Indicators Multiple Causes
<b>MPCE</b>	Ministère de la Planification et de la Coopération Externe
<b>NDC</b>	Net Domestic Credit
<b>NFI</b>	Normed Fit Index
<b>NNFI</b>	Non-Normed Fit Index
<b>OECD</b>	Organization for Economic Cooperation and Development

<b>OIT</b>	Organisation Internationale du Travail
<b>OLS</b>	Ordinary Least Squares Regression
<b>ONPES</b>	Observatoire National de la Pauvreté et de l'Exclusion Sociale
<b>OR</b>	Odds Ratio
<b>OXFAM</b>	Oxford Committee for Famine Relief
<b>PARDH</b>	Action Plan for National Recovery and Development of Haiti
<b>PAS</b>	Programme d'Ajustement Structurel
<b>PLS</b>	Partial Least Squares
<b>PNUD</b>	Programme des Nations Unies pour le Développement
<b>PV</b>	Present Values
<b>RIR</b>	Real Interest Rate
<b>RMSEA</b>	Root Mean Square Error Approximation
<b>RR</b>	Risk Ratio
<b>SEM</b>	Structural Equations Model
<b>SRMR</b>	Standardized Root Mean Square Residual
<b>SVAR</b>	Structural Vector Autoregressive Models
<b>SVEC</b>	Structural Vector Error Correction
<b>TLI</b>	Tucker-Lewis Index
<b>TPI</b>	Triennial Investment Program
<b>US</b>	United States
<b>ULI</b>	Unit Loading Identification
<b>UVI</b>	Unit Variance Identification
<b>VAR</b>	Vector Auto-Regressive
<b>VECM</b>	Vector Error Correction Model

*Dedicated to my Grandmother  
Avanise JOACHIM*



# Introduction

Addressing the informal economy in Haiti is tackling one of the diverse developmental problems facing Latin America and the Caribbean region, and it is important to look at the informal economy as a process belonging to the global economy. When the debt crisis broke out in the 1980's, the economies of Latin America and the Caribbean began to dither, the recovery was slow, long and weak, so the response of the economic agents was not long in coming.

Governments have been constrained to adopt drastic macroeconomic policies to reduce inflation and run their debt service. These policies led to growing unemployment, a decline in the production system and a dramatic increase in poverty. Hence, the Latin American and Caribbean region entered a recession period, which reduced revenues and economic resources. The poverty rate among the population increased from 38% in 1980 up to 62% in 1990 (Kliksberg, 1994).

Stringent economic decisions include the privatization of State-owned enterprises, and layoffs in public administrations; eventually, populations are left aside. At the same time globalization is in full swing, workers who had no time to adapt to the new rules of the market are laid off and companies that cannot cope with the delays rapidly fall into decay. Markets local producers find themselves in fierce competition with western producers better equipped technologically and with huge capital superiority, so they will not hold out for long in this lost battle. States seem to be choosing the most promising fields of investment, even agricultural producers are entering this era of decline, so the informal way seems to be emerging for more than one third of companies and workers (Aspilaire, 2012) that could not follow the unbridled pace imposed both by capitalism and the consequences of the crisis. Thus a process of regional deregulation began to give way to informalisation. In the meantime, poverty has waged war against these populations, unemployment is in full swing, Latin America and the Caribbean are in agony because almost all countries in the region are heavily indebted. Due to growing economic weakness of the States of the region during this period, day after day informality has risen to normality.

Latin American and Caribbean populations are heading down in the process of informalisation, and more than a quarter of a century later from 1980 they seem to be unable to cope with the current economic situation. In 2005, it was estimated that 41% of the Latin American population suffered from poverty and 17% from extreme poverty, whereas informal employment was estimated at 47% in 2003



(Lamaute-Brisson, 2005). Consequently, the study of the informal economy must be placed in its socio-political-economic context, both in its definitions and its causes, as well as in its meaning and in its conception.

The prominence of the informal economy in Haiti, where more than 80% of the population live below the poverty threshold and more than 35% are unemployed, links the informal economy, poverty and the labour market. Highlighting these interrelations, requires an assessment of this informal economy. Thus we design an evaluation of the informal economy, which is based on a diagnosis of the phenomenon by making an inventory of the theories developed on the topic.

Shadow economy and informal work in Haiti is a long-standing side of the global economy. Since 1950 informal activity has been registered in the Haitian economy (Montas, 1995). Despite the early presence of informal activity in the Haitian economy in the middle of the twentieth century, its extension is very recent in the economic record of the country.

The economic crisis of the Latin American and Caribbean region in the 1980's contributed to the exponential growth of the external debt. The costs of the debts are multiple, weak growth rate of the economy, large inflation rate, extension of poverty, loss of production resources, spread of the unemployment rate and increasing economic and social inequalities. The incapacity of government and economic authorities to bring appropriate answers to the lack of productivity, unemployment, hunger, and poverty obliges populations to undertake some mix and non-orthodox domestic, economic and financial activities to survive misery and poverty. A parallel employment market was created to bypass the lack of productivity and the loss of jobs on the formal market.

Such activities belong to the concept of the informal economy. Since the informal economy has not stopped growing, whereas the economic situations of the Latin American and Caribbean populations have changed only slightly, the informal economy has invaded from households to the office all spheres of economic activity of this region, and all the markets are concerned.

Addressing these issues necessarily involves a view of the informal economy as a whole; it should not be seen as an informal sector and an informal unit of production, but as part of the global economy including goods and services on the informal market and the informal labour market. The informal labour market includes informal workers who normally work in the informal sector and salaried informal workers of the formal sector (ILO, 2013).

It becomes very important for policy makers and economic scholars to apprehend this deep and persistent phenomenon in developing countries. Indeed, policy makers and economists advocate various reasons that justify the examination of the informal economy. The

---

expansion of the informal economy can be seen as a response of economic agents feeling that they are too pressurized by the system in place and are seeking an answer to escape from the mechanics of State regulation. If this increase is due to too high taxes and social security contributions, i.e. there is institutional sclerosis, then the shift of economic agents to the informal economy, has no other reason than avoiding excessive taxes and social security contributions (Schneider and Enste, 2000). The result may be a vicious circle where the main loops are the increase in the budget deficit or in taxes, in the size of the informal economy, and the gradual weakening of the national economy and social cohesion. A flourishing informal economy can cause serious problems for policy and economic decision makers, because labour, employment, consumption and income indicators are unreliable if the informal economy is not properly measured or evaluated.

The relationship between formal and informal economy should also be taken into account. On the one hand, a prosperous informal economy may be attractive to both local and international workers. Because it can create competition for officially registered firms. On the other hand, a part of the income earned is immediately injected into the official economy, which creates a positive effect on the formal economy (Schneider and Enste, 2000).

Exhausted in its race to pay interest on the debts contracted by the rulers on her behalf, the Haitian people had never seen or felt the palliative effects of an economy sick to the bone, feverish to the base, weakened by the various internal socio-political crises of the last thirty years. Out of 8.1 million inhabitants, Haiti included in 2001, more than 4.4 million (over 56% of its population) living below the extreme poverty line of one US dollar a day (purchasing power parity), while nearly 6.3 million (about 76%,) are below the poverty line of two US dollars a day (MPCE, 2007). The country sensitivity to external shocks has been prejudicial because of the long-lasting political instability combined with the two decades of implementation of neo-liberal policies.

Unemployment is one of the fundamental canals of informality in Haiti, due to the crucial absence of regulations and financial mechanisms driving the creation of employment. Then, implementing employment policies was more than a challenge. The absorption of urban labour surplus by the assembly industry is still one of our economic beliefs inherited from those of the seventies. The establishment of subcontracting factories was strongly encouraged from the very beginning, which constituted a tacit acceptance by the State that job creation was left to the mercy of private enterprises or to the population itself through the self-employed. However, the artisanal characteristic of the production could not bring significant value added to the economy.

A country where unemployment and poverty are combined, 40.6%

of the active population of Haiti are unemployed in 2007 (IHSI, 2010). This is an economy where job insecurity is no longer to be demonstrated and where household monetary poverty is expanding every day. Facing the inability of the State to respond to this increase in the number of unemployed, hungry and needy, poor have no choice but to seek means of survival that are not usually included in the principles of the formal economy.

Nowadays, the size of the informal economy in Haiti becomes so important it seems like informal activities dominate regular economic activities. From 1999 to 2007, the share of the shadow economy is estimated around 55% of GDP. The size of the informal economy has been calculated and evaluated between 54.8 and 57.4% of GDP (Schneider, Buehn and Montenegro, 2010). In the metropolitan area of Port-au-Prince, the capital city of Haiti, informal private companies represent 81% of the business companies (IHSI, 2010). More than 54% of the employed individuals are engaged in the informal sector (IHSI, 2010). Regardless calculation methods of the share of the informal economy, shadow economy and informal work are a large part of the economic picture of Haiti.

### **Aim and objectives**

The aim of this thesis is to analyse the relationship between the informal economy, poverty and labour market in Haiti. It starts from an operational definition of the informal economy, which represents it as the set of non-illicit activities that can generate revenues and are carried out on the margins of public regulation.

This research seeks to understand the dynamics of the informal economy in Haiti, which includes the dynamics of both the goods and services market and the labour market. Dynamics is also expressed in the weight of the informal economy in the global economic system. Understanding the forms of articulation between the informal economy, poverty and labour market allows to identify the existing causal relationships among them. But such understanding does not allow a quantitative approach of the informal economy regarding poverty and employment market. One of the most appropriate ways to tackle this useful endeavour is the use of the powerful direct econometric methods, thanks to robust theoretical approaches. There is a major gap in the empirical literature on informality in Haiti, there is also an insufficient empirical evidence for the relationship between the informal economy and poverty. To date, studies regarding the informal economy in Haiti have mostly focused on the structure and organization of the informal enterprises and the participants in the informal activities (Fass, 1988, Lamaute-Brisson, 2000), rather than exploring the link between informal activity and poverty through the labour market. Only a few analyses address the issue of poverty and

informality (ONPES and World Bank, 2014) but they do not quantify and analyse the relationship.

Lack of quantitative research on the informal economy is not the only knowledge gap to fill with regards to this topic. The other gaps include studies in informality that cover the whole country. Most authors have only taken into account the metropolitan area of Port-au-Prince and the capital city; this provides too little empirical evidence for theoretical conclusions. Importantly, there is no theoretical publication regarding the link between the informal economy and deregulation policies. Moreover, there is no theoretical paper addressing the problem of individual comparative advantage in the informal labour market. This thesis intends to fill these gaps in the literature. Therefore, the objective of this thesis is to analyse the nature of the relationship between the size of the informal economy, neoliberal macroeconomic policies, poverty and the labour market. This objective can be broken down into three sub-goals stated as three (3) working hypotheses. We tackle three main questions in the light of empirical data provided by various studies and surveys carried out on the Haitian economy, including the 2007 survey of employment and the informal economy (EEEI 2007) and some fiscal, budgetary and monetary data available. Questions constitute the guiding vector of this work, the last two of which are little examined by the previous works carried out on the informal economy in Haiti.

First, how should we understand the informal economy in the sense of the Haitian reality as an economic phenomenon, its causes and consequences? To do this we must first look at the political and economic history of Haiti, which has facilitated the emergence of the underground economy, the trajectory that it has followed from its genesis to its blossoming. We need to study the informal economy characteristics and its trajectory from an economic perspective in order to better understand its various facets and to bring out all the relevant economic observations. All of that can be summarized in the following questions. How can we measure the size of the informal economy, in relation to which macroeconomic aggregates? What are the methods used as a possible instrument for measuring the informal economy? To provide answers, we state our first hypothesis:

Hypothesis 1: The weight of the informal economy explains its impact on development and the unprepared choice of neoliberal macroeconomic policies has served to expand it. On the one hand, it is this weight that determines the importance of the informal economy in the development process. Taking into account the thesis that the informal economy has the characteristics and potential to be an important factor contributing to economic and social development (De Soto, 1994) or the position of the proponents of the theory that the informal economy is a set of economic activities generating poverty and underdevelopment (Lewis, 1954; Harris and Todaro, 1970), it is the measure of the size of the informal economy that sheds light on

these two considerations. On the other hand, only a rigorous analysis of the impact of neoliberal fiscal and budgetary policies on the growth of the informal sector will allow us to conclude.

Then, the second question contains three components that can be broken down as follows: is the informal labour market a reservoir or excluded from the labour market? To answer this question, we should answer to these two following questions: what are the characteristics of the informal workers? Are the earnings of informal workers influenced by their characteristics? To make this point we formulate the following hypothesis.

Hypothesis 2: very high unemployment among low-educated individuals coupled with low levels of education (IHSI, 2010), the overall high unemployment rate (Montas, 2005) that is higher among youths aged 15-24 (Lamaute-Brisson, 2005), household income poverty (Lamaute-Brisson, 2005) and barriers to entry to the formal labour market would be the main factors draining workers into the informal economy. In this respect, the informal economy would not only include the excluded and the forgotten on the formal labour market, but also the workers with the skills to evolve in the formal labour market.

The last question can be formulated in two sub-questions: Is there a discrepancy between the wages of formal workers and those earned by informal workers? Does the informal market allow some workers to maximize their utility?

Hypothesis 3: The informal labour market is a heterogeneous and segmented market in the sense of Fields (Fields, 1990), wherein some agents benefit from comparative individual advantages (Gindling, 1991; Maloney 1999; Perry et al., 2007). The differences in revenues between the formal and informal markets are not a subject of debate (Gunther and Launov, 2011) but in both sectors there is a large gap between the incomes of employers and employees. This gap would be the result of a difference between employees and employers in terms of human capital (education, training and experience) and / or financial capital (Dickens and Lang, 1985; Rosenzweig, 1988; Magnac, 1991). This needs to be analysed in light of Mincer's earnings function (1958, 1974). There are also large income gaps for informal workers who have accumulated human capital.

## Methodology

In chapter one we address the size of the informal economy and its impact on macroeconomic development, we also analyse its evolution in relation to the official economy. We analyse the structure of the Haitian economy. First by describing the evolution of Haiti's macroeconomic framework from 1980 to 2010, then by discussing the macroeconomic determinants of the informal economy and we present a comparative analysis of the growth of informality and GDP growth. We make an inventory of the informal economy in Haiti.

---

Using a MIMIC model (PLS model), we measure the size of the informal economy in Haiti as a percentage of GDP, according to different macroeconomic variables observed over the period 2000-2010. We highlight the main links between informal and various variables of the model (GDP growth, inflation rate, government expenditure, taxes, unemployment rate and corruption).

Chapter two analyses the effects of neo-liberal economic policies such as tax cuts, light State intervention on the market, and few State regulations rather than reduction, on the expanding of the informal economy. In order to establish the links between the triptych of informal employment, deregulation and neoliberal policies, we first analyse the relationship between the application of market liberalization and the expansion of the informal economy in Haiti. We look at the broad lines of fiscal, budgetary and monetary policies over the last 30 years. We conduct this analysis by studying the impact of fiscal, fiscal and monetary policies on the informal economy that is based on a LISREL (Linear Structural Relations) econometric model. We analyse the evolution of informal activities in relation to neoliberal and deregulatory policies. We also reassess the causes of the evolution of the informal economy generally evoked by the empirical studies (taxes and social security contributions).

Chapter three is dedicated to the analysis of poverty and earnings of the informal workers regarding human capital accumulation. We make the assumption that informal labour market includes both the those excluded from the labour market with different levels of education and social conditions and workers who have voluntarily left the formal labour market to obtain comparative advantages. Income levels, depending on whether they are below the poverty line or beyond, are the true indicators. On the basis of these data, we used a Mincer model of earnings to identify the determinants of informal incomes. To estimate the Mincer model, we used polytomic logit modelling to capture the various factors influencing income of the informal agents. The microeconomic analyses of informal employment comprise three parts. First, the general characteristics of informal workers are identified. Then, we analyse the determinants of informal gains in terms of the occupational status of workers in the market (employees, entrepreneurs and self-employed). Finally, we analyse incomes (formal and informal) in relation to the poverty line. We analyse the socio-economic characteristics of the working poor and non-poor.

The fourth chapter uses a national survey, *Enquête sur l'Emploi et l'Economie Informelle (EEEI-2007)*, to determine whether the informal labour market is competitive or segmented. We use regression mixing models and Dirichlet process model to study the segmentation and potential competitiveness of the informal labour market. First, we test the competitiveness and segmentation of the labour market by using Roy model and expanded Roy model with an estimate of a

Tobit (Heckit) model. We use the Dirichlet processes to analyse the mobility between the informal and the formal labour market, and to determine the existence of entry barriers. Finally, we separate involuntary informality (excluded from the formal labour market) of from voluntary informality (working in the formal market with comparative advantages). Hence, we determine the basic characteristics of the agents that the setting up of an employment policy could either prevent from joining the informal economy or draw them out of it.

# Chapter 1

## The shadow Economy in Haiti: a counter-cyclical impact on GDP

### 1.1 Introduction

We use a PLS (partial least squares) model, a particular type of a structural equations model (SEM) to estimate the size of the shadow economy in Haiti as a share of GDP. Econometric modelling encapsulates nine causal variables and two indicators over the period 2000-2010. It shows that the fall in GDP, the rise in taxes and government expenditures and increasing inflation are the main factors that influence the growth of the shadow economy. This growth has a knock-on effect on the unemployment rate during these last ten years. From 2007 to 2010, worse is the GDP growth per annum and better is the growth of the shadow economy; thus, suggesting that the latter may be counter-cyclical.

Growing poverty of the population in Haiti, the pivotal role played by informality, about 56.4 % of GDP between 1999 and 2007 (Schneider, 2010) and the idea that the promotion of decent work is the only way to fight against poverty, is a triptych that guides our reflection on the impact of the informal economy upon the macroeconomic variables. Informality becomes more significant in Haiti in as much as it is neglected by public policy. However, no academic work has focused on macroeconomic measurement of the size of the informal economy and its impact on GDP. This underlines the originality of this chapter that addresses a quantitative aspect but especially analytical impact of informality on the Haitian economy.

Few papers address the interesting task of estimating the hidden economy (**Schneider; 2010; Schneider et al.; 2007**) Besides this challenge, this chapter addresses the impact of the informal economy on macroeconomic variables and vice versa. Some papers provide results about the main causes of informality (**Schneider; 2010**) to our knowledge this research work is the first to tackle the relationship between informality and economic growth in Haiti. Usually, among the two types of structural equations modelling with latent variables, scholars prefer the LISREL (Linear Structural Relation) model in the



empirical literature of informal economy. This chapter is an essay to fill this gap. Our study is the first to undertake the measure of the size of shadow economy using the PLS (Partial Least Squares) model. Hence, the aim of this chapter is twofold: estimating the informal economy size and analysing the relationship between macroeconomic variables and the shadow economy.

Thanks to a PLS model, which is a particular DYMIMIC (Dynamic Multiple Indicators Multiple Causes), we study the impact of the informal economy on GDP, by calculating its size as a percentage of the latter. This model uses eleven macroeconomic variables including nine causal variables and two indicators, each of these variables corresponds to the time series 2000-2010. The choice of this period is motivated by the availability and the accuracy of the data. The PLS model allows to dispense with the problem of expected signs (for each variable) presented by the other space state models. Our model is applicable to multivariate analysis such as multiple countries with available data on informality and macroeconomic variables.

Section 1 presents the literature review and the evolution of the macroeconomic framework in Haiti from 1980 to 2010. Section 2 discusses the macroeconomic determinants of informality while making a comparative analysis of the growth of the latter and that of the GDP. Section 3 is devoted to the econometric modelling and estimates. Section 4 concludes and summarizes our results and suggests avenues of research.

## **1.2 Literature Review and Evolution of the Macroeconomic Framework of Haiti (1980-2010)**

Economic literature on the informal economy in Haiti is rather disparate, nevertheless conclusive studies exist. It is worth mentioning Fass (1988), Montas (1995) and Lamaute-Brisson (2002), as well as a survey on employment and the informal economy in Haiti in 2007 (EEEI 2007) carried out by the Haitian Institute of Statistics and Information (IHSI). But, none of these studies is focused on the macroeconomic measurement of the size of the informal economy. However, there are studies on macroeconomic measurement of the size of GDP for 156 countries including Haiti (**Schneider; 2010**) using the MIMIC model or its derivatives. Nevertheless, we did not find any study on informality using PLS modelling. The economic application of this model is due to Russet (**Tenenhaus; 1999**) showing that economic inequality leads to political instability.

### 1.2.1 The informal economy: operational definitions, institutional and criterion-referenced

The informal sector as though controversial concept is here considered the totality of the informal activities conducting by unregistered small scale units of production with low size of staff employed (Adair; 2009) This meaning is legitimate to the extent that the calculation of macroeconomic aggregates, particularly gross domestic product covers all kinds of revenues generated by all economic activities (formal or informal). Given the mode of operation, characteristics and certain criteria relating to informal activities retained by the ILO (1993, 2013) and other scholars, we can identify two different issues of definitions of the informal sector.

One approach defines the informal sector by evaluating the relation to the legal and bureaucratic framework and it defines the informal sector as being constituted of enterprises that do not conform to the legal framework (ILO, 2013). Hence there is a tight relation between non-registration and informal activities. During the fifteenth International Conference of Labour Statisticians (ICLS), the ILO defines the informal entity as encompassing the following categories: the unpaid family business workers, those who work on their own account, the employees of informal employment, people banding together in informal production cooperatives, those who produce goods for the exclusive use of their households (ILO, 2013).

The second approach defines the informal sector, according to the way that the enterprises are organised and conduct their activities, as a particular form of production (ILO, 2013). Informal can be set to modes of organization and operation which it is subject to, since informality is defined according to certain methods of operation and voluntary production and / or circumstantial informal enterprises. These are the operational definitions which have the advantage of not following the current dualistic definitions entails limiting the informal economy compared to formal. They are based on the logic of production and the diversity of informal activities grouped, these forms of definition are also used by authors such as Hugon (1980) and Charmes (1987). The criterion-referenced definitions are those who look as informal grouping of companies with characteristics such as administrative and statistical non-registration, the small size, low level of technology, or low level of organization (ILO, 2013).

The informal enterprises have certain features that are specific (administrative, commercial and technical). We can meet several distinct features that are based on location and time. The acceptance of the concept of informal sector seems to be controversial according to the fact that heterogeneity is its main characteristic. Informal is antagonistic to the formal in a wide register. If informal is opposed to formal, this is as marginal. In a broad sense, the informal sector is

embodied in the concept of non-observed economy or shadow economy which includes legal activities unregulated by the State for one reason or another, and illegal activities (Schneider, 2007) prohibited by the State as drug trafficking, smuggling, prostitution and games that simply generate the illegal economy. The shadow economy is here recognized as the entire part of the also non-observed economy added to the illegal production of the economy.

However, it should be noted that the largest and the most suitable definition of the informal sector is the synthetic definition that combines national accounts (household sector) and the types of employment of unprotected categories of workers (ILO, 1993, 2013). The ILO (1993) use three main criteria to define informal sector enterprises, the legal structure of the enterprise, the ownership and the type of set of accounts maintained. The informal sector is constituted of unincorporated enterprises (any unit engaged in the production process of goods and/or services) owned and operated by individual household members, that do not conform to the legal framework, unregistered and/or small size of staff employed with the basic objective of generating jobs and revenues to the workers (ILO, 1993). However, we go further than the informal sector definition. The context of globalization and information and communications technology leads to tight competition. Decentralization of production and reorganization of work are the main answers of a growing number of firms. Outsourcing or subcontracting arrangements are among such measures that increases flexibilization and informalization of production and employment relationships. Generally casual workers, subcontractors and agency workers are not protected by labour laws. Thereafter, the term "informal sector" has been seen to be inappropriate to reflect this complex and heterogeneous dynamic of such economic interrelations. The term "informal economy" has come to be widely accepted to circumvent the growing and heterogeneous group of workers and firms operating in both rural and urban regions informally (ILO, 2002). The informal economy comprises own-account workers in sweatshop and survival activities, paid domestic workers enrolled by households, self-employed in micro-firms working on their own or with contributing family workers or sometimes trainees/employees (ILO, 2002). This meaning is legitimate to the extent that in the calculation of macroeconomic aggregates, particularly gross domestic product include all kinds of revenues generated by whole informal economy. In this chapter for all the modelling and the analyses we use the ILO (2002) definition of the informal economy.

### **1.2.2 Theoretical approaches: Dualism, institutionalism and structuralism: towards synthesis?**

The economic literature, until the mid-1990s, grants to classify approaches of the informal economy into three main currents of thought

which are dualism, structuralism and institutionalism. From the early 2000s, a fourth stream both more analytical and synthetic emerges and tries to overcome the divergence of the other approaches about the nature and causes of the informal economy and its interrelations with the formal sector. Meanwhile, the development of a universal concept of informality, beyond the specific contexts, continues.

### **Dualistic Approach**

Dualistic thinking is embodied in the theory of extra- legality (existence of arrangements economic and social micro-contracts that are not exactly in the sense of the rule of law which makes them unusable capital and little work divided) generated by the capital restructuring strategies. Dualistic approach sees informal economy as a marginal pan with no direct link with the formal economy. It has its roots in the different works developed by Lewis (1954) and Harris and Todaro (1970).

#### **The Lewis model**

According to Lewis (1954), agricultural and rural subsistence sector does not have enough jobs to meet the labour supply. The labour supply is still outstripping demand, its marginal productivity is generally negligible (Lewis, 1954). This labour surplus will fuel the demand for labour in the capitalist sector which is urbanized and industrialized. Since the supply always exceeds the labour demand in the industrial sector, the wage rates, which is around the living wage, will still be enough to attract the subsistence workers. This is explained by the capital management mode. So, the global labour market is controlled by the capitalist sector because it hires labour from the subsistence sector to grow its capital.

However, it should be noted that in this model the market is fragmented into several smaller non-homogeneous groups of agents of the capitalist sector and the subsistence sector. Lewis (1954) argues that there is a huge gap between subsistence sector employees and the capitalist sector workers. This gap is not only expressed in the earnings differences but also in term of agent characteristics. These characteristics included the levels of qualification, education, responsibility and prestige (Lewis, 1954).

Dealing with the mechanism of price formation according to the distribution of income and the effect of inflation on the price, the balance between savings and investment, the model of Lewis was a macroeconomic scope. This scope has overshadowed the dual nature of the labour market, which was to be released later in the other models, nevertheless it has opened the way to more realistic and complete models with respect to the informal economy.

### The Harris-Todaro Model

According to this model, urban unemployment is explained by higher wages in the formal sector, because full employment expected wage is equal to the minimum wage in the sector. For the Harris-Todaro model (1970), this expected gain encourages rural workers migrating to the industrial sector and not actual earnings. So, urban unemployment is an economically rational choice by rural migrant, because his expected income pending employment is higher than his gain in the rural sector, which can be explained as follows. The urban workforce ( $W_u$ ) comprises two components, one from the permanent urban proletariat characterized by no attachment to the rural sector and the other is the labour supply from rural migrants. Let  $h(\cdot)$  denote the agricultural sector's production function where  $h' > 0$  and  $h'' < 0$ . Let's denote  $P$  the price of agricultural goods in terms of industrial goods, real wage in agricultural sector ( $S_A$ ) is given by  $S_A = P * h'$ . The real wage of industrial sector ( $S_I$ ) is given by  $S_I = h'$ , due to the profit maximization behavior of entrepreneurs who are in a situation of perfect competition, on the other hand the minimum urban wage  $S_M^-$  is such  $S_M^- \leq S_I$ .

The expected urban wage ( $S_U^e$ ) is given by  $S_U^e = S_M^- * p$ , where  $p$  is the probability to have an employment, which is equal to  $W_u/L_T$  where  $L_T$  is the total labour. The migration of rural workers will stop only when agricultural productivity in terms of manufactured goods is equal to the expected urban wage, i.e. when there the balance  $S_U^e = S_A$  is realized. To reach this situation, two cases should be considered, either in full employment ( $W_u = L_T$ ) when finding a job is no longer hypothetical ( $p = 1$ ) and  $S_U^e = S_M^-$  or the expected urban income differential-rural income is zero i.e.  $S_M^- * p = P * h'$ , what corroborates the hypothesis that migratory flow depends on the income and the probability of finding a job. Dualistic thinkers consider the informal economy as a marginal pan with no direct link with the formal economy. According to Lewis (1954), the labour market is divided into capitalist agents and subsistence workers. The differences in earnings between the two categories (capitalist and subsistence workers) are explained by skill levels, education, responsibility and prestige. For Harris and Todaro (1970) the three market situations (employment in the rural/urban and agricultural, and urban unemployment) are resulting from the imbalance in the labour market from which informality will flow. Both models suggest that informality is a transitory phenomenon that economic growth will reverse.

### The Structuralist Approach

The structuralists design the informal sector as subordinate to the formal economy whose large companies are privileged, because its

objectives are the reduction of the production costs and increased productivity, subordinate small firms and informal workers. According to structuralist thought, the informal production relations are strongly linked to the major capitalist companies, this explains this interdependence (between formal and informal sector) that economic growth cannot eliminate. This theory is based on two approaches: informal economy is a leftover of pre-capitalism constituted of marginalized individuals who work in exploitative condition to ensure their survival (Chen et al., 2004; Williams and Round, 2008) and informal employment consolidates, rather than reduces, the disparities produced by formal capitalist employment (Pahl, 1984; Williams and Windebank, 2003).

### **Theory of Institutionalism**

This theory is constituted by two approaches that taking source in the theory of extra-legality which explains informal activities as a survival strategy to face State constraints. This approach designs the informal economy as microentrepreneurs response to State bureaucratic regulation. The main theorists of this approach are de Soto (1989) and Feige (1990).

One of these approaches considers informality as a new form of activity emerging in late capitalism as a creative response to the State's incapacity to satisfy the basic needs of the impoverished masses' by the deregulated open world economy (de Soto, 1989; Hudson, 2004). The other approach states that supporting transformations and costs of different transactions of other economic agents and degrading conditions of work such low-paid and sweat-shop like conditions characterized informal employees (Feige, 1990; Sassen, 1997).

#### **De Soto's Approach**

According to de Soto (1989), the State is sole responsible for the marginalization of the masses by giving business groups certain privileges and instituting ineffective judiciary processes. Excessive market regulations imply very high legalization costs. Governments take benefit of the majority of the privileges. Small interest groups override laws to exercise entrepreneurship and ensure their survival. De Soto model was the first to implicitly defend the idea of segmentation of the informal market, arguing that informal is twofold, street vendors and informal markets themselves.

#### **Feige Approach**

According to Feige (1990), the fact of supporting transformations and costs of different transactions of other economic agents was the factor that distinguished the informal workers of the formal agents.

These transformations and transaction costs related to formal and informal sectors explain the importance of these sectors in the development process, which are the foundations of the institutionalist approach. By focusing on the role of the State in the resource allocation process, the proponents of this school emphasized the partial aspect of the neoclassical analysis framework. According to Feige (1990) there are four distinct types of informality, which are illegal activities, unreported activities, unregistered activities and informal activities themselves.

### **The Unifying Theory**

Since none of the schools of thought has managed to develop a framework where the informal can find its full expression in its own economic meaning or expression of its statistical measurability, so, a fourth school of thought was born. It advocates an approach called unifying approach that combines the previous three currents theories and thereby finishes with the cleavage between the three main thoughts. This twofold school trend, proposes initially to see the informal labour market as a labour market in multiple segments and secondly to see the informal economy as a system with a dynamic of its own.

#### **Integrated approach**

There are, according to Fields (1975, 1990), a cleavage of the economies of developing countries in the formal and informal sectors. The formal sector is modern, urban and covered while the informal sector is traditional, agricultural and non-covered. The migration sends rural informal workers to urban areas. This movement promotes the development of the urban informal sector which in itself is a two-tiered sector. But this informal sector is heterogeneous, to the extent that it receives migrants without experience or specific qualification and former employees of the formal sector with experience, capital and skills. Hence there is a cluster of informal employment, where the entrance is free and individual workers perceive low incomes. Formal work is a dream as it is desired. There is another cluster in the informal market where entry is limited, incomes are high and it is preferred to formal employment.

#### **Dynamic approach**

Beyond the segmentation of the informal market, according to the dynamic approach, there are interrelations between the formal and informal sector. These interrelations will power the segmentation phenomenon. This segmentation is not only a dual segmentation. This is a multiple segmentation through which agents go from one

sector to another according to their intrinsic characteristics and motivations. Based on the Lopez model (1989) this approach designs an inter-sectoral mobility model. The dynamic approach considers various possibilities of movement in inter-sectoral mobility, to better understand this approach, we first present the model of Lopez (1989).

According to Lopez (1989) there is mobility between the informal sector and the formal sector. In this mobility, the informal sector is upstream through its upper segment and downstream through its lower segment. This process is described in three stages in the model. First, when the workers are young, unskilled and inexperienced. Second, when some workers are experienced and have certain skills. Finally, when employees reach a certain age, with social capital and some technical knowledge.

The Lopez model, argues firstly that the informal sector has more than two segments (Fields, 2005; Chen, 2005) and secondly, there are several types of mobility, most of them are not one-way even if everyone is not concerned (OECD, 2009). Thus the informal economy comprises various segments, which are cut in multiple social networks with different types of agents. An upper segment consists of sub-segments which are constituted of microentrepreneurs, and self-employed agents. The unpaid workers in family businesses constitute an intermediate segment. Lastly, a lower segment gathers households that are struggling to survive and industrial subcontracting workers.

All segments are not open to all agents, access to different segments and the passage from a segment to another is a function of human and social capital accumulation namely education, motivation, vocational experience and skills. This passage also depends on the socio-economic characteristics of the country (macroeconomic environment, legal framework, business regulations, etc.), and on the importance of social networks. From there three types of mobility should be distinguished on the informal market, job mobility, spatial mobility, and mobility of individual income (OECD, 2009).

The unifying approach advocates a synthesis of three previous currents theories. It designs the informal labour market as a set of segments with a dynamic mobility based on a multiple segmentation of the labour market and therefore with interaction with the formal sector.

### **1.3 Macroeconomic situation and informality growth in Haiti 1980-2010**

We cannot analyse informality growth without a deep understanding of its macroeconomic context. Between 1985 and 2010 the economy was subject to some deepened reforms that can be resumed into the



application of two structural adjustment plans 1986-1989 and 1996-1999. We develop the main elements that have led the macroeconomic situation during the period 1980-2010.

### **1.3.1 Macroeconomic outlook**

From 1990 to 1994 the agricultural sector lost 32% of its plus-value (Despinas, 2008). The part of the agricultural products in exportations dropped from 30% in 1985 to 9% of the GDP in 2001 (IMF, 2002). The deficit in the balance of trade in goods and non-factor services grew from 469.3 million of dollars of the United States in 1996 to 687.3 million of dollars of the United States in 2000 (IMF, 2002). In 2010, an earthquake caused damage and losses caused estimated at 7.863 billion US dollars (PNUD et al., 2010), which represent more than 120% of gross domestic product because 1 dollar was worth 42 gourdes at this time.

#### **Liberalization of the economy**

Different important moments have marked the Haitian economy from 1980 to 2010. The stagnation period of the economy from 1984 to 1990, which we can divide in two sub-periods. The period of adjustment without funding causing the recession of 1982-1985 and the stagnation from 1986 to 1989 and resumed in 1990. The sudden decline in exports and the meteoric rise in imports between 1985 and 2011 and the demo-economic imbalances (the average real GDP growth of 0.13% between 1988-1991 and 2007-2011 (Montas, 2012), while the population growth is about 2.5% during the decade 1988-1998 (IHSI, 2005)) are the dominant macroeconomic facts of the period. The low GDP growth, the poverty due to the constant growth of unemployment and the rapidly increasing of the urban population in the early 80s paved the way for economic liberalization mechanisms. Since 1986 several reforms were implemented: reform of trade policy to liberalize trade who understood such measures as eliminating licensing requirements for imports of sugar and rice, the removal of rights imports and quotas; fiscal and monetary policies to restore and preserve the balance of public finances and balance of payments, but also a stabilization of prices and the dollar. Thus the tariffs have been reduced to 5-10% (UNDP, 2003), but also there was the banking reforms of all kinds including the liberalization of the interest rates. These measures included privatization of public enterprises producing goods and services, reform of public administration based on the downsizing, 130,000 public employees were dismissed in four months in 1994 (Lamaute-Brisson, 2005). Despite these reforms, there were no regulations of the goods and services markets and the credit market has not been touched.

### **Growth of self-employment and wage decline**

The labour market experienced a drop of employees between 1986 and 1999. This was explained by the downsizing of the public enterprises applied by the PAS (1986-1989 and 1996-1999) and an important departure of the industries between 1991 and 1994. The employment payroll in manufacturing was generally focused trend towards growth between 1983 and 1984, there has been an attrition of about 8% over the period 1987-1989 (Lamaute-Brisson, 2005), which was a significant decline in the industry's workforce. This sharp decline of the workforce in manufacturing was followed by a slow decrease from 1987 to 1989 (Lamaute-Brisson, 2005). In 1991, about 15% of the employees disappeared in the enterprises records. In 1992 approximately 40% of employees from bars, restaurants and hotels and about 28% in manufacturing have been disappeared in the records (Lamaute-Brisson, 2005). Meanwhile the share of self-employment increased from 27.4% in 1982 to 42.32% in 1986.

In 1990, urban informal employment represented 45% of total employment (Montas, 1995), in 1999-2000, 54.7% of the jobs are informal (IHSI, 2001a, 2001b) and 57.1% of jobs are informal in 2007 (IHSI, 2010). The unemployment strongly hits young people especially those who are relatively more educated (Montas, 2005). The participation rate is about 61% today in 2020 there will be 5,545,000 active on the market with a participation rate of 65.3%. The acceleration of external migration is a new model of survival for the economy. The households receive in average about 1.4 billion US dollars of transfers between 2007 and 2011 from Haitian migrants living abroad (Montas, 2012). Nevertheless, the structure of the branches of the informal underwent modifications. There are fewer and fewer craftsmen and more and more merchants and people turning to the transport sector, but also a lot of self-employed (88%), among them 63.4% are found in the trade sector, 51% of workers in the "manufacturing" sector are self-employed (Lamaute-Brisson, 2005).

### **1.3.2 Informality and GDP growth: a countercyclical relationship**

It seems, a priori, that there is a countercyclical link between informal and GDP growth. For a better understanding of the inter-relationships between informal and GDP growth, we highlight the macroeconomic factors behind the expansion of the informal and informal analyses the impacts on GDP growth.

#### **Macroeconomic factors of informal expansion**

After the loss of salaries for many workers and the sharp fall in industrial employment between 1986 and 1987 and in 1991-1992, different macroeconomic changes occur. Among them we can cite,

extension of self-employment, low growth in industrial employment, low rate of commitment in the employment system only 46.9% of people of working age are active, (IHSI, 2001), under-employment and the persistence of particularly high unemployment rate in the metropolitan area. Those are the prime factors that propelled informal activities.

In 1986, for the first time since 1971 the share of self-employed starts to grow, reaching its highest level since 1950 (Montas, 1995, 2012). From 27.4% in 1982 the proportion of self-employed shifted to 42.32% in 1986 (Montas, 1995, 2012). This observation was one of the first effects regarding the loss of wage earners and deindustrialization. Meanwhile the share of traders in the informal workers increased from 43.5% to 45.3% (Lamaute-Brisson, 2005). This is explained by the imports increase in the total supply volume. The highly concentrated oligopolistic banking sector finance the largest segments of the import market. While 15% of employees have lost records in 1991-1992, the share of informal employment in total employment has risen steadily since.

The growth of the private sector accompanied that of the informal economy: 54.8% of employed persons are working in the informal private sector. Informal private companies represent 81% of the metropolitan area of business (IHSI, 2010). Furthermore, there is no real boundary between the informal private sector and the formal private sector which provides only 1.9% of total employment. Weak employment growth, high unemployment, persistent underemployment, the narrowness of the private sector, and bank credit orientation are all factors that contributed to the growth of the informal economy.

### **Links between informality and GDP growth**

The impact of informality on GDP has not been a subject of empirical studies. But a first observation can already be made from the respective past results they have shown in the economy.

From table 1.1, during the period from 1971 to 1982, informal employment fell by an average of 0.5416%, while GDP has grown on average by 5.85%. As for 1983-1987 period, informal employment rose slightly by 2.98%, while GDP in real terms had a negative average growth of 1.25%. In 1988-1991, urban informal employment increased by 0.67%, while the GDP had a positive average growth of 0.13%. Informal employment has increased on average 1.215% and GDP shows a negative growth of 0.1321% over the period 1992-1999. In 2000-2007, the average increase of the informal economy is 0.2125% and the average GDP growth of 0.4762%. We find that any increase in informality over 1% leads to a negative GDP growth and vice versa, but it should also be noted that when the increase in informality is between 0 and 1% the same applies to the GDP growth. During the periods 1983-1987 and 1992-1999 the increase in informality causes

## 1.4. Modelling and measurement methods of the informal economy

a negative GDP growth, whereas the decline in informality induces a positive GDP growth for the period 1971-1982. It seems that the informal economy has a countercyclical impact on GDP.

TABLE 1.1: Comparison of GDP growth and increase in the informal economy (in %)

	1971-1982	1983-1987	1988-1991	1992-1999	2000-2007
PIB growth %	5.85	-1.25	0.13	-0.1321	0.4762
Informal increase	-0.5416	2.98	0.67	1.215	0.2125

Source: Computations by the author from Montas (1995, 2002) and IHSI (2007).

However, to confirm these observations we need measurement tool subjected to the rigor and econometric and statistical accuracy, therefore we build the following econometric model that allows us to better grasp the desired object.

## 1.4 Modelling and measurement methods of the informal economy

### 1.4.1 Measurement methods

Several methods have been developed for a better understanding of the informal economy. Some of them are rather based on microeconomic approaches, these methods are what experts call the direct approaches. Others are based on various sets of macroeconomic indicators and belong to what might be called indirect approaches.

#### Direct methods

Direct microeconomic approaches using different procedures that are generally grouped as follows.

The method of questionnaire surveys is based on voluntary responses from households, businesses or both (Tanzi, 1983). The household surveys to identify multiple jobs and home-based work home. Surveys of small businesses allow identification of business characteristics, income and many aspects of the production. Finally, the mixed surveys - households and businesses - identify the socioeconomic characteristics of informal workers and the segmentation of the informal labour market.

The major drawback of this method is that the results of questionnaire-based surveys are too subjective (Mogensen et al. 1985). They depend on the wording of the questionnaires in the first place, and they also depend on the answers of the interviewees, it turns out that interviewed may struggle to declare fraudulent behaviour.

The method of tax audits focuses on controls of taxes paid to the tax authorities that estimate the informal economy based on the

differences between the statements of income subject to taxes and those subject to special checks (Feige, 1986). The weakness of this method is that it does not take into account undeclared work as the special checks do not detect all fraud and false statements in the existing system.

The method of implicit labour supply requires also the use of available information in the formal and informal labour market in order to estimate the size of the informal economy where the method provides implicit work. The method focuses on comparative studies of the labour market also has weaknesses such as the possibility of underestimating the total number of people working in the informal economy, it does not account for illegal activities such as drug trafficking, gambling and prostitution and finally it does not take into account the persons who hold more than one job, because according to the literature on informality, in developing countries many people who are working in the informal economy holding a second job (Adam and Ginsburgh, 1985; Thiessen, 1997; Jensen and Slack, 2009; Gunther and Launov, 2011).

The main benefit of direct approaches is to provide detailed information about the structure of the informal economy. Among the derivatives of these approaches, questionnaire-survey remains the most used method.

### **Indirect Methods**

The method based on discrepancies between expenditures and household income takes into account the household budget and consumption surveys to make a comparative analysis between income and household consumption. According to National accounts aggregate income (GDP) must be equal to aggregate expenditure. First, there should be an independent measurement of spending upon a representative sample of the population gauging household consumption over a specified period while deducting savings from their income. Second, the difference between this measurement and aggregate expenditure (GDP) then comes from activities other than those that are reported and it can serve as an indicator for measuring the share of the informal economy (Park, 1979). Nevertheless, the weakness of this method is that it does not take into account undeclared work. Hence, adjusting income tax does not apply to non-registered job.

The method of discrepancies on the labour market and the implicit labour supply. Official turnout of the population in the national workforce are not always compatible with the actual data on the labour that evolves in the economy, this often observed gap between the official rate and the actual rate of participation can be considered as an indicator to measure the size of the informal economy into the global economy. A decrease in the labour force participation rate in the formal economy implies that the workforce increased in the

informal economy (O'Neill, 1983) and this difference is the indicator used by that method to measure the size of the informal economy.

The limitations of this method relies on the fact that the gap is not necessarily reliable because some people work in the formal and informal market at the same time; the gap in official work may have other causes as increased participation in the informal labour market.

The method of physical inputs includes two empirical models (Kaufmann and Kaliberda 1996; Lackó, 1999) considering that electricity consumption is the best indicator of economic activity as a whole, that is to say both legal and informal activities. Both models show the existence of an elasticity close to one between the Gross National Product (GNP) and electricity consumption.

However, this method can be criticized because informal activities are more labour intensive, or will use other sources of energy; hence, underestimating the informal economy. With the exception of models based on demand for cash currency, the indirect approach methods are used less and less, at a time when econometric methods appear to be the safest and most robust modelling tool in research on the informal economy. The method of monetary transactions come from the quantity theory of money and the validity of the Fisher quantity equation. The equation states  $M_1V + MV = PT$ , where  $M$  and  $M_1$  denote money supply in cash and deposits,  $V$  and  $V_1$  velocity of circulation of cash and deposits,  $PT$  product of total monetary transactions  $T$ , the general price level  $P$ . The approach developed by Feige (1979) is based on the assumption that the total GNP is traded in currency and that the relationship GNP-quantity of money is a constant. The measurement of Gross National Product (GNP) including both the official GNP and underground GNP can be determined due to the fact that the money supply which is composed of cash and deposits is easily observable.

Among many limitations of this method, we note in particular that it gave a negative measurement of the informal economy for the whole period from 1939 to 1968. The main hypothesis implies a constant relationship between the total value of monetary transactions and economic activity, which deserves to be elucidated by econometric tests. Finally, the assumption of the existence of a reference year when the informal economy is considered null is unrealistic.

The method of currency demand assumes that illegal or underground transactions are settled in cash to escape control it holds that the velocity of circulation of cash currency ( $M_0$ ) in the informal economy is the same as for the formal economy and that the base year is chosen arbitrarily (Adair, 2002). This method contains two main models.

The Cagan model first tried to establish a correlation between the demand for money and the tax burden (Cagan, 1958). Gutmann (1977) provided an extension by studying the ratio of "normal" demand for money and the actual demand for money, which is then interpreted

as an indicator of the size of the informal economy. Therefore, any increase in the size of the underground economy will cause an increase in the cash/deposit ratio in the banking institutions.

Tanzi (1983) published one of the first modelling methods built on the same basis as the demand for money method. According to the Tanzi model, a growing informal economy leads to a growth in cash money demand, wherein the dependent variable is the cash/deposits ratio ( $C/M_2$ ). To isolate excess demand for cash currency, an equation of the cash money demand is estimated. In order to identify the variables considered, all conventional factors are controlled, such as changes in income, consumer spending, savings and interest rates. Some other factors likely to influence economic agents to operate in the informal economy are included, such as direct and indirect taxes, State regulations and the complexity of the tax system.

Nevertheless, these two models have some limitations regarding their assumptions, and fall short of any theory designed for this purpose. First of all, transactions in the informal economy are not made entirely in cash (Isachsen and Strom, 1985). Second, the Tanzi model has poor specification, wherein only taxes and government regulations encourage economic agents to turn informal, ignoring other important factors can also come into play. Hill and Kabir (1996) contradict the assumption that the money circulation rate is identical in all sectors of the economy. This means an increasing in money demand is not necessarily a cause of growth in the informal economy (Frey and Pommerehne, 1984). Last, the assumption that the informal economy was negligible during a reference period is unrealistic.

### **1.4.2 Modelling methodology**

It is often criticized the methods of the indirect approach taking account of a single variable as an indicator of the weight of the informal economy. Indeed, such methods prevent to understand all the effects of informality. Now it is well established that the informal economy has impacts both the work on the money market but also on production, therefore more element to be taken into account for the well identified. Thus the modelling approach takes into account the causes and effects of informality through two groups of variables classified into causes of variables and variables "indicators". It is a method based on the statistical theory of unobserved variables or latent variables (Schneider, 2010) which takes into account several variables as causes and others as indicators of the phenomenon (Bajada and Schneider, 2007).

If economic theory justifies the choice of the variables used in modelling, however the consideration of the informal economy as a latent variable that it appears as a fixed value over time, is neither explained nor justified by theory still less by economic statistical theory.

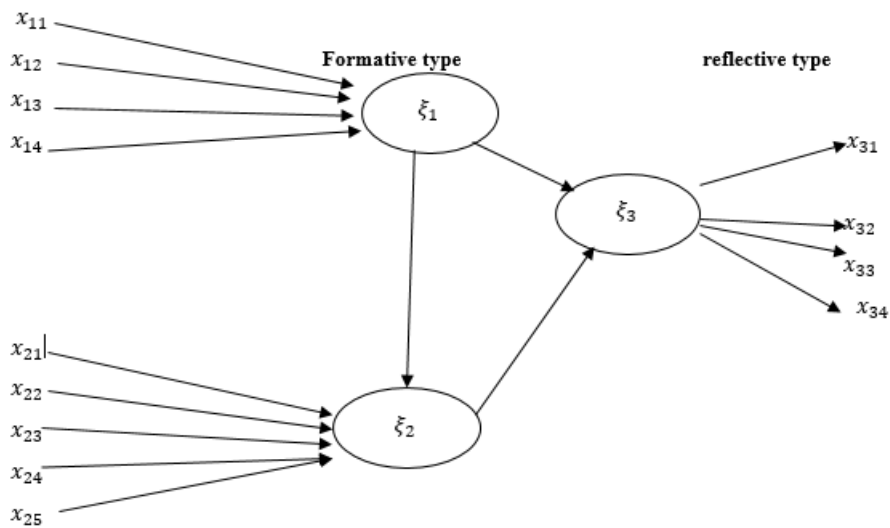
**Theoretical structural equation model with latent variables: the PLS approach**

Partial Least Squares Path Models (PLS) are latent variable models that treat  $p$  observed variables on  $n$  individuals divided into  $j$  blocks of  $k_j$  continuous variables. A structural model of relationships between variables interconnects the blocks. Variables are two types, the observed variables called manifest variables (MV) and unobserved variables called latent variables (LV) that exist through the manifest variables. To each variable block  $k_j$  we associate one latent variable. This model is composed of two sub-models:

$$\begin{aligned} X_{jh} &= \pi_{jh}\xi_j + \epsilon_{jh} && \text{(Measurement model)} \\ \xi_j &= \sum_i \beta_{ji}\xi_i + \nu_j && \text{(Structural model)} \end{aligned} \tag{1.1}$$

The measurement model links the MV and LV, it may be reflective (MV reflect their LV) or formative (LV reflects the VM associated with it) the structural model links the VL among them (figure 1.1).

FIGURE 1.1: The PLS Model



Source: Representation of the author

Two approaches can be used to estimate the parameters of a such model. The first one is the LISREL approach (linear structural relationships) which serves to test hypotheses about the relationship between observed and latent variables. In LISREL approach two sub-models can only be reflective (MV reflect their LV) this means the latent variable is a factor that we ignore the causes. The second approach is the PLS (Partial Least Squares) model. It measures the latent variables based on the manifest variables. In this case the two sub-models can be totally reflective, or the two sub-models can be mixed also this means they can be reflective and formative (LV reflects MV associated with it). Hence, latent variable is taken as a factor that we know the causes and effects are palpable. In this chapter the PLS approach is the most appropriate. This model does not require the assumptions of



normality, due to regressions, over the issue of independence between variables does not arise (Jakobowicz, 2007).

### Assumptions and estimation method

This model is valid only under the two following conditions: observations must be independent (they can be multi-levels) and variable blocks should be one-dimensional (in the reflective case).  $\epsilon_{jh}$  is an impact matrix where each  $\epsilon_{jh}$  is white noise, they are uncorrelated with the latent variable  $\xi_j$ .  $\nu_t$  represents the unexplained component is a column vector of terms of mean zero error uncorrelated with the manifest variables  $x_{jh}$ .

Latent variables are estimated in two different ways. After fixation with initial external weights

$$c_{ij} = c_{ji} = \begin{cases} +1 & \text{if } \text{cor}(\xi_i, \xi_j) > 0 \\ -1 & \text{if } \text{cor}(\xi_i, \xi_j) < 0 \\ 0 & \text{if } \text{cor}(\xi_i, \xi_j) = 0 \end{cases} \quad (1.2)$$

The external estimate from the manifest variables  $x_{jh}$  of its block  $Y_j = \sum_{h=1}^{k_j} w_{jh}x_{jh} = X_j w_j$  where  $Y_j$  must be centred and reduced,  $w_j$  a column-vector of the coefficients  $w_{jh}$ . The inner estimate  $Z_j$  form the external estimates  $Y_j$  of the standardized latent variables  $\xi_i$  related to  $\xi_j$  are defined by:  $Z_j \propto \sum_{\{i \setminus i \neq j, c_{ij} \neq 0\}} e_{ji} Y_i$ .

Where  $\propto$  means that the variable on the left is obtained by the reduction of the variable on the right. Where  $e_{ji}$  are the multiple regression coefficients of  $Y_j$  on  $Y_i$  because  $\xi_i$  explains  $\xi_j$ .

By connecting external estimates  $Y_j$  with inner estimates  $Z_j$  we obtain:

$$Y_j \propto \sum_h \text{cor}(x_{jh}, Z_j) x_{jh} \equiv Y_j \propto X_j X_j' Z_j$$

Wold (1985) gets the stationarity condition of  $Y_j$  variable sub-model (measurement) of reflective form  $Y_j \propto X_j X_j' \sum_{\{i \setminus i \neq j, c_{ij} \neq 0\}} e_{ji} Y_i$ . In the model of formative type the variable  $Y_j$  is obtained by the multiple regression of  $Z_j$  on the columns  $x_{jh}$  of the matrix  $X_j$  then reduction  $Y_j \propto X_j (X_j' X_j)^{-1} X_j' Z_j$ .

Hence the condition of stationarity of a variable  $Y_j$  in the sub-model (formative type) is:  $Y_j \propto X_j (X_j' X_j)^{-1} X_j' \sum_{\{i \setminus i \neq j, c_{ij} \neq 0\}} e_{ji} Y_i$ .

It repeats the last two steps until the method converges, and then estimated structural equations by multiple regression method. The residue normality assumption is not necessary, because of the regressions.

### 1.4.3 Theoretical justification of the economic model and variable use

It is essential to know the causes and consequences of the informal economy into the economy in general to better define in economic theory. Thus the variables that must be used to model and estimate the

weight of informality must be variables of causes and consequences (indicators) taken as needed for the proposed analysis and the scope of the existing theory.

One of the classic features of the informal economy is that it is exercised outside the fiscal and administrative frameworks regulated; it is also one of the factors supporting its expansion and sustainability (Castells and Portes 1989; Maldonado, 1995). According to Maldonado (1995), one of the comparative advantages of the informal sector resulting in reduced hourly costs which itself is due to non-payment of taxes, non-payment of social security contributions, tax evasion and concealment of income (Adair, 2009), which themselves depend on the institutional context.

The second view contains theories that use to link informal activities and corruption with regard to the economic situation: high income economies or low income countries (Schneider, 2007). Some models describe corruption and informality as complements such as a larger State corruption practice leads to a larger informal economy size (Johnson, Kaufmann and Zoido-Lobaton, 1998; Hindriks et al., 1999). Some other theories see informal activities and official corruption as substitutes, the larger is the former the smaller is the latter (Choi and Thum, 2004; Dreher, Kotsogiannis and McCorriston, 2005). The consideration of these elements is based on two empirical approaches that are the positive elasticity of labour supply with respect to the tax rates and the risks of subjective estimation and gains from tax evasion. The third approach is built on the model of Sandmo (1972), which highlights the idea of maximizing expected utility function of the fraudster whose arguments income and job settings and the wage rate, the rate employment, the probability of being caught and the penalties accompanying the possible detection.

### **Macroeconomic and Microeconomic foundations**

The first approach is explained by the importance of microeconomic foundations of the tax burden theorized by the Laffer curve, built on the assumption that increasing the tax burden induced a favourable arbitration preferably undeclared work replacing the institutionalized job or supplementing (Adair, 1985 in Adair, 2009). Thus, it is justified to use in modelling to estimate the informal economy variables such as taxes (direct and indirect) social contributions, control of corruption and quality control of public institutions and quantity of money in the economy. Taxes and social contributions are used for analysing induced arbitration in favour of informal employment, the framework of functioning of State institutions (quality control) that could explain tax evasion and fraud, the quantity of money circulating in the economy which is controlled by the regulatory bodies (Central bank and tax institution) to identify hidden income.

The second approach is built on the model of Sandmo (Allingham and Sandmo, 1972), which highlights the idea of maximizing expected utility function of the fraudster. It takes into account the different variables. These variables can be: GDP or GDP per capita) or the growth of income per capita in percentage, employment which may be reflected in the growth rate of labour or unemployment parameters, the wage rate and related variables such as the inflation rate or the number of working hours, the probability of being caught and the penalties accompanying potential detection. So many variables can incorporate modelling to estimate the weight of the informal economy.

### **Justification of econometric modelling**

Direct methods (tax audit implicit labour supply) or indirect methods (gaps, physical inputs, monetary transactions) have the disadvantage of not taking into account a single variable as an indicator of the weight of the informal economy, which prevents to understand all the effects of informality. In as much as the informal economy has an impact on all segments of the economy, so several elements must be taken into account to better identify this issue, hence the use of econometric modelling.

Because informality is not directly observable in macroeconomic aggregates while it is present throughout the economic cycle, the latent variables models are those that best reflect this unobservable characteristic. The choice of a particular model MIMIC i.e. the PLS meets two requirements. First the development of informality is influenced by several macro factors as we have highlighted them above. These factors constitute the causes of the evolution of informality. They form with a latent variable that the PLS model called the formative kind. The informal economy in turn acts on a set of other macroeconomic factors, called indicators of the impact of informality, the PLS modelling binds with another latent variable to form the reflective type.

PLS model, unlike other derivatives of MIMIC models, does not consider the informal economy as the only unobserved factor of the economy. PLS model supposes the existence of other unobservable factors in the economy that will themselves act on the unobservable informality as the fundamental element. This explains the existence of the structural model (internal) in PLS modelling, which is the main originality of this model. PLS is different from space-state type model, which gives the general form of all MIMIC models. In the structural model other latent variables are used as predictors of the main latent variable. It is probably the most flexible extensions of multivariate multiple linear regression model (Jakobowicz, 2007)., allowing it to be used when the number of observations is less than or equal to the number of variables. This is exactly the case here. PLS components produced by linear combination of the initial predictors, so that

these components used in the predictive model are not correlated. This model does not require the assumptions of normality, due to regressions (Jakobowicz, 2007).

The PLS models have some limitations. These models are valid only under specific conditions. One must be careful in building blocks for the convergence, which has not been demonstrated for two blocks, but it is found in practice. Finally, PLS underestimates the parameters of the structural model.

## 1.5 Design and Interpretation of the Model: Variables, Equations, Estimates and Results

Our PLS model (MIMIC) is designed and processed by a method with the software called PLS-PM.

### 1.5.1 Data, variables and modelling

We designed a set of eleven variables (annual series) spread over the period 2000-2010. This period of eleven years would have been a problem but PLS may be used when the number of observations is less than or equal to the number of variables without significantly altering the significance of the results. All data are from the base of the World Bank (2012b), with the exception of those where sources are mentioned.

#### Data

First we have the variables constituting the formative sub-model:

*Tax*: taxes paid including social contributions paid by the employer with the exception of consumption taxes as a percentage of GDP (source: CEPAL, ECLAC).

*GE*: Government expenditure as a percentage of GDP as a proxy variable indirect taxes as a percentage of GDP.

*IG*: index of perception of the quality of public services, the independence of these services regardless of political pressures, the quality of the development and implementation of public policies and the credibility given to the government to meet its commitments to comply with these policies, between -2.5= bad and 2.5 = excellent.

*CC*: index of State corruption, it measures the perception of the use of power for the purpose of illicit personal enrichment, between -2.5= bad and 2.5 = excellent. Both indices show the predisposition of the state to encourage formal business provided the former is high and the second low.

*logGDP*: GDP per capita (taken in logarithm) at constant prices in local currency (gourdes)

*Inf*: inflation These two variables reflect the ability of households to purchase baskets of goods, or otherwise, they are turning to the informal economy hence their inclusion.

*GDP*: GDP growth rate *per capita*

*UR*: Proportion of unemployed as a percentage of the total population (source: CIA world factbook)

*RW<sub>f</sub>*: labour force participation rate of work as a percentage of total population Then we have the variables constituting the reflective sub-model:

*logW<sub>f</sub>*: available labour force (in logarithm) as a percentage of total population  $M_1/M_2$  :  $M_1$  represents coins, notes and demand deposits,  $M_2$  represent  $M_1$  added to deposits with a lower maturity of up to two years and carry a lower payment or notice of up to three months, this is a proxy of cash/deposits ratio due to Gutmann (source: Bank of the Republic of Haiti).

The importance of both of these variables (CC and IG) resulted in their variance we have left them as they are, and the other eight variables were standardized according to the wish of the model used, as some expressed percentages and other quantities. Our PLS model consists of a formative sub-model (the latent variable reflects the manifest variables associated with it) and a sub-reflective model (the manifest variables reflect their latent variable). Our model is built with three latent variables, two of which are exogenous "Reality" and "Perception" and respectively contain four and five manifest variables, they are indicators in the formative type and one is endogenous ("ShE") which represents informality, it is the event variable in the reflective type and has two manifest variables.

## Modelling

Both model equations sets are written as follows: (Measurement model):

$$X_{jh} = \pi_{j0} + \pi_{jh}\xi'_j + \epsilon_{jh}$$

With the following hypothesis  $E(X_{jh}|\xi_j) = \pi_{j0} + \pi_{jh}\xi'_j$

Which implies that  $\epsilon_{jh}$  has a zero mean and is not correlated with  $\epsilon_j$ .

where:  $X_{jh}$  is a matrix with rows containing the manifest variables of each latent variable and as many columns as the highest number of manifest variables that contain each latent variable, here it is a matrix  $(3 \times 5)$ .  $\pi_{jh}$  is a matrix of coefficients representing the intensity of the expected changes in each group of manifest variables associated with a latent variable to a one unit change in the latent variable in question is also a matrix  $(3 \times 5)$ ;  $\epsilon_{jh}$  is an impact matrix where each  $\epsilon_{jh}, j = 1$  to 3, is white noise, it is still a matrix  $(3 \times 5)$ ;  $\xi_j$  is a column vector composed of the different latent variables. A more elaborate writing of this model can be:

$$\begin{aligned}
 x_{1h} &= \pi_{1h}\xi'_1 + \epsilon_{1h}, h = 1, \dots, 4 \\
 x_{2h} &= \pi_{2h}\xi'_2 + \epsilon_{2h}, h = 1, \dots, 5 \\
 x_{3h} &= \pi_{3h}\xi'_3 + \epsilon_{3h}, h = 1, 2
 \end{aligned} \tag{1.3}$$

Where:  $x_{1h}$  contains the variables  $x_{11} = \log GDP$ ,  $x_{12} = Inf$ ,  $x_{13} = RW_f$ ,  $x_{14} = UR$  et  $\xi_1 = Real$  ;  $x_{2h}$  is constituted by:  $x_{21} = GE$ ,  $x_{22} = Tax$ ,  $x_{23} = CC$ ,  $x_{24} = IG$ ,  $x_{25} = GGDP$  et  $\xi_2 = Percept$  ;  $x_{3h}$  is constituted by:  $x_{31} = \log W_f$ ,  $x_{32} = M_1/M_2$  et  $\xi_3 = ShE$ .

The Structural model is:  $\xi_j = \sum_i \beta_{ji}\xi_i + \nu_j$

With the predictor specification condition  $E(\xi_j|\xi_1, \dots, \xi_{j-1}) = \sum_i \beta_{ji}\xi_i$ .

$\xi_j$  is a vector already defined in the measurement model column where  $\nu_j$  is an average term of random uncorrelated zero the explanatory latent variables  $\xi_i$  in the right part of the structural equation. Most  $\beta_{ji}$  coefficients are structurally zero and  $\xi_i$  corresponding variable is not in the equation of the structural model. The  $\beta_{jj}$  are the most remarkable cases ( $\beta_{jj} = 0$ ) and the corresponding  $\xi_j$  are absent from the right side of the equation.

A more detailed writing of this model is:  $\xi_3 = \beta_{31}\xi_1 + \beta_{32}\xi_2 + \nu_j$ , car  $\beta_{11} = \beta_{22} = 0$ . Where  $\xi_3$  is the shadow economy.

### Estimates

After running the model, we obtain the following results (table 1.2) for the manifest variables and the indicators regarding their block (the latent variable that they are associated to).

TABLE 1.2: Estimation results

Latent variables	Manifest variables	estimates	t-statistic
$\xi_1$	$RW_f$	0.360	5.114***
	$\log GDP$	0.239	2.813***
	$Inf$	0.198	1.807***
	$UR$	0.368	6.142***
$\xi_2$	$GE$	0.444	5.720***
	$Tax$	0.615	5.411***
	$CC$	0.088	2.209***
	$IG$	0.022	0.806
	$GGDP$	0.013	0.060
Indicators			
$\xi_3$	$M_1/M_2$	0.060	0.257
	$\log W_f$	0.948	3.548***
***=significance level at 5%, an significant if $ t  > 1.65$ .			

Source : Computations by the author

The main equations (including only significant variables) of the measurement model are given by:

$$\begin{aligned}\xi_{1t} &= 0.36RW_t + 0.239\log GDP_t + 0.198Inf_t + 0.368UR_t \\ \xi_{2t} &= 0.444GE_t + 0.615Tax_t + 0.022CC_t\end{aligned}\quad (1.4)$$

TABLE 1.3: Coefficient of the latent variables driving the shadow economy

Path coefficients of $\xi_3$					
Variable latent	Value	SD	t	Pr> t	$f^2$
$\xi_1$	-0.857	0.116	-7.399	0.000	6.843
$\xi_2$	0.144	0.116	1.249	0.247	0.195

Source : Computations by the author

Hence, the main equation of the model is given by:  $\xi_{3t} = -0.857\xi_{1t} + 0.144\xi_{2t}$

To move to the actual estimate, we replaced in the model equation all significant coefficients of the manifest variables (table 1.3) belonging to each latent variable, and then we have:

$$ShE_t = -0.308RW_t - 0.205\log GDP_t - 0.169Inf_t - 0.315UR_t + 0.064GE_t + 0.088Tax_t + 0.012CC_t$$

Using equation above we obtain:  $Sh\tilde{E}_{2007} = 0.6377$ ,  $Sh\tilde{E}_{2008} = 0.6936$ ,  $Sh\tilde{E}_{2009} = 0.6740$  and  $Sh\tilde{E}_{2010} = 0.7825$ .

Following the Dell'Anno and Schneider process (Dell'Anno and Schneider, 2006) this index ( $Sh\tilde{E}_t$ ) must be converted to absolute values of the informal economy, with particular base year is fixed, where the informal economy takes a value for a base year, which is an exogenous estimate. We choose 2007 as the base year and the value 57.1% as the shadow economy size for this base year (Schneider, 2010), we denote it  $ShE_{2007}^*$ . The real size of the informal economy is given by:

$$Sh\tilde{E}_t = \frac{Sh\tilde{E}_t}{Sh\tilde{E}_i} ShE_i^*$$

Where  $i$  is the chosen base year and  $t$  is the given year. Thus we have the following estimates of the size of the informal economy for the years 2008, 2009 and 2010 as a percentage of GDP:

$$\begin{aligned}Sh\tilde{E}_{2008} &= 62.1\% \\ Sh\tilde{E}_{2009} &= 60.35\% \\ Sh\tilde{E}_{2010} &= 70.06\%\end{aligned}\quad (1.5)$$

## 1.5.2 Model validation

To be valid, the results found in the previous section need that the built model should verify some fundamental criteria.

### Block's Unidimensionality and Reliability

The model constructed has three blocks of manifest variables each belonging to a latent variable. The block of variables that constitute the first latent variable  $\xi_1$  does not admit values for the Dillon-Goldstein coefficient and the Cronbach's alpha (Tenenhaus et al., 2005; Jakobowicz, 2007) because it is an exogenous variable. Therefore, we cannot state about its internal consistency and its unidimensionality. But this block has a critique value (0.909) greater than 0.5 that suggests the block has acceptable internal consistency and unidimensionality. The second block of variables that form the second latent variable  $\xi_2$  has a Cronbach's alpha lower than 0.7 (table 1.4) that means the strength of the relationship between variables is not very interesting. But the block has a Dillon-Goldstein rho higher than 0.7 so the block is a one-dimensional (Jakobowicz, 2007). The block of variables that forming the latent variable  $\xi_3$  has a Cronbach  $\alpha$  greater than 0.7 (table 1.4) so there is high internal consistency in this block, and it also has a Dillon-Goldstein coefficient greater than 0.7 that means it is a one-dimensional block. Globally the three blocks have an acceptable consistency and all of the three blocks are one-dimensional because both possess coefficients Dillon-Goldstein higher than 0.7 (Jakobowicz, 2007).

TABLE 1.4: Composite reliability

Latent variable	Dimension	Alpha of Cronbach	Rho of Dillon-Goldstein (ACP)	Critical values	eigenvalues
$\xi_1$	4			0.909	2.538
					0.939
					0.154
					0.006
$\xi_2$	5	0.634	0.823	0.558	1.650
					0.852
					0.259
					0.024
					0.004
$\xi_3$	2	0.924	0.964	0.909	1.690
					0.128

Source : Computations by the author

The variables that constitute the block of the latent variable  $\xi_1$  has the first eigenvalue greater than 1 and the second smaller than 1 so this block is reliable (Jakobowicz, 2007; Tenenhaus et al., 2005), more the gap between high the first eigenvalues of each block and those that follow serve to corroborate this finding. All the two other blocks that constitute respectively the two latent variables  $\xi_2$  and  $\xi_3$  have the same eigenvalues characteristics (first eigenvalue greater than 1 and



the others smaller than 1) hence we conclude that the composite of the three blocks of the model is reliable.

Further analysis of cross-loadings values (table 1.5) that compare the values displayed by the manifest variables in their block relative to those displayed in other blocks, if their value in their respective block is superior to other values displayed elsewhere then they were assigned to the best block. Thus, this analysis shows that each of the model variables was assigned to the best block of variables as possible which helps to reinforce the fact that the model is well fitted to the data. The three blocks have all the first eigenvalue greater than 1 and the second smaller than 1 so they are unidimensional (Tenenhaus et al., 2005; Jakobowicz, 2007).

TABLE 1.5: Relationship between manifest variables and latent variables by block

	Cross-loadings			Communalities
	$\xi_1$	$\xi_2$	$\xi_3$	
RW	0.977	-0.909	-0.991	0.954
logGDP	0.782	-0.499	-0.763	0.612
Inf	0.487	-0.624	-0.421	0.237
UR	0.990	-0.944	-0.997	0.980
GE	-0.696	0.899	0.711	0.807
Tax	-0.968	0.950	0.979	0.902
CC	-0.625	0.678	0.619	0.460
IG	-0.239	0.356	0.276	0.127
GGDP	-0.004	0.193	0.037	0.037
M1/M2	-0.865	0.816	0.875	0.765
logWF	-0.989	0.931	1.000	0.999

Source : Computations by the author

For all these reasons all the three blocks are essentially unidimensional and according to the fact that correlations between the manifest variables and the first principal component are positive (appendix 1) none of the manifest variables should be removed from their belonging block in the measurement model (Tenenhaus et al., 2005). There are as many component factors as the block contains variables.

### Global Model and sub-models Quality

The goodness of fit of the global PLS model is given by the GoF (Goodness of Fit) index (Tenenhaus et al., 2005; Jakobowicz, 2007). The values of  $f^2$  (table 1.3) correspond to the indices of cross-validation redundancy (Tenenhaus, 2004; Tenenhaus et al., 2005). These indices related to the endogenous blocks measure the global quality of the structural model. When all the cross-validation redundancy indices are positive, which is the case in our model, the measurement model

has a good quality (Tenenhaus et al., 2005). The cross-validation redundancy indices show the capacity of the path model to predict, using the cross-validation, the endogenous manifest variables indirectly from a prediction of their own latent variables through the related structural relation (Tenenhaus et al., 2005).

For each block the Communalities index (Tenenhaus et al., 2005; Jakobowicz, 2007) measures the quality of the measurement model, and it was found that this model is of good quality because the communality index is 0.625 in average (table 1.6). For each endogenous block the redundancy index (Tenenhaus et al., 2005; Jakobowicz, 2007) is used to measure the quality of the structural model, the quality of the model is interesting.

TABLE 1.6: Coefficient of communalities and redundancies of the model

Variable	Type	$R^2$	adjusted $R^2$	Communalities	Redundancies	Rho D.G
$\xi_1$	Exogenous			0.696		0.896
$\xi_2$	Endogenous	0.843	0.843	0.467	0.393	0.780
$\xi_3$	Endogenous	0.983	0.981	0.882	0.867	0.937
Average		0.913		0.625	0.630	

Source : Computations by the author

The  $R^2$  for  $\xi_2$  and  $\xi_3$  are greater than 0.84 (table 1.6) hence these sub-models take into account the complexity of the model in a very satisfying way.

Since in our model, as shown in the table 1.7, the GoF indices of internal and external models are very high as well as the relative GoF which reflect a very good fit to the data model. The external model explains 96.3% of the data; the internal model is capable to explain 95.6% of the data. Relatively the model takes into account 92% of the achievable fit. In absolute the model fit the data at 76%. All those results are very significant because in the table 1.7, zero is not included between the lower bound and the upper bounds (Tenenhaus, 2004).

TABLE 1.7: Fit model quality

	GoF	SD	Critical Ratio (CR)	Lower bound (95%)	Upper bound (95%)
Absolute	0.756	0.030	25.044	0.681	0.802
Relative	0.921	0.031	29.944	0.812	0.941
External model	0.963	0.015	65.277	0.906	0.970
Internal model	0.956	0.027	36.010	0.862	0.989

Source : Computations by the author

The Communality for each variable releases the proportion of variance of the manifest variables explained by their latent variable

associated. The communality indices show that only two variables (IG and GDP) are not well explained by their latent variables (table 1.5).

The three blocks are unidimensional and the variables that constitute each of them belong to the best possible block then the three blocks are reliable. The manifest variables are positively correlated to the first principal component. The measurement model has a very good quality. The structural model has a high quality. External and internal models explain more than 95% complexity of the data. According to these observations we conclude that the estimates of the PLS model are valid so the results can be analysed.

### 1.5.3 Results analysis

We analyse the impact on growth in informality of each significant manifest variables of the model: GDP per capita, GDP, inflation, unemployment rate, direct taxes, government expenditures (proxy of indirect taxes), labour force participation rate, corruption index, and the index of perception of the quality of public services. The effects of each variable on the growth of informal activities are compared to theoretical and empirical assumptions in the literature devoted to the informal economy.

In 2008 an increase of the size of the inflation of 4.87% has accompanied the size of the informal economy increased by 5%. An important drop in the unemployment rate by about 2% and a fall of the direct taxes of 0.11% and indirect taxes (approximate by the government expenditures) increased of 0.59% have also accompanied the informal activities increase. This observation is coherent to the idea that informality and taxes evolve in the same direction (Adair, 1985; Schneider, 2005); At the same time the mass of money in circulation increased by 3.14 points, which corroborates the point that additional use of currency is a good indicator of shadow economy (Schneider, 2010). The GDP *per capita* has increased negatively of 0.48% while informality increased. This remark is contrary to the theory that stipulates an increase in informal activities leads to an increase in GDP per capita (Sandmo and Allingham, 1972). The decline in unemployment is due to the fact that part of the unemployed workforce has joined informality, which is coherent to the classic view that advocates a decrease of the workforce participation rate leads to an increase of the informal workforce (Schneider, 2010, O'Neill, 1983). The increase of the mass of currency circulating in the economy of 3.15% leads to an increase of the informal activities. Informality acted in a countercyclical way of the GDP which decreased by 2.5%. We must point out that despite the increase in informal activities the index of State corruption improved by 0.07 points. That explained expanding informal activities contribute to reduce State corruption (Choi and Thum, 2004; Dreher, Kotsogiannis and McCorrison, 2005). Surprisingly the improvement

of the index of perception of the quality of public services of by 0.3 points accompanied the increase of the informal economy, which is contrary to alternative empirical observation (Johnson, Kaufmann and Zoido-Lobato, 1998).

In 2009, the informal economy fell by about 2%, which is explained by the growth of the GDP of 2% and the sudden fall of the inflation. The unemployment rate fell by about 5.6%. That brings us to understand that the decline in unemployment is not just the preserve of the informal economy contrary to the idea stated by O'Neill (1983). It should be noted that the currency in circulation grew of 1.08% despite the decrease in informal activities, this observation reinforces the point that the growth of the mass of currency in circulation is not necessary an effect of growth in informality (Frey and Pommerehne, 1984); however, declining inflation rate by 9.5% influence significantly the decrease of informal activities. There was an increase in public expenditures (proxy of indirect taxes) of 0.33%, an increase in direct taxes of 1.1%. These two observations mitigate the point that taxes (direct or indirect) lead informal economy growth (Adair, 1985; Schneider, 2005). The improvement of the corruption control index of 0.05 points was a part of the decrease of informal activities. That is consistent with the empirical assumption that stipulates smaller informal economy correspond to smaller State bribery (Johnson, Kaufmann and Zoido-Lobato, 1998; Hindriks et al., 1999). Furthermore, we see that the index of perception of the quality of public services has dropped by 0.2 points while informal activities also dropped which is contrary to some empirical observations (Johnson, Kaufmann and Zoido-Lobato, 1998). GDP per capita has increased positively by 2.02% while informality decreased, which is contrary to the view that advocates a similar situation should lead to an increase in GDP per capita (Sandmo and Allingham, 1972). The countercyclical aspect of informal activities is still evidenced in as much as we observe a positive growth of 2.9% GDP induced the fall in the informal economy.

In 2010, the informal economy has increased by about 9.71%; simultaneously there has been a rise in inflation. The taxes have increased by 0.16%, while indirect taxes have fallen down by 0.41%. This observation makes the link between taxes (direct and indirect) and informal economy growth (Adair, 1985; Schneider, 2005) unsettled. Despite a slowdown in the economic activity due to the earthquake, the mass of money in circulation increased by 4.15%, which is consistent with the fact that Schneider (2010) considers additional use of currency as a good indicator of shadow economy. Inflation increased by about 1.33%. There was a worsening of the corruption index of by 0.10 points. This increase in State bribery that accompanied the positive growth of informal economy is consistent with the empirical observation stating the larger is the official corruption practice the larger is the shadow economy (Johnson, Kaufmann and Zoido-Lobato, 1998; Hindriks et al., 1999). The index of perception of the quality

of public services dropped by 0.03 points while informal activities also grew which is consistent with the point stating the greater is the perception of public services quality the smaller is informal economy (Johnson, Kaufmann and Zoido-Lobaton, 1998). The important increase in government expenditures by 1.33% is explained by the destruction of public service buildings in the earthquake. Nevertheless, unemployment fell by 1.4%. The negative growth of the GDP by 5.4% implied a sharp increase in informal activities, which corroborates the countercyclical aspect of the informal economy.

We should observe that between 2007 and 2010, the lower is the growth of annual GDP the higher is the annual growth of informal economy. The link between increase in informality and increase in State corruption is not checked, as we have seen informal increases have been accompanied by an improvement in the index of corruption. There is an inverse relationship between informal activities and the index of perception of the quality of public services. As for inflation, it seems to be moving in the same direction as informal economy. The amount of money in circulation is still increasing but when the size of the informal economy is increasing the growth of money in circulation is more significant. Taxes (direct) are growing regardless informality grows or falls. GDP *per capita* evolves inversely to informality. The unemployment rate and the labour supply decrease regardless of the informal movement, but they act very significantly on the informal economy. Except in 2010, government expenditures operate inversely to the trend of informality. The variables that significantly influence the development of the informal economy are unemployment, government spending, taxes, GDP, State corruption and inflation. The countercyclical aspect of the informal economy is documented.

## 1.6 Conclusion

Analysis of the results of the PLS model developed to measure the size of the informal economy, highlighted significant causalities between perceived taxes, State spending, corruption, inflation, especially GDP and the growth of informality. It should be emphasized the growing importance of the informal economy in Haiti, from 57.1% in 2007 to 64.83% in 2010. The various causes and effects of the model show that workers are engaged in the economy informal for various reasons. Among the main reasons that people joined informal activities we can underline unemployment and inflation. When inflation raises the purchase power of the individual workers decreases therefore they will try to compensate their salary losses through informal activities. Workers joined informality to provide responses to government actions (taxes and controls) especially the indirect taxes that lead the informal economy movement. Individuals work into informality to bring answers to government compartment (state corruption and quality of public services) and to adjust their economic situation to the

market conditions (rise of inflation). Informal activities have many impacts on some macroeconomic variables. The growth of informal economy leads to important increase of the currency in circulation. Contrary to many theories, the growth of informal activities does not imply the growth of individual wealth (GDP per capita). Nevertheless, this observation is coherent to the empirical view arguing that informal growth is inversely proportional to GDP per capita (OECD, 2009). It appears an antagonist relationship between informal economy and individual wealth during the period 2007-2010. The link between informal economy growth and taxes is unsettled. The most surprising observation is the countercyclical aspect of informal economy. The more the growth of informality is important the lower the growth of the GDP is. While in the most of countries of Latin America informal economy evolves in the same way as the global economy growth.

Gains and losses could be used to encourage the Haitian government to adopt employment policies of measures that can make the informal market less attractive and control inflation. The State could also control the line of its expenditures, in order to not strengthen the informal market. However, it should be noted that like any MIMIC model, this PLS model is far from perfect and could be improved by taking account of the labour market movement or the measurement of productivity by type of activity (secondary or tertiary). Forecasts on informal using this model taking into account the evolution of macroeconomic variables, is a line of research that could deepen.



## Chapter 2

# Informal Economy Growth in Haiti: A Critical Evaluation of the Neoliberal and Deregulatory Policies

### 2.1 Introduction

During the last decade many researchers and policy makers have recognized that the shadow economy has become a sizeable and growing side of the global economy. The informal economy growth in Haiti has been more remarkable after the application of the neoliberal policies. Neoliberal theorists stipulate that deregulation policies, fiscal policy relaxation and reduction of State intervention in the market are the main solutions to prevent the informal economy from expanding. This chapter aims to evaluate the neoliberal theoretical remedy against informal economy growth by studying the effects of the deregulatory policies on shadow economy. Through a LISREL model, we use time series data over the period 2000-2010 to analyse the impacts of fiscal and budgetary policies on the informal economy. This study finds that the more flexible the fiscal policy, the higher the informal economy; the more rigorous the budgetary policy, the more shadow economy grows. It shows that the fall in GDP, the rise in final consumption expenses and government expenditures are the main factors that positively influence the growth of the shadow economy. Lowering the growth of money circulation and the exchange rate has negative impacts on informal economy. Data do not confirm the neoliberal assumption "less regulation, less informality" that theorists present as a remedy against informal economy expansion. During the last 30 years the informal economy has accrued tremendous growth. This constant growth is coupled with the application of the neoliberal policies in the Haiti's economy. In this context, questioning the relationship between shadow economy and neoliberal policies is legitimate. How has the application of the deregulatory policies impacted the growth of shadow economy? Neoliberal theorists stipulate that flexible fiscal policy (lower direct and indirect taxes on investment,



lower tariffs, lower public investments) (World Bank, 2005) and rigorous budget policy (reduce banking credit for public organism, lower interest rates, regular payment of external debts) can reduce informal economy growth (Nwabuzor; 2005; London and Hart; 2004) The first neoliberal policies and deregulatory process of 1986-89 were elaborated to stem the sudden growth of informal workers from 27.4% in 1982 to 43.2% in 1986 (Montas, 1995) and the socio-political crisis of 1986. Despite these measures, the informal economy keep growing: the informal work increased to 45% of total employment in 1990 (Montas, 1995). While informal economy continues its growth, the second set of deregulatory measures are applied during the period 1996-99. Over the period 1999-2007, shadow economy represents in average 56.4% of GDP (Schneider, Buehn and Montenegro, 2010) and 64.17% of the GDP during 2008-2010 (Aspilaire, 2014).

Considering the empirical assumption of neoliberal theorists that advocates reduction of taxes, low levels of State intervention in the market and lower level of regulation are the main ways to block informal economy growth, how should we understand the continuous growth of informal economy after the application of two sets of deregulatory measures? Answering this question requires the analysis of the relationship between informal economy and the economic variables that reflect the application of the neoliberal economic policies. These variables correspond to fiscal policy, monetary policy, and budgetary policy.

Besides tackling the neoliberal empirical assumption that flexible fiscal policy and rigorous budgetary policy can reduce growing informality, this chapter re-evaluates the traditional main causes of informal activities such as higher taxes and social security burdens (Schneider, Buehn and Montenegro, 2010), trade barriers and regulations (Johnson, Kaufmann and Zoido-Lobaton, 1998).

Few papers address the empirical assumption of neoliberal theorists regarding the informal economy (Williams, 2013), but these papers use cross-sectional surveys. Besides this challenge, our chapter addresses the problem of the impact of neoliberal deregulatory policies upon the informal economy through macroeconomic variables. To our knowledge, this chapter is the first contribution assessing the neoliberal assumption that proposes the way to reduce the growth of informality using time series data. The LISREL (Linear Structural Relation) model is used in the empirical literature of informal economy to estimate the size of the shadow economy (Schneider, 2007) but the LISREL model has never been used before to establish the consequences of fiscal, budgetary and monetary policies upon the informal economy. This chapter is an attempt to fill this gap. Hence, the main release of this chapter is threefold: first to establish the consequences of neoliberal policies upon the evolution of the shadow economy; second to evaluate the neoliberal empirical assumptions (soft fiscal,

monetary and budgetary policies) as a remedy to the growth of informality and third to analyse the traditional empirical main causes of the informal economy.

We use a LISREL model which is a particular SEM (Structural Equations Model), and we study the consequences of the deregulatory neoliberal policies on the evolution of informal economy. With respect to these consequences, we analyse the neoliberal assumption regarding the informal economy and we evaluate the main “traditional” variables that represent the causes of the shadow economy. This model uses twelve macroeconomic variables; each of them corresponds to the time series over the period 2000-2010. Our model is applicable to multivariate analysis over a time period, such as multiple countries over a period of twenty years with available data on macroeconomic variables.

Section 1 presents the stylized facts of the fiscal, budgetary and monetary policies for the last 30 years. Section 2 discusses the modelling method. Section 3 is devoted to the econometric modelling and the data description. Section 4 presents estimates, compare models and provides an analysis of the issues. Section 5 concludes and summarizes our results and suggests avenues of research.

## **2.2 Fiscal, Budgetary and Monetary Policies : Some Stylized Facts**

Since 1986 the national authorities started the dismantlement of the protectionist trade system through the two-stage liberalization and deregulatory process. Liberal policies are mainly expressed in fiscal, monetary and budgetary policies. Thus, to analyse the relationship between liberal policies and informal economy, we should focus on the link between informality and fiscal, monetary and budgetary policies. Budgetary austerity, public expenditure cuts, privatization of the most profitable public enterprises, elimination of the nontariff barriers, dramatic reduction of customs tariffs and external debts, liberalization of the interest rates, and downsizing the public workforce reflect mainly liberalization and deregulatory policies.

### **2.2.1 Budgetary and Fiscal Policies**

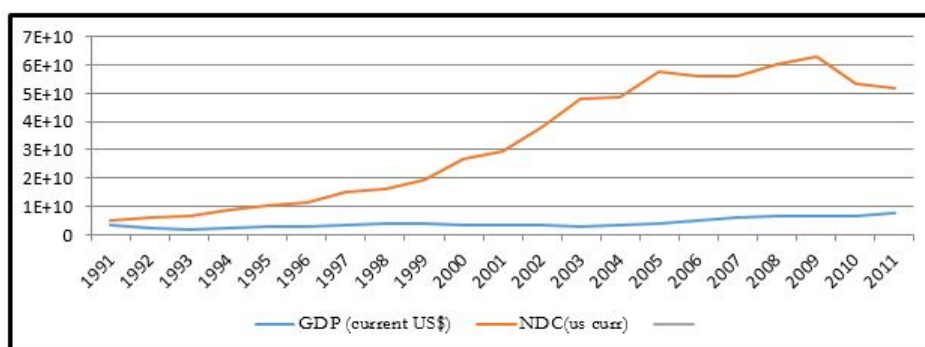
The objectives of the neoliberal process were to increase national exports following the fall of tariff barriers, to reduce the non-concessional debt of the public sector, to increase direct expenditures to development projects and to compress public expenditures (UNDP, 2002). Another objective of liberalization was the reduction of incentive for fraud; nevertheless, Haiti governments are still ranking among the most corrupt during the last two decades.

In 1986 all quantitative import restrictions (including import license) had been removed for all agricultural products. In 1987, lawmakers introduced a new 13-band tariff scheme with ad valorem average rates around 16- 20% (IMF, 2001). In 1995, all remaining import restrictions on agricultural commodities were removed. The actual tariff structure that reduces the maximum tariff rates around 50% to 0-3% for the most important basic products (rice, sugar, flour, cement) and keep a four-band tariff scheme (0, 5, 10, 15%) was adopted in 1995 (IMF, 2001). The liberalization trade dropped more than a half of the 1,600 tariff lines near a zero rate.

Approximately 60% of the budget depends on foreign aid. Taking advantage of this dependency, donors continue to impose their conditions for disbursement of funds. Budget cuts, privatization of all profitable public enterprises and state disengagement (social interventions, market control) represent the favourite conditions of disbursement for donors. In the 1990s, central government revenues represented approximately 7-9% of GDP while their expenditures have varied between 9 and 16% of GDP (IMF, 2005). These facts reflect the volatility of foreign assistance. Thus, the state gradually abandons its role in strategic sectors like quality education and adequate public health policy, which are key indicators of a well managed country. Since privatization can only reduce the state available resources, and since these processes are completed, the state no longer has the means to intervene in those areas where the needs of the population are more pressing.

Basic social needs of the population are not the priorities of the state, because it is too busy to restrict its budget on the mandatory application of the donors. According to data released by the Ministry of Economy and Finance (MEF), expenditures in the social sector fall down from 23.83% in 2001-2002 to 21.01% in 2008-2009 with a sharp fall to 13.93% between 2003 and 2004 (MEF, 2012). The government expenditures on education and public health policy respectively represented respectively 1.3% and 2.0% of the GDP for the decade 1992-2002 (IMF, 2005). The expenses incurred by the Ministry of Education, Youth and Sports were down about 1% from fiscal year (FY) 2001-2002 when they were 14.85% and FY 2008-2009, where they are of the order of 13.74%. In regard to public health, the expenditure decreased from 7.37% in 2001-2002 to 4.84% in 2008-2009 (MEF, 2012). Agriculture, which was no longer the pillar of the Haitian economy, represents only 1.6% of total public expenditures in 2008-2009 (MEF, 2012). The disengagement of the state and the shrinking of public expenditures keep roads and national port infrastructure in poor condition. This explains why the agricultural sector was unable to profit from the opportunities provided by the liberalization trade policies. Even though the net domestic credit has been increased during the last 15 years, the GDP has never grown significantly (figure 2.1). These measures have never led to significant increase of the tax revenues

and growth of income per capita is generally beneath zero (figure 2.1). That means the population welfare has never been improved.



Source: ECLAC (2012) and World Bank (2012). NDC (current US\$)

FIGURE 2.1: Evolution of Gross Domestic Product and Net Domestic Credit 1991-2011

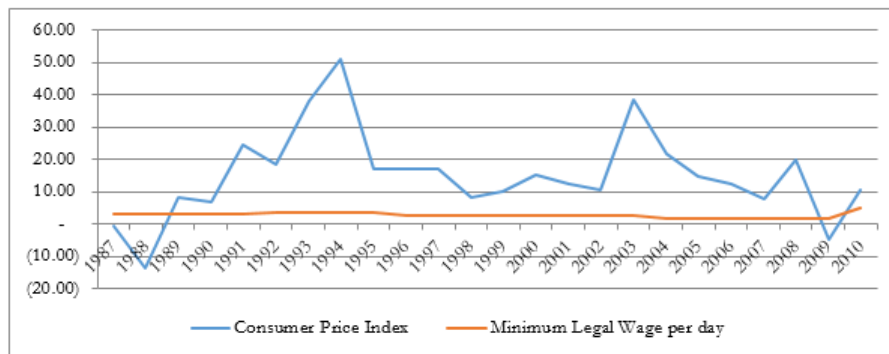
The privatization of the most profitable public enterprises and the contraction of public service staff were the main measures used to achieve these market liberalization aims. However, all of these measures have not resulted in one of the first objectives of the structural adjustment plans; namely, the reduction of the budget deficit, which was one of the most important aspects of financial programs implemented in Haiti since the political and economic crisis of 1986. The employment cut in the public sector pushed approximately 90% of workers into unemployment (Lamaute-Brisson, 2005).

These measures pushed domestic producers into unemployment as the fall in wages sent workers to unemployment or to independent work. Downsized public enterprises staff served to swell the ranks of the unemployed, and none of those parts of the population had other exit doors than the path of informal production.

The deregulatory measures contribute to increase imports and dramatically reduce exports because domestic products have failed to hold on to products from highly subsidized industrial countries. The deregulatory process eliminates the quotas, while tariffs have been reduced to a level of between 5 and 10% (UNDP, 2002), considerably low compared to those of countries Latin America and the Caribbean. Nevertheless, the elimination of licensing requirements for imports of sugar and rice and of import tariffs and quotas has contributed to pour on the US agricultural surplus on the (rice, sugar and maize). This situation contributes to the destruction of large parts of the rural economy, but the most surprising result is the reduction of tariffs on rice, which went from 35% to 3%, which increased 150% rice imports between 1994 and 2003 (OXFAM, 2005). Haiti became the fourth largest rice importer from the United States in 2009, when the country was self-sufficient rice producer until the mid-eighties. While the relative prices of rice seem to have lowered since 1995 there is strong evidence that the tradable-goods component of the CPI has

become higher since the application of deregulatory process (figure 2.2). The policy makers have never elaborated wage policy or even wages indexation practice. This explains CPI grows way faster than the wages of the workers (figure 2.2).

FIGURE 2.2: Evolution of the Consumer Price Index and the Minimum Legal wage 1987-2010



Sources: BRH (2012). Loi du 17 Avril 2003, No. 29 (2003). Loi du 06 Octobre 2009, No. 109 (2009).

Comparing to the level of 10% of tax revenues/GDP ratio of the 1980s the tax revenues/GDP ratio fell to approximately 7-9% during the decades 1992-2002 (IMF, 2005). Revenues-to-GDP increased from 7.9% of GDP to 11.9% in 2010 (World Bank, 2012a). Despite this raise, tax revenues are still low compared to international standards. The expenditures of the central government have averaged 11% of GDP over the period 1990-2005 (IMF, 2005) including large annual fluctuations in tax revenues and external financing. The main parts of the expenditures were government wages, subsidies and interest payments.

Such restrictive budgetary choices can only aggravate the precarious living conditions of the majority of the population and strengthen the socioeconomic inequalities that the country has faced for a quarter of a century. Low quality education and health interests combined with inadequate physical infrastructure are the key factors of the declining per capita income and the poor economic growth.

### 2.2.2 Monetary policy

Reduction of inflation and keeping it at a low level, eliminating the recourse of the public sector to bank credit, improving credit allocation, reinforcing financial system, consolidating required reserves for foreign currency, and reducing spread in banking intermediation were the main objectives of the monetary policy during the liberalization and deregulatory process (UNDP, 2002). Despite all the measures taken by the economic authorities, they have never mastered the inflation variation. The consumer price index presents some peaks

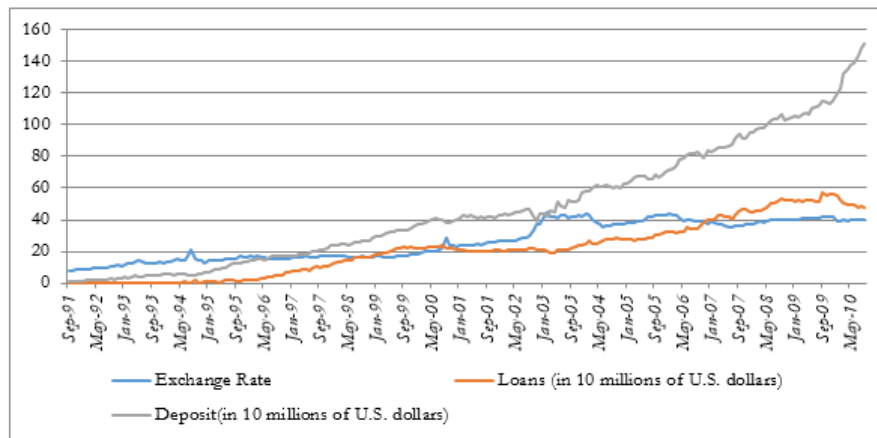
dramatically high, moreover, from one year to another the index is always greater than 15%. In an economic system where there are no unemployment benefits and social assistance from the state, the various socioeconomic crises have weakened household purchasing power. This discrepancy between inflation growth rates and the indexation of the workers' wages is shown in figure 2.3. The central bank so-called Banque de la République d'Haïti (BRH) disposes as monetary policy instruments the following tools: unremunerated reserve requirements for domestic and foreign currency deposits and issuance of BRH-bonds (weekly auctions of gourdes) bonds in maturities of 7, 28 and 91 days. The interest rate arbitrage and the credit risk analysis are important factors for loan dollarization. Commercial banks granted loans in U.S. dollars to economic agents from a fully dollarized sector of the economy or else to borrowers able to index their prices to the exchange rate such as exporters, car dealerships and fuel importers.

The exchange rate situation tends to reflect political and socioeconomic uncertainties. The exchange was maintained at 5 gourdes per U.S. dollar during the 70-80's; the political crisis period 1991-1994 was a depreciation catalyst for the gourde, reaching around 15 gourdes for 1 U.S. dollar in September 1994. The political stability period of 1995-1999 has contributed to reducing the gourde depreciation.

During the period 1997-2001 real bank credit to the private sector declined faster than the growth rate of the global economy. This leads to a fall in bank credit to the non-public sector as a ratio to GDP. During this period, debts for foreign banks did not increase significantly. This means there was no recourse to foreign borrowers. Thus, this decline in bank credit as a ratio to GDP can be understood as an indicator of the weakening of economic growth. The decline of real bank credit to the non-public sector seems to be affected by some supply and demand factors. For the supply we can enumerate the following factors. To palliate the liquidity impact of high budgetary deficits, the central bank has issued bonds in increasing amounts and interest rates (IMF, 2002). This contributes to offer commercial banks a risk-free alternative to private sector credit. This enlarged issuance of state bonds demonstrates that the budget deficit appears to have crowded out private sector credit. Several augmentations of required reserve ratios were undertaken between July 1997 and September 2001, in gourdes' liabilities from 26% to 31% and on dollar liabilities from 12% to 21% (IMF, 2002).

These considerable changes diminished the banks' capacity to lend (figure 2.3) and raise their intermediation costs, thus reducing credit supply despite growing deposit in foreign currency. The demand factors that contribute to reduced private sector credit were essentially influenced by the political situation. A massive reduction in budgetary aid, the unaccomplished foreign-funded investment projects, and the lowering of private sector investment and reduction

FIGURE 2.3: Evolution of the exchange rate, loans to private sector and deposit 1991-2010

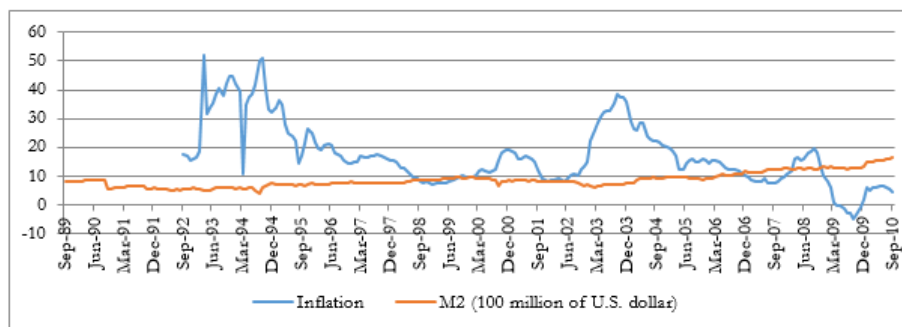


Source: BRH (2014)

in imports led to a fall in the private sector demand for bank credit.

Between the local currency (gourde) and the main foreign currency (U.S. dollar), the exchange rate fluctuated in a narrow band over the period 1991-2010 (figure 2.3). It appears that borrowers interpret the stability as the vector of central government's choice for a stable exchange rate, which is an additional incentive for interest rate arbitrage. During 2007-2010 it seems like borrowers have anticipated the exchange rate because the loans in U.S. dollars grow faster than exchange rate, which appears to be stabilized during this period. That means borrowers interpret the changing environment of monetary policy as a sign of growing macroeconomic instability. In the economic literature, inflation tends to be associated usually with higher levels of foreign currency (U.S. dollar) deposit. When comparison occurs (figure 2.3 and figure 2.4) no evidence is found of this observation.

FIGURE 2.4: Evolution of the inflation and the circulating currency (M2) 1989-2010



Source: BRH (2014) and World Bank (2012)

Higher required reserve ratios for gourde deposits entails higher



intermediation spread in gourde transactions. Demands for U.S. dollar loans have increased since 1994 and the dollar lending is continuously favoured by loan rate differentials. This explains that the monetary authorities have a part of responsibility in the high level of dollarization of the economy. This growing dollarization of the economy complicated the tight application of monetary policy due to the different money multiplier for gourde and dollar deposits. Hence the monetary policy has become less effective as counterweight to expansionary fiscal policy. It appears that inflation is guided by political uncertainties (figure 2.4), the entire peak values occur in political instability (1991-1993), political trouble period (December 2002- September 2004) and during the electoral events (November 2000, June 2008).

Liberalization and deregulatory policies were contrasted to some other structural reforms in other important areas, such as the privatization of the most profitable public enterprises. This combination deprives the population of some of the benefits of the liberal trade policy and simultaneously drops the most vulnerable part of the population in poverty.

The monetary approach of poverty refers to the principle of the existence of a living wage, which is designed according to the calculation of a threshold valued currency. This threshold is equivalent to the minimum subsistence, or the quantities of goods that a household should purchase, taking into account the general conditions in which households operate. Households whose income does not allow obtaining this minimum basket of goods are considered as monetarily poor, or those whose consumption, measured by the monetary yardstick, are below the threshold calculated. So unemployed individuals can quickly fall below the poverty line because the non-support of the state, if their savings do not allow them to live above the poverty line. The budgetary policies implemented during the mid-eighties and the nineties aim to reduce the role of the state in regulating the economy and intervene in the market. Therefore, the state was unable to support or subsidise the poor.

An economy where the employment market is unorganized, self-employment is synonymous with informal employment. It is found that the share of independent workers (57%) is higher among the working poor than the 48.7% of non-working poor (Montas, 2005). Nevertheless, since the application of the neoliberal deregulation policies, there is also a constant fall in GDP per capita. That means a growing poverty of the population (Montas, 2005). Studies have shown that there is a negative relationship between per capita income and informal employment, following this relationship there is an empirical view, arguing that informal growth is inversely proportional to GDP per capita (OECD, 2009). Therefore, individuals whose revenues are below the poverty line are seeking to obtain this minimum amount necessary to their survival. So mostly the poor (unemployed) constitute the category that joins informal employment



more often. But the working poor are those with the least valid arguments (technical or vocational training or adequate academic training on demand) to integrate the formal employment market or to create a formal business (financing and technical devices). The extent to which more than a third of informal working poor have no schooling level and only about 1.1% have a university degree (Montas, 2005). But the poor informal workers are the least well equipped in terms of labour force as more than 38% exceed 55 years (Montas, 2005). So they are the ones who find themselves hampered in the pincers of the informal economy to ensure their everydayness.

Liberalization trade and deregulation neoliberal policies reduce the state capacity to support the poorest and consolidate the social minimum acquired (employment and wages). Because of these measures the less equipped workers (independent and informal) have derived in informality, so did the unemployed poor. Based on these points we conceive a relationship between the neoliberal policies and informal economy that we model in the following section.

## 2.3 Econometric Modelling

We use LISREL, a particular of a structural equation model (SEM), to establish the impacts of neoliberal policies on the shadow economy in Haiti during the last decade. The LISREL model is a general structural equation model that contains two types of sub-models: a structural sub-model, relating endogenous to exogenous variables and to one-another. In the peer-influences model, the endogenous variables are unobserved, while the exogenous variables are directly observed, a measurement sub-model, relating latent variables (here only endogenous latent variables) to their indicators. We introduce the theoretical model.

### 2.3.1 Theoretical model, assumptions and estimation method

The structural and measurement sub-models are formulated as follows:

$$\eta_i = B\eta_i + \Gamma\xi_i + \zeta_i \quad (2.1)$$

$$X_i = \Lambda_x\zeta_i + \delta_i \quad (2.2)$$

$$Y_i = \Lambda_y\eta_i + \varepsilon_i \quad (2.3)$$

$\eta_i$  :  $(m * 1)$ - vector of endogenous latent variables

$\xi_i$  :  $(n * 1)$ - vector of exogenous latent variables

$\zeta_i$  :  $(m * 1)$ - vector of structural disturbances

$B$  :  $(m * m)$ - vector of structural parameters relating endogenous latent variables

$\Gamma$  :  $(m * n)$ - vector of structural parameters relating exogenous latent variables

$X_i$  :  $(q * 1)$ - vector of indicators of exogenous latent variables

$Y_i$  :  $(p * 1)$ - vector of indicators of endogenous latent variables

$\delta_i$  :  $(q * 1)$ - vector of measurement errors in exogenous indicators

$\varepsilon_i$  :  $(p * 1)$ - vector of measurement errors in endogenous indicators

$\Lambda_x, \Lambda_y$ : Respectively  $(q * n)$  and  $(p * m)$  vectors of regression coefficients relating indicators to latent variables

$\sigma_{ij}$  is the covariance of  $(x_i, x_j)$

$\Psi$  :  $(m * m)$  covariance matrix among structural disturbances  $\zeta_i$

$\Phi$  :  $(n * n)$  covariance matrix among exogenous latent variables  $\xi_i$

$\Theta_\varepsilon$  :  $(p * p)$  covariance matrix of  $\varepsilon_i$

$\Theta_\delta$  :  $(q * q)$  covariance matrix of  $\delta_i$

$\Sigma$  :  $((p + q) * (p + q))$  matrix of covariances among observed variables

Where:  $m$  is the number of endogenous latent variables,  $n$ : number of exogenous latent variables,  $p$ : number of indicators of endogenous latent variables, and  $q$ : number of indicators of exogenous latent variables.

Under the assumptions that the errors and latent variables are distributed following a normal multivariate distribution, all information is contained in moments of degree 1 and 2 that is to say the average and the covariance. Therefore, we can estimate the covariance matrix using the structural equations (McDonald and Hartmann, 1992).

Considering the reduced form equations, the manipulation of equation 2.1 gives :

$$\eta = (I - B)^{-1}(\Gamma\xi + \zeta) \quad (2.4)$$

$$\begin{aligned} E(\xi) &= \mu_\xi \\ E(\eta) &= (I - B)^{-1}\Gamma\mu_\xi \end{aligned} \quad (2.5)$$

$$E(\xi\xi') = \Phi + \mu_\xi\mu_\xi' \quad (2.6)$$

On the assumption that  $E(\xi\xi') = 0$  and  $E(\xi\xi') = \Psi$ , we have

$$E(\eta\eta') = (I - B)^{-1}(\Gamma(\Phi + \mu_\xi\mu_\xi')\Gamma' + \Psi)[(I - B)^{-1}]' \quad (2.7)$$

and

$$E(\eta\xi') = (I - B)^{-1}\Gamma(\Phi + \mu_\xi\mu_\xi') \quad (2.8)$$

SEM is valid when the model is overidentified and testable hence some assumptions are imposed on the model  $E(\varepsilon) = 0$ ,  $E(\sigma) = 0$ ,  $E(\varepsilon\sigma') = 0$ ,  $E(\eta\varepsilon') = 0$  and  $E(\xi\sigma') = 0$ . Moreover  $E(\varepsilon\varepsilon') = \Theta_\varepsilon$ ,  $E(\sigma\sigma') = \Theta_\sigma$  are the covariance matrices for the error measurements in variables. The model also assumed that errors in variables are

independent of the disturbances in equations that is mean  $E(\varepsilon\zeta') = 0$  and  $E(\sigma\zeta') = 0$ , this allows to relate completely equations 2.2 and 2.3 to equation 2.1.

In the simultaneous equation system, a measurement model must be imposed on the variables of the model, hence we write this model as follows

$$\begin{aligned}x &= \Lambda_x \xi + \delta \\y &= \Lambda_y \eta + \varepsilon\end{aligned}\quad (2.9)$$

Using these assumptions, we can obtain the moment structures of the measured variables and the first moments are given by

$$\begin{aligned}E(x) &= \Lambda_x \mu_\xi \\E(y) &= \Lambda_y (I - B)^{-1} \Gamma \mu_\xi\end{aligned}\quad (2.10)$$

The second moments are given by

$$\begin{aligned}E(xx') &= \Lambda_x (\Phi + \mu_\xi \mu_\xi') \Lambda_x' + \Theta_\delta \\E(yy') &= \Lambda_y (I - B)^{-1} (\Gamma (\Phi + \mu_\xi \mu_\xi') \Gamma' + \Psi) [(I - B)^{-1}]' \Lambda_y' + \Theta_\varepsilon \\E(yx') &= \Lambda_y (I - B)^{-1} \Gamma (\Phi + \mu_\xi \mu_\xi') \Lambda_x'\end{aligned}\quad (2.11)$$

Bentler (1983), by relabelling the resultant vector as  $\xi$  this means  $\xi' = (\xi', \zeta')$  and  $B$  as  $\beta$  and considering an augmented parameter  $\gamma = (\Gamma, I)$ , modifies equation 2.1 and obtains

$$\eta = \beta \eta + \gamma \xi \quad (2.12)$$

Concretely  $\eta$  and  $\xi$  are respectively  $(m * 1)$  and  $(n * 1)$  vectors and the parameter matrices  $\beta$  and  $\gamma$  are respectively  $(m * m)$  and  $(m * n)$  dimension. Bentler (1983) studies the mean and covariance structure associated with equation 2.12. On the assumption that  $(I - B)$  is invertible he expressed the dependent variables as

$$\eta = (I - B)^{-1} \gamma \xi \quad (2.13)$$

The expectation of  $\xi$  is given by

$$\mu_\xi = E(\xi) \quad (2.14)$$

and the expectation of  $\eta$  is given by the formula

$$\mu_\eta = E(\eta) = (I - B)^{-1} \gamma \mu_\xi \quad (2.15)$$

Due to the second moments about the mean it is essential to define the covariance matrix

$$\Phi = \Sigma_{\xi\xi} = E(\xi - \mu_\xi)(\xi - \mu_\xi)' \quad (2.16)$$

Following 2.13 -2.16 the cross-variance matrix is

$$\Sigma_{\eta\xi} = E(\eta - \mu_\eta)(\xi - \mu_\xi)' = (I - B)^{-1}\gamma\Phi \quad (2.17)$$

Then the covariance of is obtained by a similar process

$$\Sigma_{\eta\eta} = (I - B)^{-1}\gamma\Phi\gamma'[(I - B)^{-1}]' \quad (2.18)$$

When constants are involving in covariances that implies these covariances will be zero, then it becomes that some elements of 2.16 – 2.18 may be null. Bentler (1983) write the measures variables as follows

$$\begin{aligned} x &= M_x\xi + \delta \\ y &= M_y\eta \end{aligned} \quad (2.19)$$

Where  $M_x$  and  $M_y$  represent known matrices with zero entries except for a single unit in each row that allows to select the variables  $x$  from  $\xi$  and the variables  $y$  from  $\eta$ . Following this measure, independent measured variables are included in the vector  $x$  dependent measured variables are included in the vector  $y$ . Equation 2.19 facilitates to express the moments of the measured variables in terms of the moments of all variables. It comes that the mean and the covariance structure of the observed variables are expressed by

$$\begin{aligned} \mu_y &= M_y(I - B)^{-1}\gamma\mu_\xi \\ \mu_x &= M_x\mu_\xi \\ \Sigma_{yy} &= M_y(I - B)^{-1}\gamma\Phi\gamma'((I - B)^{-1})'M_y' \\ \Sigma_{yx} &= M_y(I - B)^{-1}\gamma\Phi M_x' \\ \Sigma_{xx} &= M_x\Phi M_x' \end{aligned} \quad (2.20)$$

By relabeling the vectors  $s' = (y', x')$ ,  $\mu' = (\mu_y', \mu_x')$  and  $\lambda' = (\eta', \xi')$  and the matrices  $M$ , a  $2 \times 2$  supermatrix where the rows  $(M_y, 0)$  and  $(0, M_x)$ ,  $\Gamma' = (\gamma', I)$  and  $B$  with rows  $(\beta, 0)$  and  $(0, 0)$  we can obtain a simpler representation of the model.

By writing

$$s = M\lambda \quad (2.21)$$

The mean and covariance structure of the measured variables  $s$  becomes

$$\begin{aligned} \mu &= M(I - B)^{-1}\Lambda\mu_\xi \\ \Sigma &= M(I - B)^{-1}\Lambda'((I - B)^{-1})'M' \end{aligned} \quad (2.22)$$

We define  $\Omega = E(ss')$  as the second moments about the origin, which, now, is easy to express in the following form

$$\Omega = \Sigma + \mu\mu' = M(I - B)^{-1}\Lambda(\Phi + \mu_\xi\mu_\xi')\Lambda'((I - B)^{-1})'M' \quad (2.23)$$

Thus, it is clear that the expressed moments are structured in terms of the unknown parameters of the model. These unknown parameters

are the elements of the matrices  $B, \Gamma, \Phi$  and  $\mu_\xi$ . In terms of LISREL notation we can write the equations 2.1 – 2.3 as follows

$$\begin{bmatrix} \eta \\ y \\ x \end{bmatrix} = \begin{bmatrix} B & 0 & 0 \\ \Lambda_y & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta \\ y \\ x \end{bmatrix} + \begin{bmatrix} \Gamma & I & 0 & 0 \\ 0 & 0 & I & 0 \\ \Lambda_x & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \xi \\ \zeta \\ \varepsilon \\ \delta \end{bmatrix} \quad (2.24)$$

The covariance matrix of the independent variables for LISREL system 2.24 described in 2.16 can be expressed as

$$\begin{bmatrix} \Phi & 0 & 0 & 0 \\ 0 & \Psi & 0 & 0 \\ 0 & 0 & \Theta_\varepsilon & 0 \\ 0 & 0 & 0 & \Theta_\delta \end{bmatrix}$$

To estimate the LISREL model we denote

$\theta$ : vector of parameters to be estimated, which contains the coefficients of all matrices

$\Sigma$ : covariance matrix of X columns

$\Sigma(\theta)$ : Covariance matrix written as a function of the model parameters.

The basic assumption of the structural equation model is:

$$\Sigma = \Sigma(\theta) \quad (2.25)$$

Using the decomposition of this matrix and non-correlation between the different variables we have:

$$\begin{aligned} \Sigma(\theta) &= \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \\ &= \begin{bmatrix} \Lambda_y(I - B)^{-1}(\Gamma\Phi\Gamma' + \Psi)(I - B)^{-1}\Lambda_y' + \Theta_\varepsilon & \Lambda_y(I - B)^{-1}\Gamma\Phi\Lambda_x' \\ \Lambda_x\Phi\Gamma'[(I - B)^{-1}]\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta \end{bmatrix} \end{aligned}$$

This matrix can be estimated by the empirical covariance matrix that we denote  $S$ , we have to find  $\hat{\Sigma}(\theta)$  such that  $\hat{\Sigma}(\theta)$  and  $S$  are as close as possible in the sense of a predefined function to be optimized. The function to be optimized has the following properties:

$F(S, \Sigma(\theta))$  is a scalar

$$F(S, \Sigma(\theta)) \leq 0 \quad (2.26)$$

$$F(S, \Sigma(\theta)) = 0 \iff S = \Sigma(\theta) \quad (2.27)$$

The function  $F(S, \Sigma(\theta))$  is continue in  $S$  and continue in  $\Sigma(\theta)$ .

For the model identification it is important that each model parameter can be expressed as a unique function of elements of the population covariance matrix. Such expression will permit to satisfy the condition of minimizing the statistical criterion (Kline, 2011). Generally, given the data and the statistical criterion to be minimized, the

population covariance matrix is estimated with the sample covariance matrix. This aspect of identification is justified because the set of parameter estimates is unique (Kline, 2011). Validating the fact that the parameters can be expressed as unique functions of the sample data is a mathematical question. This theoretical question is evaluated by resolving parameters equations in terms of elements of the sample covariance matrix.

Solving the model returns to estimate the parameters, one of the current methods consists in using a likelihood function based on the maximum-likelihood which is

$$F_{ML}(S, \Sigma(\theta)) = \log|\Sigma(\theta)| + \text{trace}(S, \Sigma(\theta)^{-1}) - \log|S| - P \quad (2.28)$$

Where:

$$P = \begin{bmatrix} \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1q} \\ \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} \end{bmatrix} & \cdots & \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \\ \vdots & \begin{bmatrix} \theta_{11}^\varepsilon & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \theta_{pp}^\varepsilon \end{bmatrix} \backslash & \vdots \\ \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} & \cdots & \begin{bmatrix} \psi_{11} & \cdots & \psi_{1m} \\ \vdots & \ddots & \vdots \\ \psi_{m1} & \cdots & \psi_{mm} \end{bmatrix} \end{bmatrix}$$

On the assumption that all of the error variables have expectations of 0, and that the observed and unobserved variables are expressed as deviations from their expectations.

Here we make the choice of the maximum-likelihood because of its strong asymptotic properties, despite of its disadvantage relatively to the assumption that the variables are multivariate normally distributed, what we verify later. When the data nonnormality occurs it is recommended to use the robust maximum likelihood also called pseudo-maximum likelihood and in small samples the covariance matrix can be singular. this situation is not a problem because the model is typically restricted then the parameters may still be estimated from the observed moments even when some of these moments were collinear (Yuan and Bentler, 1998).

### Theoretical justification of the Econometric model variables

In the theoretical model, the choice of the variables is supposed to be justified economically. The expansion of the informal economy in Haiti began between the end of the 90's and the beginning of 2000's. That period corresponds with the application of the second set of neoliberal measures to the economy. Regarding the main tasks that those measures were conceived for (macroeconomic stabilization, private investments increasing and reduction of fiscal deficit), we use some macroeconomic and other economic variables to show the

impacts of these measures on the shadow economy. Therefore, we have to justify economically the use of these variables.

### **Theory on automatic stabilization of Fiscal Policy**

Economists globally admit the effectiveness of fiscal policy as an instrument of macroeconomic stabilization (Debrun and Kapoor, 2010; Botman et al., 2006; Bullard and Singh, 2007). The economy turned to fiscal policy as their fundamental stabilization instrument for two reasons: the change in the monetary regime and the deterioration of financial conditions to the point of making monetary policy ineffective (Spilimbergo et al., 2008). Haiti's economy was concerned with both situations. The economy has turned from fix monetary regime to a flexible monetary regime since 1990. There was a deterioration of financial conditions due to the political situation, the departure of private enterprises and the privatization process of the public enterprises. The transition from fix to flexible monetary regime should be reflected in the model, hence we use real exchange rate as variable to reflect it.

A classic classification would attribute fiscal policy three main roles: allocative, redistributive and stabilizing. Contribution of fiscal policy in macroeconomic stability has three main channels, but regarding our model we treat one of them: the automatic stability. The automatic stabilization, by reduction in government saving during slowdowns periods and increase during growth periods, cushioning shocks to national expenditure (Blinder and Solow, 1974). The automatic stabilization is based on the follow facts. Proportionality of tax revenues to national income and expenditure, while independent of the business cycle, government commitments are the mainframe of public spending. Therefore, we represent the tax revenues in the model (% of GDP) by a variable called *Tax*. Downturns periods are supported by entitlement programs. To the extent that among the components of the GDP, government consumption is the less volatile (Debrun and Kapoor, 2010), the public sector is an actor of the stability because of its composition effect of domestic expenditure, based on that we introduce in the models variables such final consumption expenditures, denoted *FCE*. The fact that the government absorbs a relatively constant share of output because of its stabilizing role allows us to use a variable that represents general government final consumption expenditures that we denote *GCE*.

When the neoliberal plan was elaborated for the country, there were three main tasks of that plan, stabilizing the economy, inciting private investments and reducing the fiscal deficit. Due to that fact, in this model, we focus on stabilizing role of fiscal policy and the way that fiscal policy can be used as incitement for private investment. Thus we suppose there are no nominal rigidities; therefore, the monetary regime has no importance. The stabilizing role of fiscal policy can

be understood as a factor that contributes to reducing the variance of consumption. There are two ways to do that: by borrowing and lending or by sharing risk. Here we focus on the first way this means on the net borrowing channel, hence we are authorized to introduce as variables, **NDC**: Net domestic credit.

Government taxes individuals at a proportional tax rate in each period, this justifies the choice of taxes as variable in the model. Whether economic agents are liquidity constrained or not, the expected value of consumption depends on the variance consumption and the utility function. Here stabilization must to be understood as any policy elaborated to reducing the average variance of consumption holding constant the expected disposable income and consumption, therefore we use in this model the GDP *per capita*. The budget constraints of the fiscal authority are twofold. First, it indicates the money and quasi money growth what explained that we use the variable  $M_2$ : Money and quasi money growth. Second, budget policy focuses on three aspects of sustainability, which are solvency, growth and stability (Schick, 2005). Solvency and stability are related respectively to the ability and the capacity of government to pay its financial obligations and meet future obligations with existing tax burdens. Based on this we use a variable (net transfers on external debt) to traduce how the economy meet its financial obligations.

### **Monetary Policy: theory of liquidity preference**

It exists other alternative macroeconomic stabilization instrument such as effective monetary policy, and unrestricted access to financial instruments of economic agents. We assume that the Haiti's economy pursues a Taylor-type rule featuring consumer price index, inflation. In an open economy, *CPI*, is often the variable of interest for the monetary policy makers (Bullard and Singh, 2007). The demand for money is the second piece of the theory of liquidity preference. The most liquid asset available is money, because it represents the economy's exchange medium. Money offers higher rates of return than other assets due to these facts, money is the chosen asset hold by economic agents. The quantity of money demanded is determined by many factors, but the theory of liquidity preference puts the emphasis on one: the interest rate, because the interest rate represents the opportunity cost of holding money. Therefore, we introduce in our model variables that represent inflation, interest rate and terms of trade that we respectively denote *CPI*: Inflation, *RIR*: real interest rate and  $M_2$ : Money and quasi money growth.



### 2.3.2 Structural Equation Modelling: Theoretically and practical considerations

When it comes to analysing multivariate data and determining interdependencies and dynamic relationships between variables, only some strong models respond to the profile.

#### **Usual dynamic models: inadequate to analyse the shadow economy**

The vector auto-regressive (VAR) models are seen as a classical instrument in Econometrics to determine interdependencies and dynamic relationships between variables. But here we could not use that tool because VAR models explain the endogenous variable only by their own story (that means their past values) except deterministic regressors. Knowing the informal economy is not a direct observed variable in the economy, so it cannot explain itself by its own past values only. Furthermore, validating the VAR obliges the verification of two fundamental assumptions, which are absence of serial correlation and residuals are normally distributed. Unobserved variables cannot verify these econometric assumptions.

Segundo, we have structural vector autoregressive models (SVAR) that were conceived to overcome the shortcomings of VAR models. Structural vector error correction (SVEC) framework is closely related to the SVAR modelling approach. An important characteristic of SVAR is the exclusive focus on unsystematic policy shocks instead of systematic policy (Krusec, 2003). Nonetheless, we envisage studying the interrelations of systematic policy (budget, fiscal and monetary policies) and their impact on shadow economy growth. SVAR models oblige to find a model with instantaneously uncorrelated residuals, which supposes the variables cannot have common causes other than the factors included in the model. Among macroeconomic variables that assumption is not correct because macroeconomic variables have other common causes out of these included with the model. Our main objective is to show the relationship between macroeconomic variables and shadow economy growth, therefore SVEC is not adequate for such modelling.

Thirdly vector error correction model (VECM) is just a representation of cointegrated VAR models. Beside the problems that we explain above, there is a big risk in applying a VAR model for integrated time series like t-statistics of tests highly significant and high value of  $R^2$  explained variables, although it exists no relation between the variables. But here we want to highlight the relations between shadow economy and public economic policy in particular budgetary, fiscal and monetary policies, so VECM model is not the best one to establish these links.

### **Unobserved phenomenon: Advantages of Structural Equation Modelling**

SEM is one of the statistical models in particular an extension of general linear modelling (GLM) procedures, which can be used to evaluate the validity of a theory by using empirical data. LISREL model which is a particular SEM is very fit to evaluate a theory validity (Jakobowicz, 2007). SEM presents many advantages; we underline two of the most interesting of them. The first one is the fact that it can be used to analyse the relationships among variables construct from others (latent constructs). The second one regards the use of SEM, which is applicable for experimental or non-experimental data, as well as cross-sectional and longitudinal data or time series data. The methodology of SEM is a confirmatory approach of the structural theory to the multivariate analysis.

The goal of a SEM is to determine whether a hypothesized theoretical model is consistent with the data collected to reflect this theory. We envisage determining whether informal economy was affected by the macroeconomic policy conducted during the last decade. The consistency is evaluated through model-data fit, which indicates the extent to which the postulated network of relations among variables is plausible. Here we test in one hand if the budgetary policy variables influence the informal economy growth, and in the second hand if both variables (fiscal policy and monetary policy variables) are responsible the way that the shadow economy grows in Haitian economy. Structural equations modelling presents the advantage that a variable can simultaneously serve as a source variable (independent or exogenous variable) in one equation and a result variable (dependent or endogenous variable) in another equation, all that in the same model in a chain of causal hypotheses. In using a simultaneous equations system, we consider the fact that macroeconomic variables influence each other, and these variables also influence the indicators of the informal economy. We regard some of these variables as exogenous variables in some economics relation or as endogenous variables in the others relationships of the model, in that way all of the macroeconomic variables influence shadow economy variables. Due to the fact that structural relations consist in hypotheses about directional influences (how an exogenous variable affect another endogenous variable) or causal relations of multiple variables, we can evaluate which variable among the macroeconomic policy variables influence informal economy and in what direction they do it. SEM also allows residuals of variables (indicators) to correlate in case that the variables have common causes other than the factors included in the model.

SEM has a basic covariance structure analysis; the generalized SEM is sometimes referred to as causal modelling, because the SEM generally study the interrelations among variables and these interrelations are hypothesized to generate specific observed covariance patterns among the variables. Using equations that are estimated to

simultaneously provide to SEM a more interesting and direct way of modelling indirect effects and other complex relationship among variables. For all that SEM seems like the best model to test the hypothesis that fiscal policy and monetary policy inducted the informal economy growth during the last decade in Haiti.

## 2.4 Model Designing

Since 1986, three remarkable observations have marked the global economy in Haiti. Twenty-five years of neoliberal and deregulation policies that have changed the macroeconomic structures of inadequate monetary policy (IMF, 2001). During this period, we observe growing informality and conducting of inappropriate fiscal policy (Aspilaire, 2014; Schneider, Buehn and Montenegro, 2010). This quarter of a century is marked by increasing poverty (Lamaute-Brisson, 2005; Montas, 2005) and unsuccessful budgetary policy. Three related issues arise: the consequences of liberalization trade and neoliberal deregulatory policies on shadow economy evolution; the evaluation of the neoliberal empirical assumption advocates that flexible fiscal, monetary, and budgetary policies are the remedy of informality growth and questioning traditional empirical macroeconomic causes of the informal economy.

### 2.4.1 Data description and Model

To address the above issues, we use the econometric model elaborated previously. In this LISREL model we consider the informal economy as an endogenous latent variable explained by two others exogenous latent variable which represents in one hand budgetary policy and fiscal policy and in the second hand monetary policy.

#### Data and Software

Our data are treated with a SEM software program, LISREL (linear structural relationships) is one of the earliest SEM programs and perhaps the most frequently used. First, LISREL checks distributional assumptions, such as univariate and multivariate normality, calculating summary statistics; second, it calculates means, an asymptotic covariance matrix of variances and covariances. LISREL permits the analysis of multilevel models for hierarchical data in addition to the core models. We use R software to estimate our model. This software uses a function which calls SEM function. The SEM function is implemented from an algorithm which name is RAM reticular action model (McArdle and McDonald, 1984).

The empirical analysis is based on annual time series data over the period 2000-2010. We use 14 macroeconomic variables that represent budget, fiscal and monetary policies. We also use variables that are

very concerned by monetary policy and variables that reflect wealth and global economic transactions. The use of all of these variables is theoretically justified previously. All the data are from the World Bank (2012a), sources are mentioned for the variables that are not pooled in the World Bank database

*GDP*: GDP *per capita* (current U.S. dollar)

*CPI*: Consumer Prices Index (annual %)

*UR*: Proportion of unemployed in percentage of total labour population (source: CIA World factbook, 2011)

*NDC*: Net domestic credit (current local currency unit)

*CC*: Control of corruption index (-2.5=worst, 2.5= best)

*ER*: official Exchange rate) (LCU per U.S. dollar, period average)

*RIR*: Real interest rate (%)

*NTED*: Net transfers on external debt, long-term (current U.S. dollar)

*GCE*: General government final consumption expenditure (current U.S. dollar)

*Tax*: Tax revenue in percentage of GDP (source: ECLAC, 2012)

*FCE*: Final consumption expenditure, etc. (current U.S. dollar)

*FDI*: Foreign Direct Investment (current U.S. dollar)

*M<sub>2</sub>*: Money and quasi money growth (annual %)

*BOP*: Balance of Payments (current U.S. dollar)

### Conditions on Multicolinearity and Multivariate normality

All the initial variables are standardized for the needs of the model. When we construct SEM we should be careful with the related variables that we use as indicators of a construct.

TABLE 2.1: Multicolinearity test for indicators variables given by the correlation matrix

	Gross Domestic Product	Consumer Price Index	Unemployment Rate	Net Domestic Credit	Control of Corruption
Gross Domestic Product	1.000	-0.673	-0.876	0.653	0.874
Consumer Price Index	-0.673	1.000	0.480	-0.175	-0.658
Unemployment Rate	-0.876	0.480	1.000	-0.861	-0.626
Net Domestic Credit	0.653	-0.175	-0.861	1.000	0.447
Control of Corruption	0.874	-0.658	-0.626	0.447	1.000

Source: Computations by the author

If indicators variables are too highly correlated, the results of certain statistical tests may contain bias. When checking multicollinearity for any pair of variables (table 2.1) with a correlation coefficient higher than .85 one of the two variables should be removed from further analysis (Kline, 2005). Hence, the results (table 2.1) show that between GDP and CC one should be excluded.

In SEM when data fail the normal characteristics test, sample should be large enough to minimize the importance of sampling error of the estimation procedure. Hence, we test the normality characteristics of our 14 variables with the Jarque-Bera test (table 2.2).

TABLE 2.2: Jarque-Bera normality test of the variables

Variables	$\chi^2_{1-\alpha}$	Variables	$\chi^2_{1-\alpha}$
Gross Domestic Product	1.128 (.5688)	Net transfers on external debt	2.137 (.3435)
Consumer Price Index	3.711 (.1563)	Foreign Direct Investment	2.879 (.2369)
Unemployment Rate	.600 (.7405)	Tax revenues	0.746 (.6884)
Net Domestic Credit	1.349 (.5092)	Final Consumption Expenditure	1.109 (.5743)
Control of Corruption	0.5494	Government Consumption and Expenditure	4.162 (.1248)
Exchange Rate	2.240 (.3262)	M2	0.822 (.6627)
Real Interest Rate	.602 (.7401)	Balance of Payments	4.3259 (.115)

Source: Computations by the author. In parentheses are the p-values. Normality is accepted if p-value > .05

Results ( table 2.2) show that all the 14 variables present normal distribution characteristics, and then we avoid the risk to maximize the impact of sampling error during the estimation procedure. Now we have to verify the multivariate normality of the data.

TABLE 2.3: Shapiro-Wilk multivariate normality test

S-W	.2968
p-value	$2.19e - 07$

Source: Computations by the author. Multivariate normality is accepted if p-value > .05

Because of the result of the p-value (table 2.3) the data is rejected as multivariate normal. However, this assumption is hardly met in practice (Kline, 2005). Nevertheless, testing a model built with non-multivariate normal data may inaccurately suggest that the model is a poor fit to the data while the model is a good fit to the data and vice

versa (Teo, Tsai and Yang, 2013). When the observed variables are not from a multivariate normal distribution it is recommended to use the Satorra-Bentler robust maximum likelihood method of parameter estimation (Bentler, 2004; Yuan and Bentler, 1998; Satorra and Bentler, 1994).

### 2.4.2 Model Design

The most important step is specification, which is the representation of the hypotheses in the form of a structural equation model Specification. The main significance of the specification is in the fact results from later steps assumed that the model is basically correct. The process of specification can be described by a series of equations. The model's parameters are defined by these equations, and they are presumed to establish relations among observed (measured variable) or latent variables.

The matrices  $\Lambda_x$  and  $\Lambda_y$  in equations 2.1 and 2.3 are constituted of factors loading of the structural regression of measured on unmeasured variables. In path analysis, measured or observed variables are treated single as exact, error-free representations of the constructs of interest, which is an assumption that does not hold in economics theory. When a measured variable is not error-free, which is always true in economics, if a complex model is constructed, there is a risk that estimates of path coefficients to be biased in unpredictable ways (Bollen, 1989). To fix their error variance, if estimates of reliability for the observed variables are available, they can be incorporated into the model. Alternatively, considering that multiple observed variables measure the same latent constructs, a measurement model can be used to distinguish the common variances of the measured variables from their error variances. That is a huge point that each substantive latent variable must to be scaled, but that is also the case for each error term. Generally, a unit loading identification (ULI) constraint is used to assign scales to disturbances in structural models or measurement errors in measurement models through. Fixing the path coefficient for the direct effect of a disturbance or measurement error on the corresponding endogenous variable to equal the constant 1.0 is the interpretation of that. This specification presents the advantage that for observed endogenous variables; the unstandardized total variance will equal the sum of the unstandardized residual variance and the explained variance. The indicator with the ULI constraint is called the reference variable or marker variable. This specification consists in assigning to a factor a scale related to that of the explained variance of the marker variable. This fixes the factor variance to 1.0, and also standardizes the factor. When a factor is scaled through a UVI (unit variance identification) constraint, all factor loadings are free parameters.

We build six different models by using the equation 2.2 (which contains the two sub-models of the measurement model) and the equation 2.1.

Model 1:

$$\begin{aligned} \begin{bmatrix} Tax \\ FCE \\ GCE \\ M_2 \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix} \\ \begin{bmatrix} GDP \\ CPI \\ NDC \\ UR \\ CC \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \\ \gamma_{4j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}, \eta_{j4}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \\ \varepsilon_{4j} \end{bmatrix} \\ \begin{bmatrix} ER \\ NTED \\ RIR \\ FDI \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix} \end{aligned}$$

Model 2:

$$\begin{aligned} \begin{bmatrix} GCE \\ FCE \\ Tax \\ M_2 \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix} \\ \begin{bmatrix} GDP \\ CPI \\ NDC \\ UR \\ CC \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \\ \gamma_{4j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}, \eta_{j4}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \\ \varepsilon_{4j} \end{bmatrix} \\ \begin{bmatrix} ER \\ NTED \\ RIR \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix} \end{aligned}$$

Model 3:

$$\begin{aligned} \begin{bmatrix} Tax \\ FCE \\ GCE \\ BOP \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix} \\ \begin{bmatrix} GDP \\ CPI \\ NDC \\ UR \end{bmatrix} &= \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} ER \\ NTED \\ RIR \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix}$$

Model 4:

$$\begin{bmatrix} Tax \\ FCE \\ GCE \\ BOP \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix}$$

$$\begin{bmatrix} GDP \\ CPI \\ UR \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix}$$

$$\begin{bmatrix} RIR \\ NTED \\ ER \\ NDC \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix}$$

Model 5:

$$\begin{bmatrix} Tax \\ FCE \\ GCE \\ BOP \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix}$$

$$\begin{bmatrix} GDP \\ CPI \\ UR \\ NDC \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{32} \end{bmatrix}$$

$$\begin{bmatrix} RIR \\ NTED \\ ER \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix}$$

Model 6:

$$\begin{bmatrix} Tax \\ FCE \\ GCE \\ M_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{31} \end{bmatrix}$$

$$\begin{bmatrix} GDP \\ CPI \\ UR \\ NDC \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\eta_{j0}, \eta_{j1}, \eta_{j2}, \eta_{j3}] + \begin{bmatrix} \varepsilon_{0j} \\ \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \end{bmatrix}$$

$$\begin{bmatrix} RIR \\ NTED \\ ER \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma_{1j} \\ \gamma_{2j} \\ \gamma_{3j} \end{bmatrix} \times [\xi_{j1}, \xi_{j2}, \xi_{j3}, \xi_{j4}] + \begin{bmatrix} \delta_{0j} \\ \delta_{1j} \\ \delta_{2j} \\ \delta_{3j} \end{bmatrix}$$



Where:  $\Lambda_x = (1, \gamma_{1j}, \gamma_{2j}, \gamma_{3j})^T$ ,  $\delta = (\delta_{0j}, \delta_{1j}, \delta_{2j}, \delta_{3j})^T$ ,  $\xi = (\xi_0, \xi_1, \xi_2, \xi_3)$ ,

$$\Lambda_y = \begin{bmatrix} 1 & \gamma_{1j} & \gamma_{2j} & \gamma_{3j} & \gamma_{4j} \\ 1 & \gamma_{1j} & \gamma_{2j} & \gamma_{3j} & \gamma_{4j} \end{bmatrix}^T, \eta = \begin{bmatrix} \eta_{0j} & \eta_{1j} & \eta_{2j} & \eta_{3j} & \eta_{4j} \\ \eta_{j0} & \eta_{j1} & \eta_{j2} & \eta_{j3} & \eta_{j4} \end{bmatrix} \text{ and}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_{0j} & \varepsilon_{1j} & \varepsilon_{2j} & \varepsilon_{3j} & \varepsilon_{4j} \\ \varepsilon_{j0} & \varepsilon_{j1} & \varepsilon_{j2} & \varepsilon_{j3} & \varepsilon_{j4} \end{bmatrix}^T$$

$y$ : Matrix which contains the variables which measure mostly the budgetary and fiscal policies

$x$ : Matrix which contains the variables which measure mostly the monetary policy

$\xi$ : Matrix which contains the variables which measure the shadow economy (SE)

### 2.4.3 Strategy of Analysis

Three main strategies for specifying and evaluating model are known in the literature (Joreskog and Soborn, 1996; Joreskog and Goldberger, 1975; Goldberger, 1972). The first one is the strictly confirmatory strategy which is used to evaluate a single *a priori* model in isolation. This strategy presents two problems, it is highly restrictive and user has little recourse if model does not work well. The second one is the model generation strategy consists in fitting to data an initial model and try to modify it until it fits well. But, this strategy is potentially misleading, researchers have demonstrated that such data-driven model modifications present lack of validity (MacCallum, 1986) and are highly capitalized on chance (MacCallum and al, 1992). The use of this strategy must be conditioned, first it must be recognized that the resulting model is in part data driven; second, modifications must be fundamentally meaningful; and third, the modified model must be evaluated by fitting it to an independent sample. Very few models using this strategy respond to third condition. The third one is the alternative models strategy using multiple *a priori* models and providing comparative information about alternative explanations of the data through the *a priori* models, this strategy avoids the difficulties met in the two others strategies and protects against a confirmation bias (Kline, 2011). Here we use the third strategy in comparing three models and providing substantive explanation of the data.

Using the alternative models strategy, we have to deal with an important factor, which is the small size of our sample. We envisage comparing alternative models of varying complexity to give the best possible explanation of the data. The number of parameters of the model increases with the complexity of the model. When the model has a large number of parameters, it loses some precision when the parameters are estimated; more it does with smaller samples because small sample can have difficulty to support the estimation of some

complex models. It is recommended to favour simpler models when sample size is small.

Besides the problem of small size sample, we must verify that our data presented normality characteristics because LISREL method is based on distributional assumptions of univariate and multivariate normality of the data. However, we find that our data does not satisfy the multivariate normality assumption. To overcome this problem, we introduce a fundamental method for parameters in our SEM computer program, which are bootstrapping (small size sample).

Due to the restrictiveness of the model ( $N = 11$ ) we use a particular computer-based process called bootstrapping. Use of bootstrapping has become a usual supplement to statistical parameter as well as assisting potential small sample issues (Nevitt and Hancock, 2001). Bootstrapping is a method of re-sampling, which is divided into two general kinds: nonparametric bootstrapping, where the sample is treated as a pseudo-population and parametric bootstrapping, where samples are generated randomly from a theoretical probability density function specified in advance it is used to study the properties of particular estimators (Kline, 2011). In our modelling, we use nonparametric bootstrapping. This method generates other data sets by selecting randomly cases from the original data set, generally the same number of cases as the original. With the aim that the repeated sampling will display a comprehensive scheme of the sampling distribution (Efron and Tibshirani, 1986). Bootstrapping proposes other extensions besides drawing random samples of size  $n$ , it can draw larger sample sizes than initial sample size  $n$  this means bootstrapping can generate  $2n, 3n, 4n$ . According to other researchers, strong factor loadings can spare a small sample size with simple structure, (Guadagnoli and Velicer, 1988). A brief overview of mathematical understanding of bootstrapping is presented (**Appendix2**) to help to illustrate its methodology.

## 2.5 Estimates, Model Validation and Analysis Issues

Our structural equation model is a structural regression (hereafter SR) model, also called a full LISREL model. An SR is defined as the synthesis of a structural model and a measurement model.

### 2.5.1 Model Validation

The model was computed with the SEM function based on the maximum-likelihood estimates for general structural equation models returns; it displays parameter estimates, together with the likelihood-ratio chi-square statistic for the model, contrasting the model with a

just-identified model, which perfectly reproduces the sample covariance matrix.

The first important characteristic of the SEM model refers to the convergence (number of iterations). The Less a model needs iterations to converge better is that model for other advanced analysis. Regarding the six models (table 2.4) the second and third models are basically those who needed more iterations to converge. Despite the fact they have latent variables which are very different from their composure and their exogenous variables. The model number 5 presents the best convergence process, after that we have the model number 6, number 1, and model number 4.

In SEM one important objective is to find the most parsimonious model that is able to represent the interrelationships among variables that reflect accurate associations observed in the data (Teo, Tsai and Yang, 2013). Hence, a large degree of freedom ( $DF$ ) refers to a more parsimonious model. Therefore, results (table 2.4) show that all the six estimated models are very parsimonious.

TABLE 2.4: Indices of the model goodness-fit of data

Model	Iterations	Chi-square	S-B	$df$	SRMR	RMSEA	CFI	TLI	AIC
1	61	382.28	0.992	62	0.194	0.685	0.461	0.321	317.449
2	92	288.77	0.993	51	0.198	0.654	0.497	0.349	283.585
3	88	294.03	0.799	52	0.274	0.743	0.382	0.216	294.223
4	64	284.93	1.048	51	0.15	0.646	0.542	0.408	300.783
5	56	284.84	1.048	51	0.15	0.646	0.542	0.408	300.783
6	61	283.55	1.051	51	0.144	0.644	0.545	0.411	300.288

Source : Computations by the author; S-B: Satorra-Bentler correction factor

The Satorra-Bentler scaling correction factor, which is the average generalized design effect of nonnormality data, takes into account both nonnormality and the sampling design effect (Bentler 2003, Satorra and Bentler, 1994). The value of the Satorra-Bentler correction factor for model 1 is 0.992 (table 2.4), which means the effect of nonnormality is 0.8% lower than the conventional analysis. For the model 6, this correction factor is 1.051 (table 2.4) this means the effect of nonnormality is 5.1% higher. The Satorra-Bentler statistic (Satorra and Bentler, 1994) adjusts downward the chi-square value from standard maximum likelihood estimation. This adjustment reflects the degree of kurtosis of the data. There is a wide variety of measures and criteria for determining what constitutes good fit for SEM models (Marsh, Balla, and McDonald, 1988). To analyse the performance of various fit measures with respect to sensitivity to model misspecification, the use of three fit indices categories are recommended in the SEM literature: the absolute fit indices, the parsimonious fit indices and the comparative fit indices.

Absolute fit indices assess how well the model reproduces the data (Hair et al., 2006). The main absolute fit index is the chi-square

(Teo, Tsai and Yang, 2013). Significant values of chi-square suggest that the model does not fit well the sample data. However, the chi-square is known to be too sensitive to sample size. When sample size is small the probability level tends to be significant, consequently a non-significant  $p$ -value is uncommon, although the model fit the observed very well (Teo, Tsai and Yang, 2013). This is exactly the case of the six models that we estimate, referring to the results (table 2.4) they all have a significant  $p$ -value ( $< .05$ ). Since these increments in fit are associated with a statistically significant difference between the null model and the constructed model, so the difference between the constructed and the saturated model need not to be a source of concern due to the small sample size. Secondly, the chi-square statistic is in essence a statistically significant test which is sensitive to sample size (Bentler and Bonnet, 1980; Joreskog and Soborm, 1993). We are in the case where small sample is used; the chi-square statistic lacks power and then may not discriminate between good fitting models and poor fitting models (Kenny and McCoach, 2003). Due to the restrictiveness of the model chi-square, researchers have sought alternative absolute fit indices to assess model fit.

The other common absolute fit indices provided by the software are the standardized root mean square residual (SRMR) and root mean square error approximation (RMSEA). SRMR range from zero to 1.0 with well-fitting models obtaining values less than 0.05, all the six models that we used have SRMR values from 0.144 to 0.274 which are higher than 0.05. Nevertheless, SRMR will be higher when there are models based on small sample sizes (Kline, 2011), which is exactly the case of our models. Lower RMSEA values ( $< .05$ ) assess a good fitting model, all the six competing models present RMSEA higher than .05 level (table 2.4) but this index provides better results for large sample size model than for small sample size data (Kline, 2011).

Comparative fit indices measure whether a competing model is better than a baseline model that assumes that all observed variables are uncorrelated. One of the comparative fit indices widely used is the comparative fit index abbreviated CFI (Bentler, 1990) which is the revisited form of the NFI. The widespread use of the CFI is due to its strengths, including its relative insensitivity to models complexity (Teo, Tsai and Yang, 2013). The CFI should be Normed and varies from 0 to 1; nearer 1 are the values better the model fits the data. Among the six models that we estimate three of them present acceptable CFI values (table 2.4) because values are higher than 0.5. Tucker-Lewis Index (TLI) also called NNFI (Non-Normed Fit Index) has a preference for simpler models (Bentler and Bonnet, 1980) when models are good fitting the TLI values approach 1.0. All the six models that we estimate present mediocre TLI values. Nevertheless, in situations where small samples are used, the value of the TLI can indicate a poor fit despite other statistics pointing towards good fit (Bentler, 1990; Kline, 2005). This explains our poor TLI values.

Parsimonious indices evaluate the discrepancy between the observed and implied covariance matrix while taking into account the complexity of the model (Teo, Tsai and Yang, 2013).

There are no golden rules for assessment of model fit, reporting a variety of indices is necessary (Crowley and Fan, 1997) because different indices reflect a different aspect of model fit. Three criteria of model fitting are suggested by Schumacker and Lomax (2004). The first one is the non-statistical significance of the chi-square test, this means a non-significant  $p$ -value. All the six estimated models present significant  $p$ -value (table 2.4), but the model chi-square has many problems associated with it, it is still essential that this statistic, along with its degrees of freedom and associated  $p$ -value should at all times be reported (Kline, 2005; Hayduk and al, 2007). Generally a non-significant  $p$ -value is uncommon (Kline, 2011), hence the model can be close fit to the observed data and present a significant  $p$ -level. The second criterion regards the statistical significance of each parameter estimates for the paths. For all the six estimated models we register statistical significance for each parameter estimates for the paths (table 2.5) because their  $t$ -values are higher than 1.96 hence they are significant at the 0.05 level at least. The third criterion address the problem of magnitude and direction of the parameter estimates which are need to be consistent with the fundamental theory. By observing the estimation results (table 2.5) for all the six models, in the sub model so-called **BP** in this table all the three variables *Tax*, *Final Consumption Expenditures*, and *Government Consumption Expenditures* have positive coefficients, which is coherent to the theory on automatic stabilization of fiscal policy (Debrun and Kapoor, 2010) that we developed on the sub-section 2.3.1. For the sub-model denoted **MP** in the table 2.5, the variables *Real Interest Rate* and *Exchange Rate* when they are not fixed to be equal to 1, they present the expected sign accordingly to the theory of liquidity preference that we developed in the sub-section 2.3.1, the same observation is valid for the variables *Consumer Price Index* and  $M_2$  in the sub-model denoted **SE**. In the sub-model **SE** the values of the variables *Unemployment Rate* and *Net Domestic Credit* are coherent to the the elaborated theory on the informal economy (Schneider, Buehn and Montenegro, 2010; Schneider, 2007)

Generally, all structural equation models must be overidentified (Hoyle, 1991; Raykov and Marcoulides, 2000). This means there are more equations for the model than unknown parameters because the fit of a just-identified model cannot be properly evaluated (MacCallum, 1995). This characteristic is shown by the number of degrees of freedom if that number is zero the model is just-identified. Unknown parameters are those for which the SEM process will generate numerical values. If the model has one or more degrees of freedom, it is over-identified, then all the six models are over-identified, because all the models have at least 51 degrees of freedom (table 2.4). Hence all the six competing models can be evaluated properly.

If an alternative model is a subset of the initial model, the difference in chi-squared values between the two models is itself a statistic that we can use to test the importance of the parameters that differentiate the model (Bentler, 1980). A such statistic test is the chi-squared difference test which evaluates the model improvement.

Results (table 2.4) show statistics such as CFI and the Chi-squared difference test toward good fit but others such as SRMR, RMSEA and NNFI give indication for bad fitting models for the six *a priori* models that we have evaluated. Researchers underline that these statistics need a sample size of 200 to make their use reliable (Kline, 2011) contrary to that assumption our six models are realized on a sample size of 11. Thus, we do not allow concluding that the six models that we estimate are bad or good fitting models. Therefore, to find the appropriate model we compare six *a priori* models.

Nevertheless, practitioners noticed that statistical tests are less relevant for SEM than the other type of techniques, therefore two main reasons advocates that statistical tests play a smaller role in SEM. First, SEM allows the evaluation of the entire models, which gives a higher-interest perspective to statistical tests of individual effects in models (Kline, 2011). Second, virtually all non-significant statistics in a small size sample would be statistically significant in a sufficiently large sample (Thompson, 1992).

## 2.5.2 Estimates and Model choice

An extremely well fitting model can present the problem that all the relationships among its variables are weak or even zero, which means that effects hypothesized in the model are weak also. These weak relationships would be reflected by the resulting parameter estimates and by large residual variances for endogenous variables. Therefore, it is crucial to look at the parameter estimates and their significance even when the fit is very good (MacCallum, Browne, and Cai, 2006).

The test of individual parameter estimates for statistical significance is based on the ratio of the parameter estimate to its standard error estimate (often called *z* or *t*-value). As a rough reference, absolute value of this ratio greater than 1.96 may be considered statistically significant at the .05 level and greater than 2.58 for the .01 level. In the table 2.5, **SE** is an endogenous latent variable, **BP** and **BP** are exogenous latent variables that also influence the endogenous latent variable.

The values 1.00 of some parameter estimates in table 2.3 and table 2.4 derive from the method consisting in imposing unit loading identification (ULI) constraints. This principle means we fix loading coefficient for the direct effect of a factor on any one of its indicators to equal 1.00. The indicator with the ULI constraint is considered as the reference variable or marker variable.

TABLE 2.5: Estimates of the six a priori models

Latent Variables	Measured Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<b>SE</b>	<i>Gross Domestic Product</i>	1.000	1.000	1.000	1.000	1.000	1.000
	<i>Consumer Price Index</i>	-0.695*** (0.017)	-0.691*** (0.020)	-0.560*** (0.043)	-0.788*** (0.047)	-0.560*** (0.045)	-0.561*** (0.046)
	<i>Unemployment Rate</i>	-0.794*** (0.026)	-0.794*** (0.031)	-1.085*** (0.05)	-0.846*** (0.035)	-1.085*** (0.053)	-1.080*** (0.057)
	<i>Net Domestic Credit</i>	0.559*** (0.020)	0.559*** (0.024)	0.904*** (0.020)	M2 -0.433*** (0.016)	0.904*** (0.025)	0.896*** (0.024)
	<i>Control of Corruption</i>	0.193*** (0.034)	0.192*** (0.039)	-	-	-	-
<b>MP</b>	<i>Exchange Rate</i>	1.000	1.000	1.000	-0.363*** (0.006)	-0.598*** (0.040)	-0.597*** (0.039)
	<i>Real Interest Rate</i>	-2.079*** (0.013)	-2.245*** (0.016)	-1.671*** (0.028)	1.000	1.000	1.000
	<i>Net Transfers on External Debts</i>	2.279*** (0.011)	2.574*** (0.010)	1.664*** (0.021)	-1.450*** (0.047)	-0.996*** (0.059)	-0.998*** (0.058)
	<i>Foreign Direct Investment</i>	1.151*** (0.014)	-	M2* -0.156*** (0.029)	NDC* -0.729*** (0.005)	M2* 0.093 (0.050)	-
<b>BP</b>	<i>Tax</i>	1.000	2.080*** (0.014)	1.000	1.000	1.000	1.000
	<i>Final Consumption Expenditures</i>	1.088*** (0.032)	2.264*** (0.029)	0.951*** (0.048)	1.592*** (0.027)	0.951*** (0.048)	0.952*** (0.049)
	<i>Government Consumption Expenditures</i>	0.481*** (0.014)	1.000	0.692*** (0.039)	0.124*** (0.013)	0.692*** (0.037)	0.697*** (0.038)
	M2	-0.209*** (0.020)	-0.437*** (0.011)	BOP* -0.885*** (0.032)	BOP* -1.450*** (0.047)	BOP* -0.885*** (0.036)	-0.232*** (0.034)
<b>SE</b>	<i>Monetary Policy</i>	0.886*** (0.017)	0.900*** (0.015)	-0.142*** (0.045)	-0.078 (0.056)	0.085 (0.051)	0.083 (0.055)
	<i>Fiscal Policy</i>	0.678*** (0.033)	1.486*** (0.019)	1.004*** (0.028)	1.375*** (0.021)	1.004*** (0.028)	1.009*** (0.029)
<b>Monetary Policy</b>	<i>Fiscal Policy</i>	0.396*** (0.019)	0.726*** (0.018)	0.458*** (0.069)	-0.847*** (0.032)	-0.766*** (0.046)	-0.765*** (0.045)

Source: author's calculation. \*\*\*= significance level at 1%, an estimate is significant if  $|z| > 2.58$ ; \* = variable that replaced the initial variable in the model; in parentheses we have the standard error for each variable.

The standard errors evaluate the variability of each parameter estimate (Bentler, 1980). The different standard error values (table 2.5) show that the parameter estimates have very low variability.



Our strategy, as we stated previously, consist in using multiple a priori models and providing comparative information about alternative explanations of the data through the a priori models, besides its multiple advantages this strategy protects against a confirmation bias (MacCallum and Austin, 2000; Joreskog and Sorbom, 1996).

To analyse differences between estimated models it is usual to make use of the null hypothesis of no difference in population discrepancy function values. However, it is known that this null hypothesis is always false in practice and it presents limited empirical interest (MacCallum, Browne, and Cai, 2006). To bypass this problem, we use the test of the hypothesis of the small difference in fit Chi-square difference test.

TABLE 2.6: Indices of the model goodness-fit of data

Model dif- ferences	$\chi^2$ dif- ference (df)	Critical $\chi^2$ (non- centrality $\lambda$ )	Model dif- ferences	$\chi^2$ dif- ference (df)	Critical $\chi^2$ (non- centrality $\lambda$ )
$M_1 - M_2$	39.51 (11)	21.3514*** (.957)	$M_3 - M_4$	9.1 (1)	7.1476*** (.1056)
$M_1 - M_3$	34.25 (10)	19.97466** (.932)	$M_3 - M_5$	9.19 (1)	4.4285*** (.157)
$M_1 - M_4$	43.85 (11)	22.1060*** (1.4)	$M_3 - M_6$	10.48 (1)	6.2139*** (.712)
$M_1 - M_5$	43.44 (11)	20.93907*** (.718)	$M_4 - M_5$	0.09 (0)	3.6450 (0.561)
$M_1 - M_6$	44.73 (11)	21.7702*** (1.202)	$M_4 - M_6$	0.38 (0)	4.6322 (.816)
$M_3 - M_2$	5.26 (1)	5.8771 (.597)	$M_5 - M_6$	0.29 (0)	3.1624 (.459)
***=difference in models is significant when Critical $\chi^2 < \chi^2$ difference					

Source : Computations by the author

We use the method described in MacCallum, Browne, and Cai (2006) to calculate the  $\chi^2$  critical values. With these values we test the small difference hypothesis between the models. Results (table 4.6) show there is a small difference between models 3 and 2, 4 and 5, 4 and 6, and 5 and 6. Model number 1 has less quality than the 5 others, hence we must reject it. Model number 2 has less quality than model number 3. Models number 4, 5 and 6 present almost the same qualities. Then the best fitting model is among models 3, 4, 5 and 6. We need to do further analyses to determine which model is the best among the six estimated models.

The reduced-form  $R^2$  are printed for endogenous variables, which are regressed on the exogenous variables only in nonrecursive models. These values explain the contribution for only one variable, because all direct effects of disturbances on their respective endogenous variables are blocked and simultaneously they removed any contribution



TABLE 2.7: Reduced-form  $R^2$  for all the variables in all of the six *a priori* estimated models

$R^2$	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>GDP</i>	NA	NA	0.938	0.824	0.824	0.829
<i>CPI</i>	0.506	0.502	0.582	0.259	0.259	0.260
<i>UR</i>	0.662	0.662	0.671	0.970	0.970	0.966
<i>NDC</i>	0.328	0.328	0.275	0.673	0.673	0.665
<i>CC</i>	0.775	0.776	-	-	-	-
<i>ER</i>	0.172	0.144	0.068	0.291	0.291	0.289
<i>RIR</i>	0.745	0.723	0.560	0.812	0.812	0.811
<i>NTED</i>	0.896	0.951	NA	0.804	0.804	0.808
<i>FDI</i>	0.229	-	-	-	-	-
<i>BOP</i>	-	-	0.894	0.753	0.753	-
<i>Tax</i>	0.836	0.836	0.367	0.961	0.961	0.958
<i>FCE</i>	0.990	0.990	NA	0.869	0.869	0.869
<i>GCE</i>	0.193	0.193	0.006	0.460	0.460	0.465
<i>M2</i>	0.037	0.037	0.176	0.007	0.007	0.052

Source : author's calculation

to all other endogenous variables (Hayduk, 2006). The proportions of explained variance estimated by the  $R^2$  of the sixth model are way more interesting than the other models. According to results (table 4.7), only consumer price index and the growth of circulating money do not explain well their respective endogenous variable, contrary to the others. Between models 5 and 6 almost all the explained variances are equal except for the monetary variable; hence model 5 is also interesting to fit the data than model 6. But despite the weakness explanation variance provided by model 6 for the variable  $M_2$ , it is approximately 7.4 higher than the explained variance of model 5.

### 2.5.3 Analysis Issues

Among the indicators of the informal economy, unemployment rate and consumer price index reflect an inverse relationship with informal economy in all the six estimated models (table 2.5). Furthermore, the unemployment rate with minimum coefficients equal to -.794 (table 2.5) present higher coefficients than the consumer price index whose maximum coefficient is -.788 (table 2.5). When unemployment rate decreases informal economy increases that confirms the classic empirical assumption that advocates the less individuals are employed the more they joined informal activities. When the consumer price index goes up informal activities go down. Because when the consumer price index increases individuals need more liquidity (currency) to buy the same goods, hence, individuals have less liquidity (currency) to undertake such activities.

Net domestic credit and control of corruption evolve in the same direction as informal activities. State bribery reflects informal activities. The comportment of this indicator is coherent to the empirical assumption that stipulates smaller informal economy correspond to smaller state bribery (Johnson, Kaufmann and Zoido-Lobaton, 1998; Hindriks et al., 1999).

Allocate domestic credit contributes to increase informal activities; this means the most part of the domestic credit is going to informal enterprises. This situation is easy to understand because credit to informal enterprises provides more interest incomes because the more the credit is at risk the higher is the interest rate. When one of the models tends to reflect the informal economy through the growth of circulating currency we find the same negative sign ?? That means the less money circulates in the economy, the more people undertakes informal activities. Comparing all the six models, we observe that direct tax revenues, final consumption expenditures, and government consumption expenses, which is a proxy of indirect tax revenues, influenced positively the growth of the informal economy. That means when taxes and consumptions raised, informal economy increased also. That confirms the view of many theorists of the informal economy (Schneider, 2007). During the last twenty years, many taxes were reduced considerably, but these measures concerned import taxes only. Hence, the main tax revenues were levied from income, that reduces investment and by the same token also reduce employment opportunities. Government consumption expenditures also including external financing, that means informal economy growth and external financing evolve in the same direction (less external financing implies decrease shadow economy growth).

While the balance of payments, the circulating currency growth, the exchange rate (gourde per U.S. dollar) and the real interest rate have a negative impact on shadow economy growth (table 2.5). These results mean globally monetary policy impacts negatively informal economy growth. The balance of payment becomes the more and more negative through the years. However, if the balance of payments was positive (more exports than imports) it would reduce considerably the shadow economy growth. The exchange rate increases continuously, therefore, the informal economy growth should be altered because people would need more local currency units to import goods (the main source of informality). This could be explained by the inexistence of an employment policy. The fact that the balance of payments is negative so we import more than we export, hence a growing exchange rate coupled with the astonishing decrease of local production leads to poverty. Hence, the inexistence of employment policy drains people in informal activities. The real interest rate decreases considerably over time what contributes to increase shadow activities. Because higher real interest rates are an incentive for people

to make deposits of their money instead undertake informal activities. When real interest rates decrease people have made arbitrage between interest rate and benefits from informal activities and then they chose to invest in informal activities, which appear to provide a higher return on investment.

The Foreign direct investment has positive effect on the shadow economy. When the exchange rate is the reference variable for analysing monetary policy, the net transfers on external debt influence informal economy positively. Hence the more indebted is the country, the more important become the transfers on external debt, so the more informal activities grow. When the real interest rate is the marker the net transfers on external debt has negative impact on shadow economy growth. But the coefficients of the net transfers on external debt are way higher for the positive impact (minimum coefficient is 1.664) than for the negative impact where the maximum coefficient is -1.45 (table 2.5). That means the more the government acquitted its international debt the more informal economy grows rapidly. In all the six models the net transfers on external debt have the most important weight in the informal economy growth.

The commercial balance influences significantly the informal economy growth. When fiscal authorities drop tariff and eliminate the nontariff barriers, importation rises due to the weakness of local production what leads to deficit commercial balance. The commercial balance influences the informal economy and leads informal growth in the opposite direction. So a negative commercial balance leads to a growing shadow economy. Net domestic credit grew during the last 20 years and has impacted informal economic growth in the same direction: the most part of the domestic credit goes to the import sector. Therefore, the domestic credit serves to inflate the trade deficit, hence the larger is the deficit of the commercial balance the greater is the informal economy growth.

The exchange rate evolves contrary to the informal economy growth. In the Haitian dollarized economy (the local currency (gourde) is partially anchored on a foreign currency which is the U.S. dollar) while the national production falls exaggeratedly the exchange rate keeps growing. This anchorage brings to the Haiti's economy all the exogenous shock took from the foreign currency, because the foreign currency vulnerability increases due to the mismatch between the foreign currency liquid liabilities and the liquid foreign assets (IMF, 2005). That implies a growing inflation level and a drop of purchase power of the local economic agents. Therefore, when the exchange rate rises, inflation increases and purchase power go down what brings up informal economic growth. That explains all the positive fluctuations of the U.S. dollar have a negative effect on the informal economy.

One of the key decisions of the liberalization policies was the declining of the real interest rate. This measure obliges individuals

that have made a limited deposit on the banks to make arbitration between the gain that the real interest rate will provide and the return of an informal investment. These very limited capitals are unable to start a modern business, then, people invest them in informal enterprises. For this reason, real interest rate has the contrary effect in shadow economy growth.

Avoidance of accumulation of an unsustainable burden of external debt was also one of the main objectives of the deregulation policy. The government has incurred additional debts and has faced many socioeconomic problems, particularly the accumulation of substantial arrears on external debt between 1991 and 1994. The International Monetary Fund (IMF) demands that authorities restore solvency by restructuring unsustainable fiscal imbalances. Hence, the government increases significantly disbursement for external debt during the last decade. Due to low and volatile levels of revenues, the government has significant difficulties to provide adequate social services and social expenditures fallen. Coupled with the very limited size of public service employment, these situations lead to informal economic continuous growth.

During the years that follow the liberalization trade policies, the budget system becomes rigid with some fix objectives. Government consumption expenditures reflect the size of government; they affect economic growth in different manners. First, its impact can be perceived through their effects on the fiscal balance. One of these signs is the impact of automatic fiscal stabilizers in a rigid budget system because of a large proportion of nondiscretionary expenses such as salaries, interest payments, social entitlements and subsidies (World Bank, 2007). This reflects a negative association between economic growth and government consumption expenditures. It has been proven that informal economy and economic growth are countercyclical (Aspilaire, 2014). The negative relation between government expenditures and growth explains that informal economy and government expenditures evolve in the same direction.

The relationship between final consumption expenditures and economic growth is important because of the well-established notion of consumption led growth. Higher consumptions lead to economic growth, contrary to this theory, we observe the increase in final consumption expenditures has a positive effect on informal economy which is countercyclical. Growing consumption expenditures influence positively informal economy growth. It can be explained by the fact that the last 20 years of the economy are characterized by a continuous declining sector of production.

## 2.6 Conclusion

All the macroeconomic variables related to budget, fiscal and monetary policies and economic growth impact the informal economy

growth. The fiscal and budgetary variables, except tax revenues, which was the reference variable, have impacts that promote informal economy growth. The monetary variables and the shadow economy growth have a contrasted evolution. Because of a dollarized, unproductive and not sufficiently prepared economy to deal with these enormous changes the deregulatory policy creates some perverse effects (growing inflation, continuous increase of importation and local currency depreciation) on other concerned variables which also influence the economic structure which annihilate these negative effects on informal economy.

Through the tackling of the neoliberal empirical assumption that advocates that flexible fiscal policy and rigorous budgetary policy can reduce informality growth two main lessons should be deducted. In 1999 the last stage of the liberalization trade policy ended, almost all the taxes on import products fall down, a new fiscal policy took place to facilitate foreign investment and reduce fraud incentive. More than 10 years after the application of the neoliberal fiscal policy, informal economy keep growing and represent a share of the GDP that increases continually. The rigorous budget policy recommended by IMF consisting in reducing banking credit for public institutions, lowering interest rates, regularizing payment of external debts and diminish public expenditures and investment was applied strictly. Against all odds the informal economy presented its best growth period after the application of the neoliberal budget policy. Hence the neoliberal empirical assumption towards informal economy must be rejected in the case of Haiti.

The re-evaluation of the traditional main causes of informal activities such as higher taxes and social security burdens (Schneider, Buehn and Montenegro, 2010), trade barriers and regulations (Johnson, Kaufmann and Zoido-Lobaton, 1998) suggest two surprising approaches.

In one hand, despite the model fixed tax revenues as the reference variable of fiscal policy, the model keeps recording significant growth of the informal economy. The data suggest that tax revenues do not constitute one of the main causes of informal economy growth.

Are trade barriers and regulation determinant factors in shadow economy growth? National authorities, under the IMF demands, eliminate the nontariff barriers, reduce dramatically customs tariffs. In 1995, all remaining import restrictions on agricultural commodities were removed, from maximum tariff rates around 50% to 0-3% maximum tariff rate, for the most important basic products (rice, sugar, flour, cement). The liberalization trade policy drops more than a half of the 1,600 tariff lines near a zero rate. Budget austerity policy reduces drastically state intervention in the market. Despite all these changes and the tariffs suppression, informal economy keep increasing more than 10 years later. The time series data do not confirm trade barriers and regulation are important factors of consequence of

shadow economy growth.

These conclusions remind how it is difficult to apprehend informal economy through uniform assumptions (empirical or theoretical) regardless its main causes or consequences. Probably some causes and consequences of informal economy growth can be evident in wealthy countries and mature economies and appear to be insignificant as causes of informal economy expanding developing countries or emerging economy.



## Chapter 3

# Employment, Poverty and Human Capital: Is informal employment in Haiti a result of exclusion?

### 3.1 Introduction

This chapter analyses mostly the relationship between informal employment, poverty and human capital accumulation using a Mincer wage model, upon a sample of 5,228 active individuals from a national survey (2007), estimated by an ordinal multinomial logit model. Demographic variables, geographic setting, working hours, and education are the main factors that explain the access to informal labour market. Results show that poor workers are those with lower human capital accumulation and they regroup salaried, trainee and self-employed. Informal employment does not imply automatically poverty and low human capital accumulation. As the formal job market, informal job market includes poor and non-poor workers, low human capital accumulated employee and largely valorised employee. The estimates of the Mincer earnings models show large discrepancies among the revenues of the informal workers as regards their demographic and economic characteristics make the segmentation of informal employment a workable assumption.

Understanding the determinants of wages and earnings in the informal employment can help policy makers develop strategies in order to promote wealth, to overcome poverty and eventually to put countries on a path that increases growth and prosperity. Human capital theory shows that one of the best ways to promote wealth and reduce poverty is the incentive for individual investments in human capital. These investments will raise earnings and diminish deprivation because even low-skilled workers can take advantage of training. These insights led to viable policies increasing overall wealth. Many researchers (Barro and Sala-i-Martin, 2004) have demonstrated the very implications of human capital theory for earnings, employment and economic growth. The literature considers as informal employment all kinds of jobs, which are not protected by labour legislation,



not subject to taxation and social protection fees paid to the workers, own-account workers or employers with unregistered activity (OIT, 1993 and 2002). In Haiti 57.1% of the employment are informal (IHSI and CELADE, 2007) which explains its importance in the economy.

Another main reason to apply human capital theory to the informal employment is to evaluate the main approaches that link informal employment and poverty. According to Fields (1990) the informal labour market has a dual side. An attractive side for people with special characteristics who hope higher earnings than in the formal market, those people choose to work in the informal market. A less attractive side where people without special characteristics work for low wages. Another theory presents informal work as a disadvantaged sector of a labour market segmented by rigidities in the official economy (Harris and Todaro, 1970), where people are underemployed in micro-firms characterized by poor capacity for capital accumulation, irregular work conditions and very low rates of remuneration. Information asymmetries can oblige inexperienced young workers to choose the informal market for gaining experience to join the formal market later (Perry et al., 2007). This implies that mostly people who work in the informal employment market have chosen voluntarily to be in informality, but they have sufficient characteristics to move to the formal sector, but refuse (Perry et al., 2007), because probably they would not be in better jobs in the formal market than formal workers with identical characteristics. Because the informal market can provide jobs that are equally valorised than those that exist in the formal sector, workers compare utilities, not only earnings in choosing jobs in formal and informal sectors from a variety of work related to their skills and characteristics (Rosen, 1981; Heckman and Sedlack, 1985; Carneiro, Heckman, and Vytlačil, 2005). By comparing the advantages related to their entrepreneurial relative skills in both sectors, young and unskilled middle-age workers chose to enter the informal market to compensate for their lack of skills (Lucas, 1978). Informal employment consolidates, rather than reduces, the disparities produced by formal capitalist employment (Pahl, 1984; Williams and Windebank, 2003). Many studies have been conducted to analyse the determinants of the informal market using the wages of the workers, both in Europe (Kolev, 2000), in Africa (Adair and Bellache, 2012) and in North America (Dougherty and Escobar, 2013). In the Caribbean region, few studies have made the link between informality and poverty by using human capital theory, although it should be mentioned the study about Trinidad and Tobago (Sookram and Watson, 2008).

As for Haiti, a recent study analyses the determinants of informal employment using probit models (World Bank and ONPES, 2014). But there is no study that concerned informal economy in the Caribbean

countries that highlights the relationship between informal employment, poverty and human capital accumulation. Our chapter estimates earnings classification regarding the human capital accumulation and socio-demographic variables thanks to a Mincer wage model through an ordinal multinomial logit equation. In addition, this chapter takes the analysis one step further by assessing the relationship between informal employment, poverty and human development.

We further develop the theories that make the link between informality and poverty in various ways. These various theories raise the following questions: Do workers choose voluntarily to enter the informal employment market? Do informal workers get wages that match their characteristics? Are jobs generated in the informal market equally valued by workers as those in the formal sector? Do informal workers have special characteristics (marginalized, youth, inexperienced, low wages, and educational attainment) that are consistent with the informal market? The aforesaid theoretical approaches answer to these questions in different ways; we separate and test the applications of these explanations by using the results from the econometric model that we conceive.

This chapter estimates the relationship between wages and human development (individual returns of education and experience) on the informal market. One of the contributions of this chapter is to analyse the determinants of wages of the informal workers and examine the dynamics of informal employment. This arises from the following questions: is informal employment a generator of poverty? Are the levels of the revenues explained significantly by the workers' socioeconomic characteristics? The main purpose of this chapter is to analyse the relationship between informal employment, poverty and human capital accumulation by modelling the data and confronting the results with the various theories. We use a sample of 5,228 individuals from a national survey (IHSI, 2010) upon informal economy and employment, which contains 11,688 valid observations, including 1,716 observations of agricultural workers and 4,130 involuntary inactive by computing different ordinal multinomial logit models.

Section 1 presents the stylized facts and the linkages between poverty and informal employment. Section 2 is dedicated to the theoretical models of human capital and the contrasting theories of the informal workforce. In section 3, we present the data and the multinomial logit equations used to estimate the Mincer wage model. Section 4 analyses the determinants of the informal revenues and the relationship between informal employment, poverty and human capital accumulation. Section 5 concludes.

## 3.2 Stylized facts of poverty and the informal employment in Haiti

The issue of poverty is one of the most challenging socioeconomic problems in developing economies. Since the informal economy has been popularized in the literature, researchers have linked poverty to informal activities. For ones, informal economy represents the main path to social exclusion and poverty. In the Haitian economy, the first official measure of poverty took place during the 2000's.

According to ILO (1993) the employed persons in the informal sector should be sub-divided into two categories: those who are exclusively employed in the informal sector, and those who work both in the formal sector and in the informal sector. But this definition adopted during the 15<sup>th</sup> ICLS presents some difficulties of classification. Statistical surveys may encounter some bias of answers, firstly, casual self-employed person or very small-scale activities can be not reported even though their activity corresponds to the enterprise-based definition, and secondly statistics regarding informal sector may be influenced by errors in grouping some employed persons by status in employment (ILO, 2003). For all that we adopt the definition of the informal employment provided by ILO (2003), the informal employment is the total number of informal jobs without distinction of sector (formal sector firms, informal sector enterprises or households activities) during a given reference period (ILO, 2003).

The informal employment includes own-account workers and persons employed in their own informal sector activity, all contributing family workers, any individual who holds informal jobs in formal or informal sector firms, paid domestic workers enrolled by households, members of informal producer cooperatives, and person who produce goods exclusively for own final consumption (ILO, 2003). In Haiti the cooperatives have a particular status, they cannot be registered as formal or informal enterprise (IHSI, 2010).

### 3.2.1 Evolution of poverty in Haiti

Many studies regarding poverty in Haiti have been conducted; they address the question in the different aspects. Unequivocally, the concept of poverty contains different dimensions. Poverty could be monetary (income and consumption expenditures) and non-monetary (access to education, nutritional needs, access to healthcare, protection from violence, adequate housing, access to sanitation, to energy and others). Some of those studies address the monetary aspect (Montas, 2002; Accion, 2003; Egset and Sletten, 2004) and others address the non-monetary aspect (World bank, 1998; CNSA, 2002; UNDP, 2002). Furthermore, the national statistics organism has also conducted different surveys and studies regarding the poverty situation (IHSI, 2001, 2003).

Using the explanation of Montas (2005), we adopt the measure of poverty based on incomes or consumption levels. An individual is considered as poor when his income level is unable to cover some minimum level of expenditures to meet basic needs. This minimum level is known as the poverty line. A poverty line of U.S. 235 dollars/year has been calculated for 1986/87 (Pedersen and Lakewood, 2001). The same organism has calculated a poverty line for 1999/00 which was established at U.S. 204 dollars. Montas (2005) proposed a poverty line for 2000 that is equivalent to U.S. 344 dollars.

In 1987, the proportion of poor was approximately around 59.6% (Pedersen and Lakewood, 2001) and 65% (IBD) while in the rural region, it was around 60-65% (IBD) and 72.8% (Pedersen and Lakewood, 2001). In 1989, it was estimated around 69% (CTPEA) and the World Bank estimates approximately 81% of persons living in rural regions were poor in 1996. Over the period 1995-99 CNUCED has evaluated the poverty proportion at 79.9%. In 2001, the national statistics organism through a survey (Enquête sur les Conditions de Vie en Haïti) has found 76% of individuals were poor while the poor in rural agglomeration represented 85.2%. A nonofficial poverty line was calculated in 2001 from two surveys "Enquêtes sur le Budget de Consommation des Ménages", phases 1 and 2 (IHSI, 1986/87 and 1999/00) which contain data on household consumption and expenditure. This study (Pedersen and Lakewood, 2001) showed a fall in consumption poverty from 59.6% to 48% in 1999/00 but these results were contested (Montas, 2005) due to a methodological weakness. Although over the period 2003-06 the national poverty line was measured based on the international poverty lines of the period that means \$1 and \$2 PPP per person per day. Thereupon, the extreme poverty rate found using data from 2001 survey ranged from 48.9% (Verner, 2005) to 53.9% (World Bank and SEDLAC, 2005/06) and 55% (UNDP, 2003).

TABLE 3.1: Employment situation and consumption expenditures of the households in gourdes

Sector employment	Urban			Rural		
	Non-poor	poor	Extreme poor	Non-poor	poor	Extreme poor
Informal sector, %	69.3	74.3	56.7	34.5	25.8	15.7
Formal sector, %	24.6	9.5	2.2	4.7	4.0	1.6
Average percapita consumption	60.989	23.360	11.322	52.657	21.250	10.086

Source: ONPES and World Bank, 2014

All data (table 3.1) are from the "Enquête sur les Conditions de Vie des Ménages après le Séisme" (ECVMAS 2012) which is a postearthquake

household living conditions survey. This survey defines the informal sector as the unincorporated enterprises (not in the primary sector) that are not legally registered and/or do not keep formal accounting (ONPES and World Bank, 2014). This survey also defines informal workers as all contributing workers from household enterprises, all independent workers in the informal sector and all employees (not in the agricultural sector) without legal contracts and not benefiting from social protection (ONPES and World Bank, 2014). Records (table 3.1) show that 57% of non-poor individuals are in informal employment, about 45% of poor individuals work in the informal sector and only 21% of extremely poor individuals are in the informal market (ONPES and World Bank, 2014). Results also show that 74% of urban non-poor work in the informal sector while 56% of poor urban individuals work in the informal market and 34% of the extremely poor work also in informal activities (ONPES and World Bank, 2014). Globally, about 75% of people living in rural regions are poor, but about 41% of urban individuals are poor. Furthermore, 66% of the poor live in rural regions and only 33% of the poor are living in urban agglomeration (ONPES and World Bank, 2014). The average per capita consumption of urban non-poor individuals is about 61,000 units of local currency (gourde) while the average per capita for urban poor individual is about 23,000 of gourdes and the extremely poor living in urban agglomeration consume on average 11,000 of gourdes (ONPES and World Bank, 2014). Accordingly, there is a huge gap between average per capita of non-poor and poor and between poor and extremely poor people. Human capital is the most abundant asset of the poor paradoxically they have less access to education than non-poor. With more than 58% of poor, poverty is one of the biggest socioeconomic challenges of the Haiti's economy. Nevertheless, poverty is expanding in Haiti but it is more severe in rural regions than in urban spaces. Inequality is huge in terms of income even among poor and extremely poor. The most important employment provider for the poor, including extremely poor is the informal job market where 66% of them are involved in.

### **3.2.2 Public policies against poverty**

The wide spreading of poverty and inequality during the decade 1990-2000 obliges policy makers to address this socioeconomic problem. To face this problem, two main documents were elaborated by the public policy makers, between 2007 and 2010. The first one was the National Strategy Document for Growth and Poverty Reduction (DSNCRP) also called the Growth and Poverty Reduction Strategy Paper (MPCE, 2007), and the Action Plan for National Recovery and Development of Haiti (PARDH) in 2010. Two other strategic documents were adopted after the 2010 earthquake to decrease the poverty

rate: The Strategic Plan of Development of Haiti and the first Triennial Investment Program (TIP) 2014-2016, both elaborated in 2012.

During the beginning of the 2000's the Haiti economy faced a dual challenge of growing poverty and inequality. To fight this tendency, policy makers adopted in 2007 the first official document which aims to fight against poverty: The Growth and Poverty Reduction Strategy Paper (MPCE, 2007). This document was built in three fundamental pillars, but we only present those to which we are interesting. The first pillar is based on using agriculture and rural development, tourism, and infrastructure as growth vectors. Develop the rural regions, while guaranteeing the protection of the environment, promote the diversification of the activities which are generator of revenues, sustain the progress accomplished and ensure a less inequitable distribution of the revenues generated. Capitalize on tourism as a growing industry in the Caribbean region. The government intends to take control of urban and interurban transport by prioritizing the development of regions with high potentiality of development, connecting all the regions throughout the territory and according attention to the protection of the environment. The second pillar concerns human development by according priority to basic services such as education and training, health, and water and sanitation, bring a solution and help for persons with disabilities, childhood poverty, and HIV/AIDS, and encourage youth integration, and gender equity. Bring some amelioration in the global educational system (teachers, schools, educational programs and others). In the health system, the plan aims to draft a national health policy, establish the procedures for inspection and control health activities, and finally modernize the health information system (MPCE, 2007). To improve the low level of water and sanitation access, ancient organisms, and entities in charge of these services will be replaced by regional offices but management through private institution are envisaged. With an important proportion of people with disabilities, approximately 800,000 (MPCE, 2007), the document proposes better prevention of incapacity through the public health system (vaccination, early detection, creation of medical rehabilitation departments in the main hospitals and home for the disabled in different departmental administrative districts) and nutritional programs. To take in charge the childhood poverty the policy makers will remove obstacles to investment that benefit young person, public support to families and communities and encouraging the participation. Young people represent more than 50% of the overall population so the reform aims to improve their living conditions by fostering their economic and social-cultural integration, and to create complexes for physical and sporting activities. Despite the promising approaches of the DSNCRP where the phases of application were divided into two phases, the acceleration of the economic growth and the control of social development awaited are far away. The earthquake has changed many parameters in the second phase of this plan,

and a new document has been redacted in 2010 the Action Plan for National Recovery and Development of Haiti (PARDH). The objective of the PARDH was to bring immediate responses to the decreasing situation of people after the 2010 earthquake, but it also aims to tackle the main structural causes of poverty and under-development in Haiti. The PARDH was constituted of four main pillars, among which we present two of the four pillars that present interest for us. The Economic rebuilding conceives to conduct the agricultural sector in the modernization era to protect families against food scarcity and to avoid hunger; the development of the professional construction sector to protect households against earthquake and hurricanes, the promotion of the manufacturing industries and plan the development of the tourist sector. The social rebuilding pillar aims to guarantee access to education for all children, to build vocational courses and university to make the offer adequate to the demands of economic modernization; to give priority to effective health systems that can give minimum overage in the entire country; finally, it conceives a social protection system for the less fortunate workers and families. The Action Plan for National Recovery and Development of Haiti (PARDH) was elaborated to fight all the aspects of multidimensional poverty. As we stated previously, poverty has a monetary aspect considering income and consumption expenditures and a non-monetary aspect which concerns access to education, nutritional needs, access to healthcare, social protection, adequate housing, access to sanitation, and others. The PARDH aims to provide food security, access to education for every child, resistant and protected houses, access to healthcare and social protection for the most vulnerable workers and families.

### **3.2.3 Linkage between informality and poverty**

Informal economy becomes more significant in developing countries since its importance has been acknowledged in the literature. In Haiti, from 2000 to 2010 informal economy represents between 55.4% and 70.1% of the GDP (Schneider et al., 2010; Aspilaire, 2014). Approximately 55% of employed persons are working in the informal sector; informal private enterprises represent 81% of business in the metropolitan sphere (IHSI, 2010). Despite of the conjoint growth the linkage between informality and poverty still concerns stylized facts. In the informal economy literature, linkages between informality and poverty are controversial. Nevertheless, we can classify the different ideas into two groups that we call the positive and the negative according to their position about informal economy. According to the "negative" thinkers, informal work is a transitory state and a marketplace where workers gain only subsistence earnings (Lewis, 1954; Harris and Todaro, 1970) and are working in sweatshop like conditions without a significant coverage of social protection (Sassen,

1997). For these theorists, the informal economy is a leftover of pre-capitalism constituted of marginalized individuals (Chen et al., 2004; Williams and Round, 2008). Informality is a new form of activity emerging in late capitalism as a creative response to the incapacity of the State to meet the basic needs of the impoverished masses in a deregulated open world economy (de Soto, 1989; Hudson, 2004). Consequently, the informal economy becomes a generator of poverty, hence informality has negative impact on growth and poverty reduction. The positive view argues that not all the informal workers are concerned with poverty and it exists a positive relationship between informal employment and poverty alleviation (Orlando, 2001; Cling et al., 2010). These theorists advocate that a significant part of household income is earned by working in the informality, hence, without the informal employment, poverty would be more intensive. Furthermore, some authors show significant linkages between poverty intensity and informal employment, they also find that the most important portion of the revenues of the extreme poverty families is earned from the informal employment market (Cartaya, 1994). Individuals in poorer situation and lower-income groups are more likely to work in the informal market (Agarwal and Dhakal, 2010), this means informal activities are a source of subsistence for poor and marginalized people (Leonard, 1994,1998, Williams and Windebank, 2001), even if wealthy groups tend to purchase a greater part of informal labour (Williams, 2008; Williams and Winderbank, 2002). For some authors, informal workers have special characteristics, poor and marginalized, youth, inexperienced, low wages, and low-educational level. For others, workers choose informal jobs related to their skills and characteristics. Where workers gain only subsistence earnings and are working in sweatshop like conditions without a significant coverage of social protection. For some theorists, the informal market is constituted of marginalized individuals. All these approaches make linkages between poverty and informal employment and they also bring evidence on the role of human capital in the determination of the poverty of individuals who are working in the informal market. Consequently, before we analyse the relationship between poverty, human capital and informal employment, we make a literature review of human capital theory in the following section.

### **3.3 Literature review of human capital theory**

Labour economics define human capital as a set of fundamental characteristics such as skills, education, motivation, work attitudes, which are prone to increase worker's productivity. According to the labour economics literature, there are five main approaches of human capital.



### 3.3.1 The main theoretical models

The first approach (Becker, 1975) is twofold. On the one hand, it argues that demand conditions are similar for everyone and supply conditions are the only cause of inequality (egalitarian aspect). This aspect assumes a similar capacity for everyone to benefit from investment in human capital, particularly education, it considers that the distributions of investment in human capital are uniform basically. In that view human capital is a one-dimensional object like a package of abilities and knowledge of the process.

On the other hand, it assumes that supply conditions are identical, but differences among workers come from demand conditions ("elite" aspect). The elite aspect makes the assumption that opportunities are identical for workers, but the differences in the distributions of earnings and investments in human capital are from demand conditions, and they are explained by the differences in the capacity of everyone to benefit from investment in human capital (Becker, 1962, 1964, and 1975). The second approach (Gardner, 1983) considers human capital as a multidimensional of many abilities and skills that is a matrix of knowledge. For Gardner, some people can have some good skills and abilities in some dimensions and being very unskilled in some other dimensions. A third approach (Nelson and Phelps, 1966; Schultz, 1961, 1962) contends that human capital is a factor that should take place within disequilibrium situations or adapt to an ongoing changing environment. The fourth approach (Bowles and Gintis, 1976) argues that human capital refers to the capacity of workers to adapt to organizations or hierarchical structures wherein they are working. The last one (Spence, 1973), takes human capital as a factor measuring the abilities and skills of workers without considering these as independent characteristics that are important in the production process. The three basic theoretical models were designed by Mincer (1958, 1962, and 1974), Becker (1962, 1964, and 1975) and Ben-Porath (1967). In addition, Acemoglu and Autor, (2011) introduce the so-called separation theorem in human capital investment in order to separate human capital stock and supply decisions from consumption decisions.

#### Becker's Model

The first approach (Becker, 1975) is twofold. On the one hand, it argues that demand conditions are similar for everyone and supply conditions are the only cause of inequality (egalitarian aspect) therefore same investment (in human capital) should provide identical earnings for each investor. On the other hand, it assumes that supply conditions are identical, but differences among workers come from demand conditions ("elite" aspect). Becker concluded with the fact that both supply and demand conditions varied and different investment in human capital for the same amount could provide more earnings

for some and less earnings for others. This model can be used to measure the adjusted rates and then to demonstrate the contribution of schooling to explain the distribution of earnings.

### Ben-Porath's Model

Becker (1962, 1964) and Mincer (1958, 1962) bring a nuanced view of the life cycle of earnings brought by Becker (1962, 1964) and Mincer (1958, 1962) which makes a connection between the life cycle of earnings and time profile of investment in human capital. They argue that people have foregone current earnings in their youth to make their main investments in themselves, because they have a longer period to receive a return on their investments. From this view Ben-Porath (1967) shows how the determination of the optimal path of investment is influenced by the function of production composed of people own abilities (innate or acquired), market constraints and opportunities and the quality of co-operating inputs; he also shows how the various properties of the technology affect the life cycle of earnings. This model serves to qualitative analysis of the effects on behaviour of two types of parametric disturbances: changes in market prices and variations in the technical parameters consist of the variables appearing in the production function.

### The separation theorem

The Separation Theorem permit to consider human-capital decisions separately from consumption decisions (Wetzstein, 2013) supposes the partial equilibrium schooling decisions and establishes a basic model in continuous time. Assuming that in all the markets, there is never any arbitrage opportunity, the separation theorem demonstrates that schooling decisions maximizes the net present discounted value of people. Let  $\mu(c)$ , a function that satisfies some standards such as strictly increasing and strictly concave, be an instantaneous utility function of an individual with planning horizon of  $T$  (it is allowed that  $T = \infty$ ),  $\rho > 0$  the future discount rate,  $\nu \geq 0$  the constant flow rate of death (Acemoglu and Autor, 2011). The objective function of an individual with such utility function at time  $t = 0$  is given by:

$$\max \int_0^T \exp(-(\rho + \nu)t) \mu(c(t)) \quad (3.1)$$

It is supposed that this person has some innate human capital denoted by  $h(0) \geq 0$ , and the evolution over time of his human capital is given by the following differential equation:

$$\dot{h}(t) = G(t, h(t), s(t)) \quad (3.2)$$

Where  $s(t) \in S(t) \subset [0, 1]$  is the fraction of time that the individual spends for investments in schooling, and  $G : \mathbb{R}_+^2 \times [0, 1] \rightarrow \mathbb{R}_+$

determines the evolution of human capital as a function of time, the individual's accumulation of human capital and schooling decisions. It is assumed that the individual faces exogenous given prices for human capital and exogenous sequence of wage per unit of human capital. If  $w(t)$  is the non-human capital labour supplying to the market buy the individual than the exogenous sequence of wage per unit is given by  $[w(t)]_{t=0}^T$ . Assuming that the individual has one margin choice, which is between schooling and labour market, therefore his labour earnings at time  $t$  are expressed by:

$$W(t) = \omega(t)(1 - s(t))(h(t) + \omega(t)) \quad (3.3)$$

where  $1 - s(t)$  represents the time spent by the individual supplying labour to the market. The last assumption concerned the interest rate (denotes  $r$ ) on the individual savings that supposed to stay constant over time. Considering the fact that in the presence of perfect capital markets, the best human capital accumulation decisions are those that maximize the lifetime budget set of the individual (Acemoglu, 2013), using equation 3.3, the lifetime budget constraint of the person is:

$$\int_0^T \exp(-rt)c(t)dt \leq \int_0^T \exp(-rt)\omega(t)(1 - s(t))(h(t) + \omega(t))dt \quad (3.4)$$

The statement of the separation theorem can be as follows (Acemoglu and Autor, 2011): Suppose that the instantaneous utility function  $u(\cdot)$  is strictly increasing and strictly concave. Then the sequence  $[\hat{c}(t), \hat{s}(t), \hat{h}(t)]_{t=0}^T$  is a solution to the maximization of 3.1 subject to (3.2,  $s(t) \in S(t)$  and 3.4 if and only if  $[\hat{s}(t), \hat{h}(t)]_{t=0}^T$  maximizes

$$\int_0^T \exp(-rt)\omega(t)(1 - s(t))(h(t) + \omega(t))dt \quad (3.5)$$

subject to 3.2 and  $s(t) \in S(t)$ , and  $[\hat{c}(t)]_{t=0}^T$  maximizes 3.1 subject to 3.4 given  $[\hat{s}(t), \hat{h}(t)]_{t=0}^T$ . This is the separation of human capital stock and supply decisions from consumption decisions.

### 3.3.2 The Mincer Wage Model

For Mincer (1958) the economic analysis of personal income distribution should be based on the implication of the theory of rational choice.

Globally the Mincer model is based on three fundamental assumptions. First, it is assumed that the earnings of an individual with  $S$  years of schooling do not depend on age. Second, if there are no post school investments, the present values (PV) of lifetime incomes are the same across individual regardless of schooling. The third one assumes that the number of years spent at work is independent of the number of years of schooling. Denoting  $E(S, t)$  the earnings at

time  $t$  of an individual with  $S$  years of schooling, the calculation of the present values of earnings of an individual who enters the labour market after  $S$  years of schooling with  $r$  the discount rate of future earnings is given by:

$$V(s) = \int_s^R E(S, t) \exp(-rt) dt \quad (3.6)$$

Regarding the assumption that the earnings of an individual with  $S$  years of schooling do not depend on age,  $E(S, t)$  does not depend on time  $t$ , therefore it can be wrote  $E(S, t) = E(S)$  and:

$$\begin{aligned} V(s) &= \int_s^R E(S, t) \exp(-rt) dt = E(S, t) \int_s^R \exp(-rt) dt \\ &= E(S) (1/r) [\exp(-rS) - \exp(-rR)] = (E(S)/r) [\exp(-rS) - \exp(-rR)] \end{aligned}$$

The second fundamental assumption which state that  $V(s)$  should not depend on  $S$  and therefore  $V(s) = V$ . According to the third fundamental assumption, work lifetime is identical for everyone, which means for some  $T$ ,  $R = S + T$ . Using those elements of information:

$$\begin{aligned} V(s) &= V = (E(S)/r) [\exp(-rS) - \exp(-rS) \exp(-rT)] \\ \Leftrightarrow rV &= E(S) [\exp(-rS) - \exp(-rS) \exp(-rT)] \\ rV &= E(0) [1 - \exp(-rT)] = E(S) [\exp(-rS) - \exp(-rS) \exp(-rT)] \\ &= E(S) \exp(-rS) [1 - \exp(-rT)] \\ \Leftrightarrow E(0) &= E(S) \exp(-rS) \\ \Leftrightarrow \ln(E(S)) &= \ln(E(0)) + rS \end{aligned}$$

Where  $E(S)$  the hypothetical earnings of an individual who does not make any is post school investment in human capital after the completion of  $S$  years of schooling. The last equation is the schooling model of Mincer. Beside investments in human capital, age and experience have important impacts on earnings; due to that fact Mincer expands the schooling model to a post-school investment model (Mincer, 1974) to make the econometric analysis of the distribution of earnings. Mincer assumes that all individuals have the same abilities and identical opportunities to occupy any place in the labour market. The differences are based on the amount of training acquired on the market by workers; which should be compensated for the costs of their post-school investment, for that the present values of life-earnings must be equalized at the time that the worker chooses an employment. He also adds another assumption regarding the return to post-school investment (denoted  $r_t$ ) which is assumed to be equal in all periods  $t$ . It is assumed that an individual dedicates a fraction  $k$  of his time to invest in human capital and a fraction  $(1 - k)$  to actual work, but that fraction of time is a time function that Mincer writes as follows:

$$k(t) = k_0(1 - t/T) \quad (3.7)$$

$k_0$  is the investment ratio during the starting period of experience,  $t = 0$ ,  $T$  is the total period of positive net investment. The growth in earnings is given by the following differential equation:

$$\dot{E}(S, t) = r_t k(t) E(S, t) \Leftrightarrow \frac{\partial E(S, t)}{\partial t} = r_t k(t) E(S, t) \quad (3.8)$$

The resolution of this differential equation implies:

$$\frac{1}{E(S, t)} \frac{\partial E(S, t)}{\partial t} = r_t k(t) \Leftrightarrow \frac{\partial \ln E(S, t)}{\partial t} = r_t k(t) \quad (3.9)$$

$$\Leftrightarrow \int \frac{\partial \ln E(S, t)}{\partial t} = \int r_t k(t) \Leftrightarrow \ln E(S, t) = r_t \int k(t) \quad (3.10)$$

For the specific solution, we have:

$$\ln E(S, t) = C + r_t \int_0^t k(x) dx \quad (3.11)$$

The smaller bound of the integral has the value zero because the growth of earnings starts when school is over or school is forsaken. In the absence of post-school investment in human capital we had on the first model:

$$\ln E(S) = \ln E(0) + rS \quad (3.12)$$

Inserting into the differential equation  $k(x) = 0$ , we obtain:

$$\ln E(S, t) = C \quad (3.13)$$

Using 3.1 and 3.2 we get:

$$C = \ln E(0) + rS \quad (3.14)$$

and we have the solution of the differential equation:

$$\ln E(S, t) = \ln E(0) + rS + r_t \int_0^t k(x) dx \quad (3.15)$$

therefore,

$$\ln E(S, t) = \ln E(0) + rS + r_t k_0 \left[ x - \frac{x^2}{2T} \right]_0^t = \ln E(0) + rS + r_t k_0 t - r_t k_0 \frac{t^2}{2T} \quad (3.16)$$

If an individual works only part of the time:

$$Y(S, t) = (1 - k(t))E(S, t) \Leftrightarrow \ln Y(S, t) = \ln(1 - k(t)) + \ln E(S, t) \quad (3.17)$$

$$\ln(1 - k(t)) = \ln\left(1 - k_0\left(1 - \frac{t}{T}\right)\right) \quad (3.18)$$

Using Taylor expansion around  $t = T$

$$\ln\left(1 - k_0\left(1 - \frac{t}{T}\right)\right) = \ln(1) + \frac{K_0/T}{1}(t - T) - \frac{1}{2}\left(\frac{k_0}{T}\right)^2(t - T)^2$$

$$\begin{aligned}
&= \frac{K_0}{T}t - k_0 - \frac{1}{2}\left(\frac{k_0}{T}\right)^2t^2 - \frac{1}{2}\left(\frac{k_0}{T}\right)^2T^2 + \left(\frac{k_0}{T}\right)^2Tt \\
&= \frac{K_0}{T}t - k_0 - \frac{1}{2}\left(\frac{k_0}{T}\right)^2t^2 - \frac{1}{2}(k_0)^2 + \left(\frac{k_0}{T}\right)^2t \\
&= -k_0 - \frac{1}{2}(k_0)^2 + \left[\frac{K_0}{T} + \frac{(k_0)^2}{T}\right]t - \frac{1}{2}\left(\frac{k_0}{T}\right)^2t^2
\end{aligned}$$

Hence:

$$\ln Y(S, t) = \ln E(0) + rS + r_t k_0 t - r_t k_0 \frac{t^2}{2T} - k_0 - \frac{1}{2}(k_0)^2 + \left[\frac{K_0}{T} + \frac{(k_0)^2}{T}\right]t - \frac{1}{2}\left(\frac{k_0}{T}\right)^2t^2 \quad (3.19)$$

$$\ln Y(S, t) = [\ln E(0) - k_0 - \frac{1}{2}(k_0)^2] + rS + \left[r_t k_0 + \frac{K_0}{T} + \frac{(k_0)^2}{T}\right]t + \left[-\frac{1}{2}\left(\frac{k_0}{T}\right)^2 - \frac{r_t K_0}{2T}\right]t^2 \quad (3.20)$$

The last equation obtained above is the canonical Mincer equation by denoting:

$$\begin{aligned}
\beta_0 &= \ln E(0) - k_0 - \frac{1}{2}(k_0)^2 \\
\beta_1 &= r \\
\beta_2 &= r_t k_0 + \frac{K_0}{T} + \frac{(k_0)^2}{T} \\
\beta_3 &= -\frac{1}{2}\left(\frac{k_0}{T}\right)^2 - \frac{r_t K_0}{2T}
\end{aligned}$$

We obtain the canonical Mincer equation which has the following form:

$$\ln(Y_t) = \beta_0 + \beta_1 S + \beta_2 t + \beta_3 t^2 + \varepsilon_t \quad (3.21)$$

Mincer deduces three main conclusions from that model (Mincer, 1974). When schooling is symmetrically distributed, earnings are it implies a positively skewed. Only a strong negative skewed of schooling can avoid a positive skewed distribution of earnings. Second, the dispersion in the distribution of schooling dictates the way that the relative dispersion and skewness of the distribution of earnings. Finally, the larger of inequality and skewness in the earnings are influenced by the rate of returning to school.

### 3.4 Data and Econometric Approach

Our main goal is to analyse the relationship between earnings and some socio-demographic variables such as education, experience, training and working hours, some demographic variables such as age, marital status, and geographic setting (place of residence) through the Mincer model for informal workers.

### 3.4.1 Multinomial Logit Model

We denote the variable  $y$  as the revenues (range of values) of the informal workers taking on the values  $1, 2, \dots, J$  where  $J = 54$ , it constitutes the number of ranges of the values, let  $X$  (a vector of dimension  $1 \times k$  with first-element unity) denote the vector of conditioning variables (education, experience, experience squared, working hours, training, age, marital status, sector of activities and geographic setting) where  $X' = (1, x_1, x_2, \dots, x_k)$ . Our main interest is how ceteris paribus changes in the variables of the vector  $X$  affect the response probabilities of the revenues ( $Y$ ),  $P(y = j/X)$ ,  $j = 1, 2, \dots, J$ . We consider each range of revenues as exclusive alternatives the workers have their revenues in one of the designed ranges. Based on the individual characteristics of each worker, we can define a level of utility for each alternative. We assume that the utility function is the sum of the workers' characteristics (observed variables) and an error term (unobserved component). Denoting  $\beta_j$  the  $k * J$  vector of coefficients, we obtain the following utility levels for each alternative  $j \in 1, 2, \dots, J$ :

$$U_j = \beta_j^T x_j + \varepsilon_j \quad (3.22)$$

An alternative  $k$  will be chosen if and only if  $\forall k \neq j, U_k > U_j$  which implies the following  $J - 1$  conditions:

$$U_k - U_j = (V_k - V_j) + (\varepsilon_k - \varepsilon_j) > 0, \forall k \neq j$$

Due to the fact that  $\varepsilon_j$  are unobserved, the choices of the workers can only be modelled in probability terms. The  $J - 1$  conditions can be rewritten:  $\varepsilon_j < (V_k - V_j) + \varepsilon_k, \forall k \neq j$  Therefore, the general expression of the probability of choosing alternative  $k$  is:

$$P(y = U_k/\varepsilon_k) = P(U_k > U_1, \dots, U_k > U_J)$$

$$P(y = U_k/\varepsilon_k) = F_{j-1}(\varepsilon_1 \leq (V_k - V_1) + \varepsilon_k, \dots, \varepsilon_J \leq (V_k - V_1) + \varepsilon_k)$$

Where  $F_{J-1}$  is the multivariate distribution of  $J - 1$  error terms (all the  $\varepsilon_j, \forall j \neq k$ ). Three fundamental hypotheses are considered (McFadden, 1973). The first hypothesis assumes that the error terms are independent, that implies the univariate distribution of the unobserved terms can be written as follows:

$$P(U_k > U_j) = F_{j-1}(V_k - V_j + \varepsilon_k), \forall j \neq k$$

Where  $F_j$  represent the cumulative density of  $\varepsilon_j$ . The conditional and unconditional probabilities are respectively given by

$$P(y = U_k/\varepsilon_k) = \prod_{j \neq k} F_{j-1}(V_k - V_j + \varepsilon_k) \quad (3.23)$$

$$P(y = U_k/\varepsilon_k) = \int \prod_{j \neq k} F_{j-1}(V_k - V_j + \varepsilon_k) f(\varepsilon_k) d\varepsilon_k \quad (3.24)$$

The second hypothesis regards the distribution of the error terms. It assumes that each  $\varepsilon_j$  follows a Gumbel distribution (smallest extreme value distribution), by denoting  $\mu$  the location parameter and  $\gamma$  the scale parameter that implies:

$$f(\varepsilon_j) = \frac{1}{\gamma} e^{\frac{\varepsilon_j - \mu}{\gamma}} e^{-e^{\frac{\varepsilon_j - \mu}{\gamma}}} \quad (3.25)$$

and

$$P(\varepsilon_j \leq t) = \int_{-\infty}^t \frac{1}{\gamma} e^{\frac{\varepsilon_j - \mu}{\gamma}} e^{-e^{\frac{\varepsilon_j - \mu}{\gamma}}} d\varepsilon_j = e^{-e^{\frac{t - \mu}{\gamma}}} \quad (3.26)$$

When  $V_j$  contains an intercept the totality scale of utility of the error terms cannot be identified, therefore, without loss of generality we assume that  $\mu_j = 0, \forall j$ . Hence, only a natural normalization (one of the  $\gamma_j$  should be equal to 1) helps to identify the  $J - 1$  scale parameters. The homoskedasticity hypothesis, which is the third one implies that the scale parameter of the smallest extreme value distribution is identical for all the alternatives because the location parameter is not identified in the error terms. Previously, the normalisation fixed the value of one of the  $\gamma_j$  to 1, therefore the homoskedasticity assumes that  $\gamma_j = 1, \forall j \in 1, \dots, J$ . Denoting  $F$  the cumulative distribution function and  $f$  the probability density function of the standard Gumbel distribution ( $\mu = 0, \gamma_j = 1$ ), this hypothesis allows the conditional and unconditional probabilities wrote above to be expressed as follow:

$$P(y = U_k/\varepsilon_k) = \prod_{j \neq k} F(V_k - V_j + \varepsilon_k) \quad (3.27)$$

$$P(y = U_k/\varepsilon_k) = \int \prod_{j \neq k} F(V_k - V_j + \varepsilon_k) f(\varepsilon_k) d\varepsilon_k \quad (3.28)$$

The probability that an alternative  $k$  better than one another is given by:

$$P(\varepsilon_j < V_k - V_j + \varepsilon_k) = e^{e^{(V_k - V_j + \varepsilon_k)}} \quad (3.29)$$

According to the first hypothesis, the probability of a worker to choose alternative  $k$  is:

$$P(y = U_k/\varepsilon_k) = \prod_{j \neq k} e^{e^{(V_k - V_j + \varepsilon_k)}} \quad (3.30)$$

The unconditional probability is the expected value of the previous probability with respect to  $\varepsilon_k$ :



$$P(y = U_k) = P(y = U_k/\varepsilon_k) e^{-\varepsilon_k} e^{-e^{-\varepsilon_k}} d\varepsilon_k = \int_{-\infty}^{+\infty} \left( \prod_{j \neq k} e^{-e^{(V_k - V_j + \varepsilon_k)}} \right) e^{-\varepsilon_k} e^{-e^{-\varepsilon_k}} d\varepsilon_k \quad (3.31)$$

$$P(y = U_k) = \int_{-\infty}^{+\infty} e^{-\sum_j e^{(V_k - V_j + \varepsilon_k)}} e^{-e^{-\varepsilon_k}} d\varepsilon_k \quad (3.32)$$

Using the change of variable such as:

$$u = e^{-\varepsilon_k} \Rightarrow du = -e^{-\varepsilon_k} d\varepsilon_k$$

The unconditional probability is given by:

$$P(y = U_k) = \int_0^{+\infty} e^{-u \sum_j e^{(V_k - V_j)}} du \quad (3.33)$$

$$P(y = U_k) = \frac{e^{V_k}}{\sum_j e^{V_j}} \quad (3.34)$$

$$P(Y = j/X) = \frac{e^{V_k}}{\sum_j e^{V_j}} \quad (3.35)$$

Using  $V_j = \beta_j^l x_j$  the multinomial logit model is written as follow:

$$P(Y = j/X) = e^{(x_i \beta_j)} / \sum_{j=1}^J e^{x_i \beta_j} \quad (3.36)$$

For all the possible set of choices (represented by the equation  $\sum_{j=1}^J P(Y_i = j)$ ) the probabilities  $P(Y_i = j)$  sum to 1, then only  $J - 1$  of the probabilities can be calculated. Thereupon, as it is a system of  $J$  equations with  $J - 1$  independent unknowns the multinomial logit equation 3.36 meets a problem of indeterminateness. To solve this problem that allows an infinite number of values of  $\beta_j$  a convenient normalization is necessary. Practitioners set to zero all the parameters that represent the reference level of each categorical variable. If we denote  $j = J$  the reference level the identification condition of the equation is

$$\beta_1^j = \beta_2^j = \dots = \beta_k^j = 0$$

under such normalization equation 3.36 becomes

$$P(Y_i = j) = 1 / [1 + \sum_{j=1}^{J-1} e^{x_i \beta_j}] \quad (3.37)$$

$$P(Y_i = j) = e^{x_i \beta_j} / [1 + \sum_{j=1}^{J-1} e^{x_i \beta_j}] \quad (3.38)$$

Thereafter the normalization the probabilities are uniquely determined so the equation 3.38 constitutes a system of  $J - 1$  equations in the  $J - 1$  unknown probabilities. From equations 3.37 and 3.38 we obtain the logarithm of the ratio of the probability of outcome  $j = j$  to that of outcome  $j = k$  as follows

$$\log\left[\frac{P(Y_i = j)}{P(Y_i = k)}\right] = \sum_{m=1}^M (\beta_{jm} - \beta_{km})X_{im} \quad (3.39)$$

The logarithm of the risk-ratio, which is the logarithm of the ratio of the probability of outcome  $j$  to that of outcome  $k$ , is independent of the other choices. We must distinguish the risk ratio (RR) from the odds ratio (OR) which refer to the probability of an outcome divided by 1 (the sum of the probabilities over all the choices) minus the probability of that outcome. The relationship between RR and OR can be written as follows:

$$OR = \frac{RR_j P(Y_i = 1)}{1 - RR_j P(Y_i = k)} \quad (3.40)$$

Since the density of  $y$  given  $x$ , the best method to estimate of the multinomial logit model is the maximum likelihood method (Wooldridge, 2002). The conditional log likelihood is given by the following formula:

$$\ln(L(\beta)) = \sum_{i=1}^1 \sum_{j=1}^J y_{ij} \ln[P(y = j/x_i)] \quad (3.41)$$

Subject that  $P(y = j/x_i)$  is well specified,  $\hat{\beta}$  that maximizes  $\ln(L(\beta))$  is an efficient estimator of  $\beta$ . Under the assumptions that the underlying utility function is linear in its parameters the log likelihood function is globally concave (McFadden, 1973). McFadden also shows that  $\hat{\beta}$  is consistent and asymptotically normal and:

$$\sqrt{n}(\hat{\beta} - \beta) \rightarrow MN(0, n\Sigma^{-1}) \quad (3.42)$$

with  $\sigma$  is the covariance matrix of  $\beta$ .

And the coefficients  $\hat{\beta}_{kj}$  are obtained by the following equation:

$$\hat{\beta}_{kj} = \frac{\partial \ln[P(j/x)/P(J/x)]}{\partial x_k} = \frac{\partial [P(j/x)/P(J/x)]}{\partial x_k} \frac{1}{P(j/x)/P(J/x)} \quad (3.43)$$

The estimator of the variance of the estimator is given by:

$$\hat{V}(\hat{\beta}) = (E(-\frac{\partial^2 \ln L(\hat{\beta})}{\partial \beta \partial \beta'}))^{-1} \quad (3.44)$$

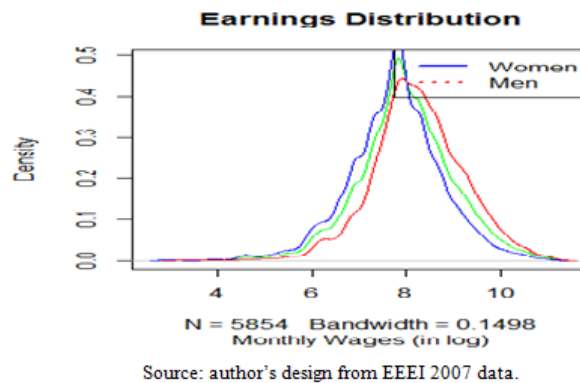
### 3.4.2 Data

Our empirical analysis is based on a national survey on employment and the informal economy in Haiti, conducted over the period October 2007- March 2008 by the national office of statistics (IHSI). The survey encapsulates two steps: a household employment survey and a survey regarding the informal production units identified by the employment survey. The survey adopt the ILO definition of the

informal employment elaborated during the International Conference of Labour Statisticians (ICLS) in 2003 but it excludes the cooperatives in this category because of their status in the Haitian legislation (IHSI, 2010). The unit of observation is a given individual in each year. We use a sample of 5,228 individuals from a national survey (IHSI, 2010) upon informal economy and employment, which contains 11,688 valid observations including 1,716 observations of agricultural workers and 4,130 involuntary.

The kernel density estimates of monthly revenues earned by workers (figure 3.1) shows small differences in densities between revenues earned by men and women. The overlap of the density functions of the revenues (figure 3.1) show that not all female workers earn less than male workers.

FIGURE 3.1: Figure 3.1: Densities of monthly log-wages for workers according to the gender



The empirical analysis is based on the data are pooled from the survey (IHSI, 2010). We use the following variables *Rev*: monthly revenues of each interviewed individual in local currency unit (gourde) ([0, 1265], (1265, 2485], (2485, 5000], (5000, 10000], (10000, 30000], and (30000, 65000]) *Sex*: gender of each person (man and woman) *Age*: Age of workers when the survey took place ([15, 25], (25, 45], and (45, 65]) *Geog*: geographic settings (rural and urban) *MS*: marital status divided in four modalities (single, married, free union and separated) *Edu*: educational level (illiterate, elementary school, middle and high school, and university) *WHr*: number of hours working per week ((3, 15], (15, 30], (30, 60], and (60, 122]) *Ind*: industry of activity (manufacture, trade, building and construction, services, and others) *SPC*: socio-professional category (trainee, salaried, self-employed, and entrepreneur) *Exp*: number of years of experience ([2,5], (5,10], (10, 20], and (20, 43]) *Hhs*: the size of the household of each person interviewed during the survey ([1, 3], (3, 6], (6, 10] and (10, 29)). *Wpl*: the place of work (premise, without premise and employers local).

The independent variables that we use regard demographic characteristics, human capital, employment and geographic setting. Some

of the variables are categorical (gender, marital status, schooling level, range of working hours and industry), some of them are binary (gender, geographic settings) and the continuous variables (age and years of professional experience) are transformed into categorical variable. The response variable is the revenues of the workers, which is originally a continuous variable. For the needs of the model we transform it into an ordinal variable (range of the revenues where the reference range is the range below the extreme poverty threshold).

TABLE 3.2: General characteristics of the informal workers

SEX	%	Industry	%	MARITAL STATUS	%
men	45.98	manufacture	13.51	single	25.03
women	54.02	building and construction	6.12	married	33.46
AGE	years	trade	49.33	free union	33.29
1 <sup>st</sup> quartile	28	services	29.74	separated	8.2
Median	36	others	1.3		
Average	37.2			Household size	
EDUCATION	%	Socio-professional category	%	Median	5.00
Illiterate	5.73	salaried	32.93	Maximum	29.0
Elementary school	6.51	trainee	6.88	Average	5.29
Middle and high school	16.71	Self-employed	57.58		
University	14.77	entrepreneur	2.6	Working Hours/week	
				Mean	48.11
REVENUES	gourdes	Market situation	%	Median	48.00
Median	2713	unemployed	15.31	1 <sup>st</sup> quartile	31.00
Mean	4606	Inactive	30.49	Maximum	122.00
1 <sup>st</sup> quartile	1500	employed	54.15		

Source: author's calculations using the data of the survey EEI 2007

As we aim to analyse the relationship between poverty and informality through human capital accumulation we analyse the extent of poverty among the group of individual workers. Poverty has many impacts on the workers (formal and informal). There are more poor among informal workers than formal workers. The poor represent almost the half of the informal workers (table 3.2), the proportion of the extremely poor among informal workers is nearly equal to the proportion of the poor. This underlines the impact of poverty on the

informal workers' lives. Among the informal workers, there are more poor women than non-poor, contrary to the men where the non-poor are the double of the poor. For rural informal workers the poverty is more remarkable because the proportion of poor is greater than the proportion of non-poor. There are more rural informal workers extremely poor than poor rural informal workers. The situation for underemployed and fully employed informal workers is almost similar when we compare poor and non-poor workers, but there are more extremely poor workers than poor workers for the underemployed.

TABLE 3.3: Poverty extent among the formal and informal workers gender

	Formal Workers			Informal workers		
	Extreme poor	poor	Non-poor	Extreme poor	poor	Non-poor
Men	9.06%	12.57%	78.36%	15.88%	20.87%	63.25%
Women	25.41%	26.07%	48.51%	27.35%	26.66%	45.99%
Urban	15.7%	18.59%	65.66 %	20.27%	23.33%	56.40%
Rural	29.17%	22.92%	47.92%	30.59%	27.20%	42.21%
Employed	20.00%	25.42%	54.58%	20.68%	24.51%	54.81%
Underemployed	11.16%	13.58%	72.25%	24.38%	23.73%	51.88%

Source: author's calculations using the data of the survey EEEI 2007

Globally formal workers have a better situation than informal workers regarding the poverty extent and nature. By observing the informal workers, we conclude that for almost all the groups considered except for the men, we find that the proportion of extremely poor is generally almost equal to the proportion of poor. Nevertheless, for women and rural informal workers the proportion of the extremely poor is greater than the proportion of the poor. For formal and informal workers, rural and female workers are more severely concerned with poverty than the other groups because the proportion of poor is more important than the proportion of the non-poor.

### 3.4.3 Model designing

We model our data by using an ordinal logit model which is a multinomial logit model. We study the relationship between the workers' revenues (an ordinal variable with 5 modalities) and K explanatory qualitative variables that we described earlier. We use the ordinal logistic regression with equal slopes, hence the main equation of the model can be written as follows:

$$\begin{aligned}
\log\left[\frac{P(Y \leq j / Pop.i)}{P(Y > j / Pop.i)}\right] &= \beta_{0j} + \begin{bmatrix} \beta_1 \\ -\beta_1 \end{bmatrix} \times \begin{bmatrix} men \\ women \end{bmatrix}' + \begin{bmatrix} \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ -\beta_2 - \beta_3 - \beta_4 - \beta_5 \end{bmatrix} \times \\
&\begin{bmatrix} ind_1 \\ ind_2 \\ ind_3 \\ ind_4 \\ ind_5 \end{bmatrix}' + \begin{bmatrix} \beta_6 \\ -\beta_6 \end{bmatrix} \times \begin{bmatrix} urb \\ rur \end{bmatrix}' + \begin{bmatrix} \beta_7 \\ \beta_8 \\ \beta_9 \\ -\beta_7 - \beta_8 - \beta_9 \end{bmatrix} \times \begin{bmatrix} ms_1 \\ ms_2 \\ ms_3 \\ ms_4 \end{bmatrix}' + \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ -\beta_{10} - \beta_{11} - \beta_{12} \end{bmatrix} \times \\
&\begin{bmatrix} spc_1 \\ spc_2 \\ spc_3 \\ spc_4 \end{bmatrix}' + \begin{bmatrix} \beta_{13} \\ -\beta_{13} \end{bmatrix} \times \begin{bmatrix} prem \\ noprem \end{bmatrix}' + \begin{bmatrix} \beta_{14} \\ \beta_{15} \\ \beta_{16} \\ -\beta_{14} - \beta_{15} - \beta_{16} \end{bmatrix} \times \begin{bmatrix} edu_1 \\ edu_2 \\ edu_3 \\ edu_4 \end{bmatrix}' + \\
&\begin{bmatrix} \beta_{17} \\ \beta_{18} \\ \beta_{19} \\ -\beta_{17} - \beta_{18} - \beta_{19} \end{bmatrix} \times \begin{bmatrix} exp_1 \\ exp_2 \\ exp_3 \\ exp_4 \end{bmatrix}' + \begin{bmatrix} \beta_{20} \\ \beta_{21} \\ -\beta_{20} - \beta_{21} \end{bmatrix} \times \begin{bmatrix} age_1 \\ age_2 \\ age_3 \end{bmatrix}' + \begin{bmatrix} \beta_{22} \\ \beta_{23} \\ \beta_{24} \\ -\beta_{22} - \beta_{23} - \beta_{24} \end{bmatrix} \times \\
&\begin{bmatrix} whr_1 \\ whr_2 \\ whr_3 \\ whr_4 \end{bmatrix}' + \begin{bmatrix} \beta_{25} \\ \beta_{26} \\ \beta_{27} \\ -\beta_{25} - \beta_{26} - \beta_{27} \end{bmatrix} \times \begin{bmatrix} hhs_1 \\ hhs_2 \\ hhs_3 \\ hhs_4 \end{bmatrix}'
\end{aligned}$$

This model is built in order to estimate ordinal response. We use the revenues of the workers as an ordered response, with categories extreme poverty threshold, poverty threshold and three different classes above the poverty threshold. All of the estimates of the model arise from focusing on the cumulative distribution of the response. In our model, we write the constant explicitly so we assume that the predictors do not contain a column of ones. Hence, if an explanatory variable increases by one then all transformed cumulative probabilities are augmented by the affected coefficient. That implies this model is more parsimonious than a usual multinomial logit or a hierarchical logit model.

### 3.5 Analysis issues

We analyse the informal employment market composition regarding the theory that stipulates informal workers are those who are excluded from the formal employment market. First, we study the main factors that influence the wages of each worker socio-professional category. Second, we analyse the results regarding the characteristics of poor and non-poor informal workers.

### 3.5.1 Determinants of revenues in the informal employment market

From the observed results (table 3.4) we have found that education, gender, age, experience, working hours, socio-professional category, geographic setting, industry and workplace are related to the revenue levels. Comparing other revenue levels to revenues, lower than the extreme poverty threshold (less than 1265 gourdes, or 1.23 U.S. dollars in PPP of 2005) only workers who earn revenues higher than extreme poverty and lower than the poverty threshold (2485 gourdes or 2.41 U.S. dollars in PPP of 2005) can gain wages without any human capital factor. Therefore, any worker without human capital and socio-demographic characteristics cannot gain revenues higher than the poverty threshold.

Global fitting model 3.5 records show that the model is globally well specified because the probability of the likelihood ratio test is lower than 0.00 which assumes that the ordinal multinomial logit model is well fitted to the data (beyond 5%). Then we can reject the null hypothesis that the model without predictors is as good as the model with predictors. The coefficients of the intercept levels (table c.1) show significantly that informal workers beyond the poverty threshold earn money regardless their geographic position, their marital status, their activity sector, their age, their schooling duration and their number of years of experience. In the table (3.6) for each qualitative variable it misses one level, that level represents the reference for comparison with other levels of the corresponding variable. Post-school returns grew faster than schooling returns, because for the wages of workers who have attended university are at least 4 times as high as the wages for those who did not. The odds of gaining higher wages are 5.39 times greater for informal workers attending university than for illiterate workers while the odds for workers attending middle and high school are 1.27 higher. The odds for the experienced workers (higher than 10 years) having higher revenues than less experienced workers are 1.41 times larger. Workers with less than 10 years of experience have odds 1.14 more chances to gain higher wages than workers with less than 5 years of experience. Individuals aged between 25 and 45 years old are positive and significantly related to gaining higher revenues in informal market than workers, which ages are in the range 15 and 24 years. The odds of having revenues in the higher ranges are disadvantageous of 0.61 for females than males. Socio-professional category impacts significantly the range of revenues of the workers. When people work as a trainee or apprentice they have odds 0.19 times higher than salaried workers to be in higher revenues ranges. Furthermore, entrepreneurs and independent workers have chances 3 times higher than salaried to gain higher revenues. But self-employed and salaried workers have approximately the same chances to be in identical revenues ranges. Working for 30-60 hours or

TABLE 3.4: Multinomial logit estimates of the extended Mincer wage model of informal workers

Explanatory Variables	Variables levels	Estimates		Explanatory Variables	Variables levels	Estimates	
		$\beta$	$exp(\beta)$			$\beta$	$exp(\beta)$
Intercept Ref:[0, 1265]	(1265, 2485]	1.0374*** (.1748)	2.8218	Industry Ref: Manufacturing	Building and Construct.	.7618*** (.1301)	2.1421
	(2485, 5000]	-.2218 (.1743)	.801		Trade	-.0416 (.0831)	.9592
	(5000, 10000]	- 1.7982*** (.1757)	.1659		Services	.1852* (.085)	1.2034
	(10000, 30000]	- 3.1143*** (.1805)	.0444		Others	.05 (.2295)	1.0512
	(30000, 65000]	- 5.6439*** (.2321)	.0035				
Age Ref:[15-25]	(25-45]	.3082** (.1053)	1.3609	Gender Ref: Male	Female	- .4996*** (.06)	.6067
	(45-65)	.2469 (.1847)	1.28				
Geog.setting Ref: urban	rural	- .2188** (.0688)	.8035	Workplace Ref:permise	Without premise	- .4351*** (.060)	.6472
Schooling Ref:illiterate	Elementary school	- .2543*** (.0724)	.7754	Marital status Ref:Free union	Married	.0872 (.0640)	1.0911
	Middle and high school	.2406*** (.0625)	1.272		Separated	.1146 (.095)	1.1214
	University	1.6858*** (.1496)	5.3967		Single	- .1424" (.0779)	.8672
Socio- professional Ref:salaried	trainee	- 1.6526*** (.1177)	.1915	Working hours/week Ref:[3, 15]	[15, 30]	.0957 (.1114)	1.1004
	Self- employed	-.0276 (.0779)	.9727		[30, 60]	.4903*** (.1)	1.6328
	entrepreneur	1.1056*** (.1646)	3.021		[60, 122]	.4654*** (.1069)	1.5926
Experience Ref:[2, 5]	(5,10]	.1344 (.1255)	1.1438	Household size Ref:(1, 3]	(3,6]	.0412 (.0621)	1.042
	(10,20]	.3518* (.1465)	1.4216		(6,10]	-.0016 (.0736)	.9984
	(20,43]	.3341" (.2008)	1.4107		(10,29]	-.1036 (.1452)	.9015

Source: author's calculations using the data of the survey EEI 2007

60-112 hours per week increase one's odds of approximately 1.6 of perceiving higher revenues than working for less than 15 hours per week that means the wages of informal workers grow significantly when



TABLE 3.5: Multinomial logit estimates of the extended Mincer wage model of informal workers - Global fitting model

Log likelihood full model	Log likelihood-intercept model	Chi-square	p-value
7561.129	8131.73	12.4043	.00

Source: author's calculations using the data of the survey EEEI 2007

they accumulate more working hours per week. Geographic setting is a significant predictor of the workers' ranges of revenues the odds of being in a higher range of revenues for workers living in a rural region are 0.8 higher than workers living in an urban agglomeration. Workers who live in larger households have fewer chances to gain higher wages than workers living in smaller households. Industry also proves to be a significant predictor of being in one of the revenues ranges. This predictor is more important for workers in building and construction and services with respective advantageous increase in odds of 2.14 and 1.20 compared with workers in the manufacturing sector. When informal workers undertake activities without premise they have fewer chances (0.64 times higher) to gain higher wages than those who have a premise for the activities. According to the results, the wages of informal workers are significantly determined by education, experience, age, socio-professional category, geographic setting, gender, industry, and number of hours working per week. Even if the marital status has one significant modality which is being single (it provides fewer chances to be in higher revenues ranges than living in free union) nevertheless the marital status and the size of households of the workers are not significant factors in the variation of the revenues. But their inclusion helps to increase the explanatory power of the model. For a better understanding of the roles of human capital and socio-demographic characteristics in the determination of informal workers' revenues, we must disaggregate informal workers' revenues for salaried and independent workers in order to compare them with revenues of similar categories of formal workers.

### 3.5.2 Human capital and socio-professional characteristics on the informal earnings

We compare the impacts of human capital on revenues for three categories of workers, formal salaried, informal salaried and independent informal. According to the results table 3.6 for the formal, informal salaried and independent workers only those who have revenues beyond the poverty threshold earn money regardless their education, gender, socio-professional category, experience, geographic position, marital status, sector of activities, age and workplace. Nevertheless, independent informal workers can earn revenues hardly

higher than the poverty threshold, regardless their human capital and socio-demographic characteristics. Women are less likely to assign to higher revenues regardless if they work in formal or informal sector.

TABLE 3.6: Ordinal multinomial logit estimates of the Mincer wages model by category of workers

Explanatory Variables	Variables levels	formal salaried		informal salaried		independent informal	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Gender</b> Ref: Men	<i>Women</i>	-0.3788* (0.1824)	0.6846	- 0.3396** (0.1166)	0.7121	- 0.6425*** (0.0757)	0.5259
<b>Marital status</b> Ref: Free union	<i>Married</i>	0.6945** (0.2211)	2.0027	0.2175 (0.1356)	1.2429	0.0405 (0.0748)	1.0414
	<i>Separated</i>	0.8524* (0.34690)	2.3453	0.14933 (0.2637)	1.1611	0.1225 (0.1042)	1.1304
	<i>Single</i>	0.5132* (0.2505)	1.6707	-0.0776 (0.1415)	0.9253	-0.1376 (0.1022)	0.8714
<b>Geog. setting</b> (Ref: urban)	<i>Rural</i>	- 0.9651** (0.33767)	0.3809	- 0.5536** (0.1825)	0.5748	- 0.2035** (0.0776)	0.8158
<b>Schooling</b> Ref: Illiterate	<i>Elementary school</i>	- 0.9957*** (0.2884)	0.3694	-0.3296 (0.2075)	0.7192	- 0.2625** (0.0797)	0.7691
	<i>Middle and high school</i>	1.3753*** (0.2479)	3.9563	0.7972*** (0.1307)	2.2194	0.0816 (0.0762)	1.085
	<i>University</i>	2.6378*** (0.316)	13.9836	2.6108*** (0.2167)	13.6103	1.1331*** (0.2568)	3.1054
<b>Working hours/week</b> Ref: [3, 15]	<i>[15, 30]</i>	-0.2835 (0.382)	0.7531	0.2547 (0.2317)	1.2901	0.0899 (0.1349)	1.0941
	<i>[30, 60]</i>	0.669 (0.3486)	1.9524	1.0782*** (0.2185)	2.9395	0.2641* (0.1187)	1.3022
	<i>[60, 122]</i>	0.4558 (0.3801)	1.5775	1.0421*** (0.2409)	2.8351	0.3021* (0.1255)	1.3527
<b>Industry</b> Ref: Manufacturing	<i>Building and Construction</i>	2.7354* (1.4585)	15.4161	0.7683*** (0.2153)	2.1560	0.9217*** (0.208)	2.5136
	<i>Trade</i>	—	—	0.0434 (0.1871)	1.0444	-0.0476 (0.1035)	0.9534
	<i>Services</i>	0.1391 (0.4944)	1.1492	0.2307 (0.1448)	1.2595	0.1268 (0.1155)	1.1352
	<i>Others</i>	0.2494 (0.682)	1.2833	0.1169 (0.2953)	1.1241	-0.1743 (0.4608)	0.84
<b>Experience</b> Ref: [2, 5]	<i>(5,10]</i>	-0.0344 (0.4138)	0.9662	0.3496 (0.2385)	1.4185	0.0078 (0.1713)	1.0079
	<i>(10,20]</i>	0.36427 (0.4814)	1.4395	0.5198* (0.2741)	1.6816	0.3107 (0.196)	1.3644
	<i>(20,43]</i>	0.37612 (0.6505)	1.4566	1.0464** (0.3956)	2.8475	0.1509 (0.2524)	1.1629

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%;

\*\*= 1%; \*= 5%; =10%.

Contrary to formal salaried and independent informal workers,

in comparison to informal salaried workers aged between 15 and 24 years, the workers aged between 25 and 45 years have more chances to earn higher revenues than those aged between 45 and 65 years. However, this observation is not significant for formal salaried and independent informal workers. While comparing the impacts of the marital status of formal salaried on their revenues with informal salaried and independent informal workers, we observe this predictor is significant and more important for female workers. When a formal salaried is married, or separated he has at least chances which are two times higher to earn more money than those who had never been married. Rural workers are significantly penalized in favour of urban workers. Nevertheless, this observation is more significant for formal workers whose have odds 0.38 times higher than formal urban workers. Globally, the three models are significant (table C.2). When the comparison occurs among illiterate workers and other groups, attending university is the biggest chance to earn higher wages for all of three categories of workers. However, this fact is way more important for formal and informal salaried than for independent informal workers because salaried workers (formal and informal) have 13 times more chances to earn higher wages than other workers. Furthermore, informal independent have odds 3 times higher than the odds of other workers for gaining higher revenues. Contrary to independent informal, for formal and informal salaried attending middle or high school is more benefit to earn higher revenues. Attending elementary school offers less opportunity of gaining higher revenues than attending elementary school in all the three categories. The number of working hours per week is not a significant predictor of the range of revenues for formal salaried but it is significant for informal salaried and independent to work more than 30 hours per week. Because working more than 30 hours increases the odds to earn higher wages than working less than that. Moreover, for informal workers this is equivalent to work for 30 to 60 hours per week and to work for 60 to 122 hours per week. Working in building and construction is significant for all the three categories (informal and formal salaried and independent informal). But working in building and construction in the formal sector makes the odds to gain higher wages 13 times greater than the odds in other sectors of activity. While working in the same sector in the informal market gives odds about two times greater than other sectors to gain higher wages. The workplace is significant only for independent informal and when the independent undertakes an activity without premise the odds of earning higher wages decrease. Experience and household size are significant for informal salaried, that means when salaried workers are experienced and live in small size households their odds to gain more are greater. But the odds of earning higher revenues increase with higher number years of experience for formal and informal salaried. In all the

three categories of workers the odds of gaining higher revenues decrease with a larger size of the household. Similarly, to formal salaried workers' revenue distribution of informal workers appears to obey to human capital and socio-demographic characteristics. For informal salaried age, gender, education, experience, working hours per week, household size, and industry are revealed to have significant impact on the determination of the revenues. Moreover, beside gender, workplace, and geographic setting only education level and working hours per week are important to determine the range of revenues of independent informal workers. Hence, we observe that independent informal workers are influenced by economic factors, whether than by socio-demographic factors. We deduce that the explanation of the informal workers' revenues is twofold. On one hand, we have a socio-demographic understanding related to informal salaried. Beside some key significant variables common to the three categories of workers, informal salaried revenues are influenced by age, household size, and marital status. In the other hand, there is a strictly economic understanding related to independent and informal entrepreneur. Conjointly with the key factors, the variables that impact the revenues of independent informal workers are workplace and industry, which appear to be rationally strategic variables of market positioning.

### 3.5.3 Market exclusion and Informal employment

Analysing informal employment as a result of the market exclusion obliges to answer to two key questions: is human capital impacted the revenues of informal poor workers? Do informal workers with important human capital accumulation earn most important revenues than workers with low human capital?

#### Poverty in the employment market

Testing the poverty approach of informal employment market returns to answer three fundamental questions: Are the poor informal workers' characteristics impacted their wages? Does human capital make differences between poor informal workers and poor formal workers?

According to the results (table 3.7) poor formal and informal workers do not earn revenues regardless their human capital and socio-demographic characteristics because the coefficients of the intercept are not significant. Middle and high school impacts hardly the revenues of poor workers. The odds of formal workers to earn the poverty threshold revenues are 2.5 times greater for those who attended middle or high school than illiterate workers (table 3.7). The odds of informal workers to earn poverty threshold revenues instead of extreme poverty revenues are 1.2 times greater for those who attended middle or high school than illiterate workers. Working for

TABLE 3.7: Ordinal multinomial logit estimates of the Mincer wages model of poor workers

Explanatory Variables	Variables levels	formal poor workers		informal poor workers	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Intercept</b>	(1265, 2485]	0.8591 (1.4029)	0.6846	-0.0132 (0.2907)	0.9868
<b>Gender Ref: Men</b>	<i>Women</i>	-0.4335 (0.4544)	0.6482	- 0.3445** (0.1068)	0.7086
<b>Age Ref: [15-25]</b>	(25-45]	-0.7842 (0.6424)	0.4564	0.2172 (0.1701)	1.2426
	(45-65]	-2.7840* (1.2782)	0.0617	-0.0706 (0.3221)	0.9318
<b>Geog. setting Ref: urban</b>	<i>Rural</i>	-0.4 (0.4939)	0.6703	-0.1014 (0.108)	0.9036
<b>Schooling Ref: Illiterate</b>	<i>Elementary school</i>	-0.4192 (0.3784)	0.6575	-0.1785 (0.1143)	0.8365
	<i>Middle and high school</i>	0.93* (0.4704)	2.5345	0.209* (0.11)	1.2324
	<i>University</i>	-0.4262 (1.5846)	0.6529	-0.1511 (0.4855)	0.8597
<b>Working hours/week Ref: [3, 15]</b>	[15, 30]	1.1414 (0.7539)	3.1311	0.545** (0.1766)	1.7246
	[30, 60]	1.4070* (0.6829)	4.0836	0.5882*** (0.1609)	1.8008
	[60, 122]	1.2798* (0.6988)	3.5959	0.5389** (0.1734)	1.7141
<b>Socio-professional category Ref: salaried</b>	<i>Trainee</i>	-2.962*** (0.791)	0.052	- 2.1071*** (0.1984)	0.1216
	<i>Self-employed</i>	16.606 (2014.505)	162900	-0.338* (0.149)	0.7131
	<i>Entrepreneur</i>	49.723 (0.014)	3920000	-0.4616 (0.4775)	0.6303
<b>Experience Ref: [2, 5]</b>	(5, 10]	0.6307 (0.5583)	1.8789	0.2735 (0.1927)	1.2146
	(10,20]	1.225 (0.788)	3.4042	0.2208 (0.2351)	1.2471
	(20,43]	2.7638* (1.2557)	15.86	0.3778 (0.3436)	1.4591

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; "=10%.

more than 15 hours per week is very significant in the determination of poor informal workers' revenues. When workers have cumulated more than 15 hours of work per week they have odds 1.7 times greater than other workers to earn poverty threshold revenues instead of

extreme poverty threshold revenues. Trainee/apprentices and self-employed workers are most likely assigned to extreme poverty revenues than salaried informal workers. These observations show that the earnings of poor informal workers are not impacted by human capital because education and experience are not significant predictors for revenues ranges. As gender and socio-professional category impact poor informal workers' revenues, they are not sufficient to describe their revenue distribution. Hence the earnings of informal workers lower than the poverty threshold are impacted by factors different of the human capital accumulation of the workers. Wages of formal workers are explained by the same factors that explained the revenues of informal workers. Only the experience ranges higher than 20 years make a small difference between the explanatory variables of formal and informal poor workers. Hence, we conclude that human capital accumulation does not make a difference between poor formal workers and poor informal workers. When we combine these two observations: poor informal workers do not earn revenues regardless their human capital accumulation and socio-demographic characteristics and human capital do not impact significantly the poor informal workers' gains. There is evidence that poor workers in the informal economy are influenced by external factors. That means workers enter informality because they are poor, but their poverty is not mostly determined by their socioeconomic characteristics.

### **Comparing informal and formal non-poor workers' characteristics**

A deeper view of the relationship between poverty and informal employment obliges to answer to the following question: Does informal employment consolidates disparities among workers in the sector? The formal market gives few opportunities to the younger workers with high level of schooling, 61% of unemployed people have attended the high school (Lamaute-Brisson, 2005), and more than 70% of informal workers attended at least the high school level, therefore we understand the proliferation of young and educated workers in the informal market.

Results recorded (table 4.8) show that non-poor formal and informal workers would not significantly earn revenues higher than 10 thousand gourdes without their human capital development and socio-demographic characteristics because the coefficients of the intercept mean the odds of being in higher range (>10,000 gourdes) are lower than .08. That means non-poor workers have less than 8% of chances to earn more than this amount regardless their human capital accumulation. Education level has high significance and huge impacts on the revenue levels of formal and informal non-poor workers. The odds of formal workers to earn revenues higher than 10 thousand of gourdes are 5.8 times greater for those who attended

TABLE 3.8: Ordinal multinomial logit estimates of the Mincer wages model of non-poor workers

Explanatory Variables	Variables levels	formal non-poor workers		informal non-poor workers	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Intercept:</b> Ref: [5000, 10000]	(10000, 30000]	-2.7555* (1.52)	0.0635	-2.57*** (0.4783)	0.0765
	(30000, 65000]	- 5.0175** (1.5403)	0.0066	- 5.1763*** (0.5014)	0.0056
<b>Gender</b> Ref: Men	<i>Women</i>	-0.1651 (0.2811)	0.8478	- 0.3488** (0.1294)	0.7055
<b>Age</b> Ref: [15-25]	(25-45]	16.1023 (895.961)	9843087.59	0.1788 (0.2772)	1.1957
	(45-65]	17.2115 (895.9615)	29845660.191	0.4355 (0.4291)	1.5457
<b>Geog. set- ting</b> Ref: ur- ban	<i>Rural</i>	-0.871 (0.869)	0.4185	-0.1636 (0.1765)	0.849
<b>Schooling</b> Ref: Illiterate	<i>Elementary school</i>	0.7474 (0.9843)	2.1115	-0.0669 (0.1794)	0.9353
	<i>Middle and high school</i>	0.6984 (0.4738)	2.0105	0.0445 (0.1366)	1.0455
	<i>University</i>	1.7621*** (0.4902)	5.8246	1.4041*** (0.2084)	4.0719
<b>Working hours/week</b> Ref: [3, 15]	[15, 30]	-1.0139 (0.6729)	0.3628	-0.2804 (0.2807)	0.7554
	[30, 60]	0.2976 (0.5871)	1.3466	0.2202 (0.2441)	1.246
	[60, 122]	0.2863 (0.6648)	1.3315	0.2021 (0.2574)	1.2239
<b>Socio- professional category</b> Ref: salaried	<i>Trainee</i>	-0.6277 (1.2057)	0.5338	-17.0357 (704.1875)	0.0
	<i>Self- employed</i>	0.1762 (1.2567)	1.1927	0.0857 (0.1548)	1.0895
	<i>Entrepreneur</i>	0.4019 (0.7112)	1.4946	1.1335*** (0.2166)	3.1067
<b>Experience</b> Ref: [2, 5]	(5, 10]	-15.4937 (895.9616)	0.00	0.4318 (0.4239)	1.5401
	(10,20]	-15.1064 (895.9617)	0.0	0.7641* (0.4514)	2.1471
	(20,43]	-16.0003 (895.9621)	0.0	0.4092 (0.5415)	1.5057

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; " =10%.

university than illiterate workers. When informal workers attend university their odds of earning higher revenues are 4.07 times greater than illiterate workers. Even if the number of working hours is not

significant, but the results (table 4.8) show when formal and informal workers work full time (more than 35 hours per week) their revenues grow up. When informal workers evolve in the sector of building and construction they have significant monetary advantages. Because their odds to earn higher revenues (greater than 10 thousand of gourdes) are 1.66 times larger than other sectors. Being an entrepreneur for informal workers is most likely and significantly assign to higher revenues than other socio-professional categories. Informal entrepreneurs have odds 3.1 times greater than others to earn revenues higher than 10,000 gourdes. In the informal employment market gaining in experience is significant to earn high revenues. When informal workers have cumulated more than 10 and less than 20 years of experience their odds to earn revenues higher than 10,000 gourdes are 2.2 times greater than other workers. Being married provides an advantage to non-poor workers (formal and informal) to earn more than the reference range revenues. But this situation is most profitable to formal workers whose odds of being in a higher range are 2.2 times larger than non-married formal workers. Despite its non-significant results for age (table 4.8) show that being young is a disadvantage in the informal market because the odds for older workers to earn higher revenues are greater. The impact of household size on the revenues levels is very significant in the informal market. The larger is the living household of the informal non-poor workers the smaller are his odds of earning revenues higher than 10,000 gourdes. More the household of informal workers regroups people less this worker gets a chance to earn higher revenues. These observations show that human capital accumulation (educational level, experience), socioeconomic characteristics (working hours, socio-professional category, working industry) and demographic profile (gender, marital status) impact significantly and considerably the earnings of non-poor informal workers. Revenues of formal non-poor workers are influenced by socio-demographic characteristics and educational level, but some other socioeconomic variables such as working hours, socio-professional category and household size also impacted the earnings but not significantly. Highly educated and experienced workers, who live in the small household, which are entrepreneurs who are working in building and construction sector earn the highest revenues. This situation shows evidence that the informal sector consolidates disparities among its workers. Therefore, we allow concluding that a portion of informal workers are not informal because they are excluded from the formal employment market. Two main observations deserve to be emphasized. Poor informal workers' revenues are not determined by their socioeconomic and demographic characteristics. Preferably their poverty is explained by the lack of human development, very limited educational level and their socio-professional category. Revenues of the highest remunerated informal workers are determined by the human capital accumulation and the socio-demographic profile.



These two observations lead to twofold conclusion. On one hand, an important portion of the informal workers is working in informality voluntarily because they are excluded from the formal employment market, due to their weak human capital accumulation. On the other hand, there are workers that evolve in informality voluntarily because they have characteristics to enter the formal market, but they refuse this option due to a rational view.

### 3.5.4 Conclusion

The main significant determinants of the revenues of the informal workers can be divided into positive and negative factors. The factors that influence positively the wages are experience, (more than 10 years), number of hours working per week (it provides higher wages when this number is larger than 30), age (between 25 and 45 years), educational level (university), socio-professional category (entrepreneur) and manufacture (building and construction and services). The negative determinants are gender (women), geographic setting (living in rural regions) educational level (elementary school), marital status (single), workplace (without premise) and socio-professional category (trainee). Common empirical assumptions related to the informal literature have linked informality to poverty in developing countries. One of the main ideas of these assumptions stated that informal workers are poor individuals excluded from the formal labour market. Among the factors pointed out previously which of them determine the revenues of the poor informal workers and which of them give indication related to the characteristics of non-poor informal workers? Socio-demographic profile has an important role in determining the revenues of informal salaried. Social status is more important for formal salaried workers than for informal salaried. Age and experience are significant factors in the determination of the revenues of informal salaried workers. Informal salaried workers who are living in rural regions are more penalised than formal salaried workers living in identical regions. The number of working hours is not significant in the determination of the revenues of formal salaried workers contrary to the informal salaried workers. Moreover, the revenues of independent and informal entrepreneurs are influenced by human capital and socioeconomic variables. There is a very large gap between the revenues of the informal entrepreneurs and the revenues of salaried and independent informal workers. If education is a more important factor formal salaried workers for gaining higher wages than for informal salaried. Overall, human capital accumulation is more important for the independent informal workers. Because for informal salaried besides the educational level, age and experience are also significant factors in determining the gains contrary to the formal salaried. This is a mix evidence. First informal employment is not mostly constituted of marginalized individuals who work in

exploitative condition to ensure their survival contrary to the view of some theorists (Chen et al., 2004; Williams and Round, 2008). Informal workers who earn lower wages, present the lowest level of human capital accumulation and higher revenues are earned by informal entrepreneurs with higher levels of human development. This implies two important considerations. First, a portion of informal workers does not have characteristics to evolve in the formal employment market that explains they are excluded from it. This category contains workers that are struggling in the informal market because of their lack of human development and they constitute the poor of the market. These informal workers with poor capacity for capital accumulation and very low rates of remuneration correspond to the view of Harris and Todaro (1970). Second, informal workers are not globally marginalized individuals excluded from the formal market. Based on the observation that the higher revenues are earned by workers with more interesting socioeconomic profile and human capital accumulation, we highlight what follows. Contrary to some theories (Harris and Todaro, 1970), many people employed in informal activities are mainly characterized by high capacity for human capital accumulation and they earn important revenues. Accordingly, to the view of some researchers (Pahl, 1984; Williams and Windebank, 2003), informal employment rather than reduce, consolidates the disparities (produced by formal capitalist employment market) among workers, that means informal workers with higher human capital earn larger revenues than others, hence this portion of informal workers are not informal because they are excluded from the formal employment market. These two points constitute another mix evidence for the voluntary aspect of informal employment. The first point, contrary to the view of Perry et al., (2007) shows that some workers (with low capital accumulation) evolve in informality because they are excluded from the formal employment market. The second point implies that some people who evolve in the informal sector have chosen voluntarily to be in informality, because they have sufficient characteristics to move to the formal sector but refuse (Perry et al., 2007). Which is the case with the workers with higher human capital accumulation and higher revenues. The earnings of informal workers lower than the poverty threshold are impacted by socioeconomic factors and the human capital accumulation of the workers. That means the poverty situation of the workers is explained by their socioeconomic profile and their lack of human development, hence workers enter informality because they are poor. The gains of non-poor informal workers are significantly and mainly influenced by human capital accumulation, socioeconomic characteristics and demographic profile. High level of schooling (for the model of the non-poor workers) and age (for the model of the informal salaried workers) are significant factors in determining informal revenues. While being youth and highly educated are two main characteristics of the individuals that

formal market has excluded (Lamaute-Brisson, 2005). These observations constitute the mixed evidence in support of the theory that stipulates informal workers are individuals that formal market has excluded. The estimates (all the models) show that earnings increase significantly with high school level, full time work, building and construction sector and urban positioning. There is a large discrepancy between the revenues of the workers who have attended university and the earnings of those who have attended middle or high school. Basically, earnings differentials among non-poor informal workers increase with educational level, socio-professional category, gender, marital status, experience and household size. The increase is much more pronounced with education. Returns to education and to experience are lower for the less educated informal workers. Nevertheless, some socio-demographic factors that have not influenced the revenues of the poor workers, influenced negatively the revenues of the non-poor informal workers, this is the case with the household size. Inversely an economic factor such as the number of hours working per week does not influence the revenues of non-poor workers but influence significantly the revenues of the poor workers. Due to all these facts, we can conclude that a portion of informal workers do not take informality as exit but they are excluded from the formal labour market. Another portion is likely assigned to informality for other reasons than exclusion because they have all the characteristics to enter the formal employment market. Some empirical studies consider informal employment market as a multi-segmented market with its own dynamic of development (Fields, 1990; Chen, 2005). According to that approach informal sector has linkages with formal economy, there are variations on labour market segmentation, and it exists many possibilities of labour mobility from formal to informal market and vice versa. Mix evidence for poverty theories and rationality approaches of informal employment, absence of evidence for economic discrepancies among informal workers open the door to that theory. Regarding the significance of schooling level, experience, number of working hours and industry in the wage distribution, two assumptions are workable: informal market has its own development dynamics and informal employment market has multiple segments therefore it exists entry barriers for this market. These assumptions can constitute a research track to deepen in analysing an eventual multiple segmentation of the informal employment market.

## Chapter 4

# Investigating labour market segmentation with Dirichlet Process model: the informal labour market in Haiti

### 4.1 Introduction

The most common a priori assumption in the empirical classic literature advocating the informal sector is homogeneous has been challenged. Now policy makers and analysts debate on whether informal labour is an attractive opportunity or a struggling sector of last resort? This chapter applies a Bayesian semi-parametric model (Lo, 1984) of the labour market of Haiti to test the empirical relevance of this hypothesis. We model the distribution of individuals within each segment as a continuous process via a separate Dirichlet process mixture model for the population density of each segment. We use a Bayesian semi-parametric inference to analyse the market structure dictated by the Dirichlet process prior and to analyse the different unobserved segments of the informal market.

The expansion of informal employment in developing countries, the high wages of informal workers with high human capital and the gap of revenues between workers with equal work have triggered two main questions. Is the informal employment a voluntary choice of the individuals as an employment opportunity or are workers dropped into informal activities due to entry barriers into the formal sector? Answers to the above questions interest both researchers and policy makers because they will bring light over the following most important question issue: whether the income gap between formal and informal sectors results from market segmentation or whether theories regarding integrated labour market are still appropriate despite the observed gap between wages. Public policies which aim to reduce informality or want to create less volatile jobs are concerned with these answers.

The paradigm "lower wages, lower returns to education and experience" which is remarkable in the Haitian informal market (chapter 3) does not necessarily imply market segmentation (Heckman and

Hotz, 1986; Pratap and Quintin, 2006). Segmented market means that workers are not free to move from one side to another one (Dickens and Lang, 1985; Basu, 1997) this is the main question for us. According to Gunther and Launov (2011), analysing the labour market segmentation leads to two questions: are individuals poor because they evolve in the informal sector? This implies that informal market is a segmented labour market. Or are they evolving in informality because they are poor? This means that the informal market is an integrated labour market.

The different approaches of informal market can be summarised by two main ideas. First, informality is a survival sector to escape involuntary unemployment (Harris and Todaro, 1970; Lucas, 1978; Pahl, 1984; Williams and Windebank, 2003; Chen et al., 2004; Perry and al. 2007; Williams and Round, 2008). Second, informality is a deliberate choice of workers based on revenues or utility maximisation (Rosen, 1981; Heckman and Sedlack, 1985; Carneiro, Heckman, and Vytlačil, 2005; Gunther and Launov, 2011). Many of the classic studies (Gindling, 1991; Pratap and Quintin, 2006) upon the developing countries provide evidence of a competitive market, which means that informal employment is a deliberate choice based on comparative advantage considerations. Several authors have disputed the dominant *a priori* assumption of the empirical literature regarding the homogeneity of the informal sector (Fields, 2005; Perry and al., 2007; Gunther and Launov, 2011) and design a different structure of the informal labour market. Although some of these researches underline the heterogeneous character of informal employment (Fields, 2005; Perry and al., 2007), their main concern is whether the informal work is voluntary or not. Few studies have focused on a theory of the existence of the heterogeneity of the informal sector. One of the references, studies on this domain (Cunningham and Maloney, 2001) has some important limitations regarding the labour market configuration. The most complete study regarding the segmentation of the informal labour market (Gunther and Launov, 2011) analyses the unobserved heterogeneity of the informal market and proposed a method that can be classified as finite mixture regression with sample selection. But the model of Gunther and Launov cannot make test for the additional segment (Gunther and Launov, 2011).

By questioning the twofold view of the relationship between formal and informal sides of the labour market, this chapter argues that these two sides of the global labour market can have a different relationship than being integrated or segmented. We argue that the labour market for relatively skilled workers may be well integrated with both formal and informal sides. This integrated market offers desirable jobs with various characteristics from which workers can make their choice. But the same labour market for relatively unskilled workers (inexperienced and low educated) place entry barriers in the formal side and oblige them to subsist in the informal sector. In

the study of the informal labour market segmentation, the main idea of the clustering analysis is to uncover the group structure of the observations and identify variables that best distinguish the different groups.

Determining the number of clusters (segments) is a long-standing issue in all the clustering methods, including mixture models (McLachlan & Basford, 1988; Banfield & Raftery, 1993). To deal with this problem, researchers generally use finite mixture models with an unknown number of components, such as the reversible jump Markov Chain Monte Carlo (MCMC) method (Richardson & Green, 1997; Tadesse et al., 2005) and continuous time Markov birth-death processes (Stephens, 2000), which allow to create and delete components. These methods are based on the assumption that the Gaussian process used is stationary. Such model is a finite-dimensional model because the distribution of the Gaussian process is determined by the parameters in the mean and the covariance structures. Therefore, an alternative method is necessary to fit these models. Researchers propose models that define mixture distributions with a countable infinite number of components. One of the best ways to fit these models is the use of a Dirichlet process prior for the mixing proportions (Antoniak, 1974; Ferguson, 1983). Various MCMC sampling methods for fitting Dirichlet process mixture models have been developed (Escobar, 1994; MacEachern, 1994; Escobar & West, 1995; MacEachern & Muller, 1998).

Hence, in this framework we propose a Bayesian clustering model because the variable distributions are unknown and we based it on an infinite multinomial sampling with a Dirichlet prior (Sethuraman, 1994). A Dirichlet process is an infinite-dimensional generalization of finite-dimensional Dirichlet distributions; hence, it allows to dispense with the disturbing hypothesis of stationarity of the Gaussian process.

Many aspects of our paper are new to the empirical literature on informality. Most papers regard the segmentation of the informal labour market as it depends only on income maximization. Our paper chapter analyses the informal labour market as it can be explained by revenue rationality, utility maximization, and human capital accumulation. Second, our paper chapter is the first to propose the introduction of the infinite mixture model, semi-parametric Bayesian, Dirichlet process, Markov Chain Monte Carlo (MCMC) and Blocked Gibbs sampling in the empirical literature of informality. Third, the model can be applied to time series data on the labour market.

We use the Dirichlet process twice. To determine the number of segments of the labour market, we first model the distribution of individuals within each segment as a continuous process via a separate Dirichlet process mixture model for the population density of each segment. Second, we use the expectation-maximization algorithm to estimate the global Dirichlet process model and evaluate the characteristics of each side of the labour market.

The model is applied to a national survey data from Haiti (IHSI, 2010) upon informal economy and employment between 2007 and 2008, which contains 30,337 observations, wherein a sample of 14224 individuals is pooled. As in Gunther and Launov (2011), our proposed model is applicable to any cross-sectional household survey that collects data on employment, wages and individual characteristics. Second, the model can serve to segment other population data based on social or financial similarities. Third, the model permits to test explicitly the voluntary aspect of informal work.

The structure of the chapter is organized as follows. Section 2 presents a literature review of the types of informal labour market (segmented and integrated). Section 3 develops the Dirichlet prior process model and justifies its foundation. Section 4 details the two estimation processes used to estimate the double Dirichlet process. Section 5 presents the data and the empirical results and analyses. Section 6 concludes.

## 4.2 Literature Review on labour market segmentation and mobility in Haiti

The economic literature identifies three main reasons that dissuade individuals enter the formal labour market. First, in order to maximize their revenues, the workers escape the government direct taxes and indirect taxes from labour protection laws. The inefficiencies and rigidities that usually accompany such protection measures, persuade workers to evade these direct and indirect taxes by evolving in the informal sector. Second, informal workers can maximize their utility because of the various aspects of informal work they can find some interesting advantages such as greater flexibility or availability. Third the incapacity of workers with low human capital accumulation to bypass the entry barriers in the formal sector. In order to withdraw themselves in the involuntary long-term unemployment these workers operate in the informal sector. The first two reasons (wages rationality and utility maximization) that can deter workers to entry the formal market lead the hypothesis of a competitive labour market that integrates formal and informal workers. The third factor of deterrence of entering the formal market (entry barriers in the formal market due to low capital accumulation of the employment seekers) constitutes the base of the segmentation hypothesis of the labour market.

Hence the empirical literature on informal employment market identifies two opposing theories. The segmented labour market theory stipulates that workers with low human capital accumulation try to escape long-term and involuntary unemployment imposed by the formal market where they encounter entry barriers. The competitiveness hypothesis sees the informal market as a place where individual

workers choose deliberately to evolve for two main reasons: revenues rationality and utility maximization. In this section we make a literature review of each type of informal labour market and we present the Haiti's labour market factual situations.

### 4.2.1 The segmented Labour market

The dual labour market is based on the premise that the wage level and working condition are not available to the employment seekers until they decide to bypass the barrier of entry (Mazumdar, 1983). The segmentation theory considers informality as a survival sector to escape involuntary unemployment (Harris and Todaro, 1970; Lucas, 1978; Pahl, 1984; Williams and Windebank, 2003; Chen et al., 2004; Perry and al., 2007; Williams and Round, 2008). One of the main ideas of the segmentation approach of the informal economy stated that informal work is a transitory state (Harris and Todaro, 1970) that the economic growth will reduce over the long term. This theory views, informal economy as a negative integral part of the official economy. This theory is based on many approaches. Different theorists define the informal economy as a leftover of pre-capitalism constituted of marginalized individuals who work in exploitative condition to ensure their survival (Lewis, 1954; Chen et al., 2004; Williams and Round, 2008). For others, informal employment consolidates, rather than reduces, the disparities produced by formal capitalist employment (Pahl, 1984; Williams and Windebank, 2003). Some theorists of the segmented labour market consider informality as a new form of activity emerging in late capitalism as a creative response to the state's incapacity to satisfy the basic needs of the impoverished masses created by the deregulated open world economy (de Soto, 1989; Hudson, 2004). The other theory states that employees in informality often work under degrading conditions, are low-paid and are working in sweatshop like conditions (Sassen, 1997). Nevertheless, analysing segmented labour markets does not imply a wage differential analysis because wage differences do not imply necessary market segmentation. All the theories related to the dual labour market (formal and informal) are based on the wage differential between the two sectors (Fields, 2005). Some papers on dual labour market are based on human capital theory to establish market segmentation (Schultz, 1961, 1962; Becker, 1962; Mincer, 1962, 1974). For these theorists, in a dual labour market different wages must be allocated in different sectors to comparable workers. Globally for the segmented labour market view, informal workers with identical characteristics than those in the formal market should earn higher wages in the formal sector. Nevertheless, we find few evidence in the empirical literature for such view.

The hypothesis of the segmented labour market is based mainly on the existence of wage differentials between the sectors (Rosenzweig,



1988; Maloney, 1999). Two arguments rule out this dualistic view (Magnac, 1991). These tests are inappropriate if their fundamental base is the comparison of the wage functions between formal and informal sectors with OLS techniques (Cain, 1976). The hypothesis that the choice of the operating sector of individual workers is determined by the distinct offered wages leads to a selectivity bias (Gronau, 1973). The hypothesis of the linear distribution of worker characteristics implies that wages are paid in different sectors following the characteristic distribution. The implicit prices of the productive characteristics vary systematically across sectors for one or both of the following reasons. First, in each sector, each agent presents particular skills and characteristics (Rosen, 1978), and second, individual skills are grouped by sector (Heckman and Scheinkman, 1987).

While it is a usual view in the development literature that differences in individual and job characteristics alone cannot account for earnings differentials between sectors, the evidence on this question is mixed. For most estimates of earning functions for developing nations find evidence for the existence of significant differences across sectors (banerjee, 1983; Rosenzweig, 1988; Pradhan and Van Soest, 1995), but these studies are based on strong parametric hypotheses that may lead to inappropriate results. Moreover, researchers underline that the results are sensitive to the wage equation chosen to specify (Heckman and Hotz, 1986; Maloney, 1999) and OLS estimates are biased and inconsistent because individuals may select sectors on the basis of observed and unobserved characteristics that also affect earnings. Pratap and Quintin (2004) fulfil the gap of using semi-parametric techniques to sepecify a wage function for developping countries. They find no evidence for the assumption that observably identical workers earn higher wages in the formal sector than in the informal sector. Due to the important limit of the segmented labour markets assumption, we analyse further the hypothesis that labour markets(formal and informal)are competitive.

#### **4.2.2 The integrated informal market**

The integrated labour market model, that some theorists called the unified labour market model, has two particular features: each labour market (formal and informal) clears, and wage equalization leads to full intermarket equilibrium. To make the analysis easier this theory supposes that the economy has two sectors, the modern and traditional sectors. All workers are supposed to be identical and they are wage maximizers, and so they choose to work wherever the wage is higher, be it in modern or in traditional sector. The main assumption of the integrated labour market theory stipulates that for a given type of worker both sectors of the economy provides the same wage. Some theorists consider informal work as the result of the evasion of the protection fees by the large modern enterprises through

subcontracting to workers without paying the protection fees (Portes, Castells and Benton, 1989). The protected workers in the formal sector are constituted of protected workers who enjoyed high wages, vacation and pension and job security. Contrary to the informal workers who are evolving in labour-incentive activities without social security and non-wage benefits (De Almeida et al., 1995). According to this point segmented market is the effect of the decisions of the public authorities and the distortions imposed by the unions interventions (Maloney, 1999). Then individuals evade the formal market to escape to state interventions. The theory of a competitive informal market views informality as a deliberate choice of workers based on revenues or utility maximisation (Rosen, 1981; Heckman and Sedlack, 1985; Magnac, 1991; Carneiro, Heckman, and Vytlačil, 2005; Gunther and Launov, 2011). Many of the classic studies (Gindling, 1991; Pratap and Quintin, 2006) upon the developing countries provide evidence of a competitive market, which means that informal employment is a deliberate choice based on comparative advantage considerations. In a fully integrated market, almost all the workers can take benefit from comparative advantages. Therefore, we should not use observable wage differential nor potential wage differential as evidence for the hypothesis of segmented market.

### 4.3 The Competitive and Segmented Models

Many models have been elaborated in order to analyse the relationship between formal labour market and informal labour market. Some of these models consider formal and informal labour market as complements, some others advocate that formal and informal labour markets are competitive. This section is devoted to the development of the two categories of models.

#### 4.3.1 The Segmented Model

Various models have been developed for analysing segmented labour markets, but all of them are based on the seminal Roy model (Heckman and Sedlacek, 1985). The Roy model denotes  $f(s|\theta)$  the density function of workers' skills, where  $\theta$  is the vector of parameters. Then, the proportion of the workers evolving in sector  $g$  is the proportion of the workers whose skills endowment lies in  $\Delta_g$  is given by:

$$p(g|s) = \int_{\Delta_g} f(s|\theta) ds$$

The aggregate supply of task to segment  $g$  is given by the integration of the micro supply over  $\Delta_g$ :

$$T_g = \int_{\Delta_g} t_g(s) f(s|\Theta), g = 1, 2$$

This model assumes that the density of skills  $f(s|\Theta)$  and the task functions  $t_g(s)$  are such that  $(t_1, t_2)$  are log-normally distributed with mean  $(\mu_1, \mu_2)$ . Denoting  $(u_1, u_2)$  a mean zero normal vector, and  $\pi_g$  the equilibrium price of task  $g$  in sector  $g$ , the model assumes that agents can choose between two possible sectors:

$$\ln(w_1) = \ln(\pi_1) + \mu_1 + u_1$$

and

$$\ln(w_2) = \ln(\pi_2) + \mu_2 + u_2$$

The choice of the workers is based on the comparison of the wages between the sectors. Economic agents with endowment  $s$  enter the sector where higher wages are paid. Assuming that prices are positive the set  $\Delta_g$  is defined as

$$\Delta_g = \{s : \pi_g t_g(s) \geq \pi_j t_j(s), g \neq j\}$$

Denoting  $\sigma^* = \sqrt{\text{var}(u_1 - u_2)}$  and  $c_g = [\ln(\pi_g/\pi_j) + \mu_g - \mu_j]/\sigma^*$ ,  $g \neq j$ , the Roy model finds

$$p(g|s) = P(\ln(w_g) > \ln(w_j)) = \Phi(c_g), g \neq j, g, j = 1, 2$$

In this equation  $\Phi()$  denotes the cumulative distribution function of a standard normal variable.

An extended version of the Roy model is proposed by Heckman and Sedlacek (1985) for segmented market. This model assumes that workers are not only revenue maximizers but they are also utility maximizers when they choose their sectoral activities. The model develops a general non-normal model for unmeasured tasks  $(\mu_1, \mu_2)$  and it includes a non-market production segment as an alternative to the usual labour markets (formal and informal). Denoting  $U_g$  the utility of evolving in sector  $g$  where the third segment corresponds to the non-market sector, a worker chooses to enter in sector  $g$  if the utility of this sector is greater than the utility of the other sectors

$$U_g > U_j, g \neq j, g, j = 1, 2, 3$$

Let  $X$  denote a vector that contains  $Z_g$  the vector of the measured consumption specificities of the sector and the household characteristics,  $Y_g$  the vector of skill characteristics,  $\pi_g$  the log task prices. Heckman and Sedlacek propose the following equation

$$\ln U_g = \gamma_g X + \nu_g, g = 1, 2, 3$$

It is assumed that  $X$  is distributed independently of all the  $\nu_g$  and  $(\nu_1, \nu_2, \nu_3) \sim N(0, \Sigma_\nu)$ . This specification leads to the Thurstone

multivariate probit model (Heckman and Sedlacek, 1985).

### 4.3.2 The competitive Model

The competitive model assumes that formal and informal labour market are fully integrated. Magnac (1991) develops another extension of the Roy model, unlike Heckman and Sedlacek (1985), he proposes two versions: the competitive market and the segmented market. In the competitive model Magnac (1991) assumes that workers can choose among four choices, working in the formal sector or in the informal sector, being unemployed and do not participate in the labour market. Magnac (1991) defines the potential wages as follows

$$\text{Segment 1 } \ln w_{1i} = X_i^0 \delta_1 + t_{1i}$$

$$\text{Segment 2 } \ln w_{2i} = X_i^0 \delta_2 + t_{2i}$$

$$\text{Shadow wage } \ln w_i^* = (X_i^0 Z_i) (\delta_0^* \delta_1^*)' + t_i^*$$

Where  $X^0$  is a vector of  $s$  independent variables and  $(\delta_1, \delta_2, \delta_0^*) \in (\mathbb{R}^s)^3$ ,  $Z$  is a vector of  $p - s$  independent variables that represents the home productivity. The model assumes that, conditional on  $X^0$  and  $Z$ ,  $(t_{1i}, t_{2i}, t_{3i}) \sim N(0, \Sigma_i)$ .

The competitive model is represented by the following equations:

$$\text{Segment 1 } \ln w = X \delta_1 + t_1 + \varepsilon_1 \text{ if } X \alpha_0 + \nu > 0, X(\beta_0 - \alpha_0) + (u - \nu) < 0$$

$$\text{Segment 2 } \ln w = X \delta_2 + t_2 + \varepsilon_2 \text{ if } X \beta_0 + u > 0, X(\beta_0 - \alpha_0) + (u - \nu) > 0$$

$$\text{Unemployed if } X \beta_0 + u > 0, \text{ or } X \alpha_0 + \nu > 0$$

$$\text{No participation if } X \beta_0 + u < 0, \text{ or } X \alpha_0 + \nu < 0.$$

Denoting  $X = (X^0 Z)$  and  $\varepsilon_1, \varepsilon_2$  the measurement errors of the observed wages. Magnac (1991) defines:

$$\alpha_0 = \begin{bmatrix} \delta_1 - \delta_0^* \\ -\delta_1^* \end{bmatrix}, \alpha_0 = \begin{bmatrix} \delta_2 - \delta_0^* \\ -\delta_1^* \end{bmatrix}$$

He also defines two sets of constraints:

$$\beta_{0j} - \alpha_{0j} = \delta_{2j} - \delta_{1j}, j = 1, s$$

$$\beta_{0j} = \alpha_{0j}, j = s + 1, p$$

The first set of constraint implies that in competitive markets, the differences in the offered wages in each sector are similar to the difference between the conditions of entry in the segments. The second set of constraints means that only human capital accumulation influences the choice of sector. Magnac (1991) concludes that the well-known comparison of wage functions across segments is not a test of the competitive character of the labour markets.

## 4.4 Bayesian Model: Dirichlet Process

In this chapter, we show that the informal labour market can be segmented on the one hand and competitive on the other hand; therefore, informal workers can evolve in different possible latent segments of this labour market. Despite many important details that household surveys can bring in, data only provide information regarding whether the workers belong to the formal or informal sector and it is not possible to observe and separate specific segments of the informal work. Like Gunther and Launov (2011), we assume that each segment of the informal is determined by a unique wages equation. To avoid making various *ad-hoc* assumptions contrary to Gunther and Launov (2011), we identify sub-segments of the labour market using observable variables (wages, geographic settings, industries, human capital, etc.) and unobservable prior knowledge (determination of a prior probability function for each individual with a set of particular characteristics and skills). We do not want to make guesses about the number of segments that can cause serious biases.

### 4.4.1 The Model

Suppose that a respondent's answer to a question with categorical answer and he chooses  $\theta^k$ , we denote that by  $\Theta = \theta^k$ . Let assume that  $N$  individuals are observed answering this question  $N_k$  times  $\theta^k$ . A logical estimate of the probability of this answer is then that  $\hat{P}(\Theta = \theta^k) = N_k/N$ . A multinomial sampling is assumed, thus the observed variable  $\Theta$  is discrete, having  $r$  possible states  $\theta^1, \dots, \theta^r$ . Hence the likelihood function is given by:

$$P(\Theta = \theta^k | g) = g_k, k = 1, \dots, r \quad (4.1)$$

Where  $g = \{g_2, \dots, g_r\}$  are the parameters (probabilities) and  $g_1 = 1 - \sum_{k=2}^r g_k, g_k \leq 0, \forall k$ . Let  $D = \{\Theta_1 = \theta_1, \dots, \Theta_N = \theta_N\}$  a dataset which admits as sufficient statistics  $\{N_1, \dots, N_r\}$  where  $N_k$  determines the number of times that  $\Theta = \theta^k$  in the dataset  $D$ . The complete dataset has for likelihood:

$$P(D|g) = \text{Multinomial}(\cdot | g) = \frac{1}{C} \prod_{k=1}^r g_k^{N_k} \quad (4.2)$$

Where  $C$  denotes normalization constants. Hence the maximum likelihood estimate is:

$$g_k^{ML} = N_k/N \quad (4.3)$$

We obtain the very intuitive result that the parameter estimates are equal to the empirical counts. When it occurs that some or many counts have small values (when  $N < r$ ) then their probabilities might

be (incorrectly) estimated to be zero. Thus a Bayesian approach could be fitted.

#### Dirichlet Prior

In a Bayesian approach, an a priori distribution can be defined for  $g$ . In this case, a conjugate prior is the reasonable choice, then a Dirichlet distribution:

$$P(D|g) = \text{Dir}(\cdot|\alpha_1^*, \dots, \alpha_r^*) \equiv \frac{1}{C} \prod_{k=1}^r g_k^{\alpha_k^*-1} \quad (4.4)$$

Where  $\alpha^* = \alpha_1^*, \dots, \alpha_r^*, \alpha_k^* > 0$  now we should re-parameterize some of the variables as follows:

$\alpha_0 = \sum_{k=1}^r \alpha_k^*, \alpha_k = \alpha_1^*/\alpha_0, k = 1, \dots, r$  and  $\alpha = \{\alpha_1, \dots, \alpha_r\}$  such that:

$$\text{Dir}(\cdot|\alpha_1^*, \dots, \alpha_r^*) \equiv \frac{1}{C} \prod_{k=1}^r g_k^{\alpha_0 \alpha_k - 1} \quad (4.5)$$

But the understanding of parameters becomes evident when we note that:

$$P(\Theta = \theta^k | \alpha^*) = \int P(\Theta = \theta^k | g) P(g | \alpha^*) dg = \int g_k \text{Dir}(g | \alpha^*) dg = \alpha_k$$

The posterior distribution is another Dirichlet:

$$P(g|D, \alpha^*) = \text{Dir}(\cdot|\alpha_1^* + N_1, \dots, \alpha_r^* + N_r) \quad (4.6)$$

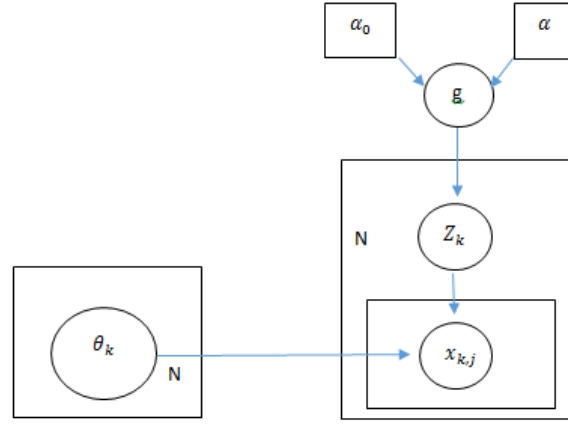
Since a Dirichlet distribution is conjugate to multinomial sampling, the posterior distribution of  $\theta$  is also given by a Dirichlet with the difference of the updated parameters  $\alpha_0 \alpha_k + N_k$ . The probability for the next data points is given by:

$$P(\Theta_{N+1}) = (\theta^k | D, \alpha^*) = \int g_k \text{Dir}(g | \alpha_1^* + N_1, \dots, \alpha_r^* + N_r) dg = \frac{\alpha_0 \alpha_k + N_k}{\alpha_0 + N} \quad (4.7)$$

We observe that the increase of  $N_k$  gives the identical result as with the maximum likelihood approach and the prior becomes negligible. Let assume that it is not possible to observe the values of the variables directly, but only some derived quantities (noisy measurements)  $X$  with some probability  $P(X|\Theta)$ . Let  $D_k = \{x_{k,j}\}_{j=1}^{M_k}$  be the observed measurements of the  $i$ -th answer and let  $P(x_{k,j}|\theta_k)$  be the probability distribution of the variable values. We can pay attention to two factors. We can be interested in calculating the probability of the actual possible values of the variable  $P(\Theta_1, \dots, \Theta_N | D)$  or in interfering the property of the different values of the variable by determining  $P(g|D)$ . Now, due to the fact that the  $\Theta$  are missing we are dealing with a problem with missing data. Since a such problem is relevant also in Dirichlet process.

To bypass the problem of missing data the introduction of an auxiliary variable is necessary. Then we introduce a vector of auxiliary variables with states  $Z^1, \dots, Z^r$  we obtain:

FIGURE 4.1: Multinomial sampling with a Dirichlet prior, noisy observations and the auxiliary variable



Source: author's design

$$P(Z = z_k | g) = g_k, k = 1, \dots, r \quad (4.8)$$

$$P(\Theta = \theta^k | Z = z^k) = \delta_{j,k}, k = 1, \dots, r \quad (4.9)$$

The following predictive distribution contains all we have learned about the model based on the data:

$$P(\Theta_{N+1} | D) = \sum_{\theta_1, \dots, \theta_N} P(\theta_1, \dots, \theta_N) P(\theta_{N+1} | \theta_1, \dots, \theta_N) \approx \frac{1}{S} \sum_{s=1}^S P(\theta_{N+1} | \theta_1^s, \dots, \theta_N^s)$$

Where  $\theta_1^s, \dots, \theta_N^s \sim P(\theta_1, \dots, \theta_N | D)$  (Monte Carlo Approximation).

Now the most important is to generate samples from the posterior distribution. Generate independent samples would be ideal, but it is usually infeasible. In Markov Chain Monte Carlo (MCMC) the next generated sample depends on the previously generated sample.

The vector of parameters  $g$  does not concern by the sampling process while we can integrate this vector out this process is called the collapsed Gibbs sampling (Ishwaram and James, 2001), which is a particular form of an MCMC process. In Gibbs sampling all variables are initialized in an appropriate way, thus the replacement of  $\theta_1 = \theta_1$  by a sample of  $P(\theta_k | \{\theta_i = \theta_i\}_{i \neq k}, D)$ . Each  $\theta_k$  depends on its data  $D_k = \{x_{k,j}\}_j$  but is independent of the remaining data given the samples of the other  $\theta$ . The fact that subsequent samples are not independent is an unquestionable problem for collapsed Gibbs sampling. Nevertheless, a non-collapsed Gibbs sample can be used if the  $\theta$  are fixed in 4.8. By doing this the main advantage is given  $g$  independent sample can be generated from the auxiliary variables in a block. For

this reason, such sampling is called blocked Gibbs sampling. Thus, a sample can be generated from  $Z_k/g, D_k, k = 1, \dots, N$  as follows:

$$g/Z_1, \dots, Z_N \sim \text{Dir}(\alpha_1^* + N_1, \dots, \alpha_r^* + N_r) \quad (4.10)$$

where  $N_k$  indicates the number of times that  $Z_k = z^k$  in the current sample.

Now the  $\theta_k$  are treated as random variables. This means, when  $z$  indicates which value (modality) the variable (answer) takes  $\theta_k$  corresponds to a value associated with the  $k - th$  modality of the variable. Now we should put a prior on  $\theta_k$  with two hyperparameters that we denote  $h$  and learn  $\theta_k$  from data set. For the reasons that we have explicated earlier, a reasonable prior for the probabilities might

$$P(\pi|\alpha_0) = \text{Dir}(\cdot|\alpha_0/r, \dots, \alpha_0/r) \quad (4.11)$$

When  $M_k = 1$  and typically  $r \ll N$  equation 4.11 becomes a typical mixture model.

#### 4.4.2 Implementing a Dirichlet Process: Theoretical foundation and Practical Reasons

The usually known classification methods (agglomerative hierarchical clustering and divisive hierarchical clustering) do not allow choosing the number of clusters (Bellanger and Tomassone, 2014) or even knowing which one is accurate in a given situation. One of the best ways to classify objects is the use of clustering probabilistic model (Dahl and Newton, 2007). The mixture models belong to the family of such models.

The finite mixture model assumes that the data are from different subpopulations, each of them is modelled separately by a density function  $f_k(x, \theta_k)$  where  $x$  represents a  $k$ -dimensional vector of worker characteristics (wage, educational level, experience, demographic characteristics, . . .) and  $\theta_k$  an unknown vector of parameters. A vector  $\lambda = (\lambda_1, \dots, \lambda_n)'$  is introduced where  $\lambda_i = k$  if  $x_i \in k^{th}$  subpopulation. Now the clustering problem can be solved by estimating  $\theta = (\theta_1, \dots, \theta_k)$  and  $\lambda = (\lambda_1, \dots, \lambda_n)'$  to maximize the following likelihood function  $L(\theta, \lambda) = \prod_{i=1}^n f_k(x_i, \theta_{\lambda_i})$ . In order to obtain a known likelihood function generally researchers chose  $f_k(x, \theta_k)$  as the continuous density function of a  $k$ -dimensional Gaussian distribution with parameters  $(\mu_k, \Sigma_k)$ . Referring to our model this consideration means when  $x \in x$  is a continuous variable,  $p(x|g) = N(x, \mu_g, \sigma_g^2)$  where  $N(x, \mu_g, \sigma_g^2)$  is a probability density function of a normal distribution. This method was shown less effective than the discretization approach when  $p(x|g)$  is not Gaussian (John and Langley, 1995). Discretization consists in partitioning the domain of  $x$  into  $k + 1$  intervals as a pre-processing step. That implies  $x$  can be treated as a discrete variable with  $k + 1$  possible values. Let  $y_{gj}$  be the number of class  $g$



where  $x = X_j$  and  $n_g$  the number of elements belonging to class  $g$  given a data set, the expected value of  $p(x = X_j|g)$  is given by

$$\hat{p}(x = X_j|g) = \frac{\alpha_j + y_{gj}}{\alpha + n_g} \quad (4.12)$$

Let  $I_j$  denotes the  $j$ -th discretized interval, hence the classification process in a Bayesian classifier model with discretization uses  $\hat{p}(x \in I_j|g)$  as an estimator of  $\hat{p}(x|g)$  for each continuous variable in equation 4.1. Using the equation 4.1, based on the Dirichlet assumption the vector of the estimates  $\hat{p}(x \in I_j|g)$  admits a Dirichlet prior (Hsu, Huang and Wong, 2003). In a Dirichlet Process Mixture Models (DPMM), an particular stick-breaking process the random distribution function sampled from a DP is almost surely discrete (Blackwell, 1973; Blackwell and MacQueen, 1973).

Let denote  $f(x|g)$  the conditional PDF of  $x$ . When the function  $f(x|g)$  is not integrable everywhere, we have  $p(x \in I_j|g) = \int_{x \in I_j} f(x|g)dx \approx \sum_{D \subseteq I_j} \hat{p}(x \in D|g)$  where  $D$  is disjoint sub-interval in  $I_j$ . One way to increase the precision in estimating  $\hat{p}(x \in I_j|g)$  is to make  $f(x|g)$  integrable everywhere in each sub-interval  $D$ . That can be done by augmenting the number of sub-intervals in  $I_j$ . Hence the value of the estimate of  $\hat{p}(x \in I_j|g)$  can be as close as possible to the real  $p(x \in I_j|g)$  regardless the initial density function  $f(x|g)$ .

Using the property of perfect aggregation of the Dirichlet distribution deriving  $\hat{p}(x \in I_j|g)$  can be achieved by estimating the  $p(x \in D|g)$  such that each disjoint sub-interval in  $D$  is integrable everywhere. The estimate of  $\hat{p}(x \in D|g)$  is independent of the shape of the  $f(x|g)$  in interval  $I_j$  when  $\int_{x \in I_j} f(x|g)dx$  is fixed. Therefore, discretization methods including Dirichlet process are more applicable than the methods by assuming normal priors for continuous variables. The Dirichlet prior transforms the finite mixture into an infinite mixture then it allows dispensing with the assumption of stationarity of the Gaussian process derive from the normal density function.

## 4.5 Estimating the Dirichlet Process

We show earlier that the multinomial sampling with Dirichlet prior is a typical mixture model. Then to estimate the mixture model we use the Expectation-Maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) that we describe further. But for such algorithm the number of clusters should be known in advance or fixed. To avoid had hoc assumptions about the number of clusters to fix in the EM algorithm, we use a Dirichlet process which we estimate with the cut-point method to find the number of clusters to use in the EM algorithm.

### 4.5.1 Estimating the number of clusters by the Cutoff Method

The limits of the usually known classification methods (agglomerative hierarchical clustering and divisive hierarchical clustering) that we prove above oblige to choose a more accurate model to analyse the possibility of segmentation of the informal labour market. Hence we use one of the well-known methods among the Dirichlet process models the cut-point method.

Our observed data are described by a matrix that we denote  $X : (n * m)$  where  $X_{il}$  represent the value for individual  $i$  for a variable  $l, i = 1, \dots, N$  and  $l = 1, \dots, L$ . The values of  $X_{il}$  are categorical and continuous. Following Rossi, Gilula and Allenby (2001) we use the cut-point method. We model it on the basic point that  $X_{il}$  represent discrete values of underlying continuous  $X_{ij}$ . We obtain that all the  $X_{ik}$  values are pooled on  $R$ -point scale. When we assume that the information are pooled on a  $R$ -point scale that means we are on the continuum where the values have precise meaning. This implies that the continuous scores are rounded to the nearest of the  $R$  integers.

$$x_{il} = r \implies c_{r-1} < y_{il} \leq c_r \quad (4.13)$$

Hence  $c_0 = -\infty < c_1 < \dots < c_R = \infty$  define a partition of the real line  $(\mathbb{R}, \mathcal{B})$  into  $R + 1$  bins. Then the data vector is transformed into a count vector as follows:  $x_i = (x_{i1}, \dots, x_{im})'$  defines  $y_i = (y_{i1}, \dots, y_{im})'$  such  $y_{ir} = \sum_{j=1}^m I_{(c_{r-1} < x_{ij} \leq c_r)}$  where  $I_{(\cdot)}$  is the indicator function. The equation 4.2 is that in a cluster  $K$  the observations  $x_i$  are iid and follow a multinomial mixture distribution given by:

$$p(x_i | \theta_k) = \frac{(\sum_{r=1}^m x_{ir})!}{\prod_{r=1}^m x_{ir}!} \prod_{r=1}^m (\theta_{kr})^{x_{ir}} \quad (4.14)$$

Where  $\theta_{kr} = F_k(c_r) - F_k(c_{r-1}) \geq 0$ , and  $\sum_{r=1}^{R+1} \theta_{kr} = 1$  for  $k = 1, \dots, K$ . We should mention also  $F_k(c_r) = \sum_{t=1}^k \theta_{kt}$ , hence if we obtain the estimates of multinomial probabilities which are derived from an underlying continuous multivariate normal distribution, we also obtain the estimates of the component CDFs at the cut points. The conjugate prior for  $\theta_k = (\theta_{k1}, \dots, \theta_{k,R+1})$  is a Dirichlet given by

$$p(\theta_k | b) = \frac{\Gamma(\sum_{r=1}^{R+1} b_r)}{\prod_{r=1}^{R+1} \Gamma(b_r)} \prod_{r=1}^{R+1} (\theta_{kr})^{b_r-1} \quad (4.15)$$

Where  $b_r = b/R + 1$  and  $b$  controls the number of characteristics assigned to each cluster. Referring to the theoretical model we have  $G_0 = \text{Dirichlet}(b_1, \dots, b_r, b_{r+1}), G \sim DP(\alpha, G_0), \theta_k \sim G_0$  and  $x_i \sim \text{Multinomial}(n_i, \theta_i)$ . Here  $\theta$  leads the *a priori* expected number of segments.

### 4.5.2 Estimating the model parameters by the Expectation-Maximization Algorithm

Globally in a general model class of finite mixtures of regressions, the mixture is supposed to have  $K$  components where each of them is distributed according to a parametric distribution. Each observation has an a priori probability  $\pi_k$  to come from the  $k$ -th component, this probability is the weight of the concerned component. The weighted sum over the  $K$  components constitute the mixture distribution. When the weights of the components depend on further variables, these are called concomitant variables. Unlikely, to the theoretical model which assumes that revenues, human capital and other demographic variables and sector of activity separate the group of workers, we show that mixture of regressions is adequate to estimate model workers' heterogeneity and segmentation.

By denoting  $\phi = (\pi_1, \dots, \pi_k, \theta_1, \dots, \theta_k)$  the vector of all parameters for the mixture density function  $h()$ , let denote the response variable (earnings) by the vector  $y$ , and by  $x$  the vector of the predictor variables and  $\omega$  the vector of the concomitant variables. Let  $x_d$  and  $y_d$  describe two subsets of variables for  $d = 1, \dots, D$  and densities and let  $f_k$  be the component specific density given by a product over  $D$  densities defined for  $x_d$  and  $y_d$ .  $\theta_k = (\theta_{kd})_{d=1, \dots, D}$  define specific parameters of the component. Denoting  $\alpha$  the vector of parameters of the concomitant variable model, for the component weights  $\pi_k$  we have  $\forall \omega$

$$\sum_{k=1}^K \pi_k(\omega, \alpha) = 1 \text{ and } \pi_k(\omega, \alpha) > 0, \forall k \quad (4.16)$$

Hence forth, the model class is written as

$$h(y|x, \omega, \phi) = \sum_{k=1}^K \pi_k(\omega, \alpha) f_k(y|x, \theta_k) = \sum_{k=1}^K \pi_k(\omega, \alpha) \prod_{d=1}^D f_{kd}(y_d|x_d, \theta_{kd}) \quad (4.17)$$

When the number of components of a finite mixture model is known the expectation-maximization algorithm (EM) is the most used method for maximum likelihood estimation (Grun and Leisch, 2008). The EM algorithm treats the unobserved classes as the missing values and a missing data augmentation program is applied to the model. The algorithm supposes that an unobserved multinomial variable  $z_n = (z_{n1}, \dots, z_{nK})$  where  $z_n \in \{0, 1\}^K$  and  $z_{nk} = 1$  if observation  $(x_n, y_n)$  comes from component  $K$  and  $z_{nk} = 0$  otherwise. Moreover, for all  $n$ ,  $\sum_{k=1}^K z_{nk} = 1$ . This data augmentation is done by estimating the a-posteriori probabilities  $\hat{p}_{nk}$  which is given by

$$\hat{p}_{nk} = \frac{\pi_k(\omega_n, \alpha^{(i)}) f(y_n|x_n, \theta_k^{(i)})}{\sum_{k=1}^K \pi_k(\omega_n, \alpha^{(i)}) f(y_n|x_n, \theta_k^{(i)})} \quad (4.18)$$

This calculation corresponds to the E-step. Given the estimated values for  $\hat{p}_{nk}$  which are functions of  $\psi^{(i)}$ , in the M-step the new estimates  $\psi^{(i+1)}$  is obtained by maximizing the following quantity

$$Q(\psi^{(i+1)}|y|\psi^{(i)}) = Q_1(\theta^{(i+1)}|y|\psi^{(i)}) + Q_2(\alpha^{(i+1)}|\psi^{(i)}) \quad (4.19)$$

Where

$$Q_1(\theta^{(i+1)}|y|\psi^{(i)}) = \sum_{n=1}^N \sum_{k=1}^K \hat{p}_{nk} \log(f(y_n|x_n, \theta_k^{(i+1)})) \quad (4.20)$$

And

$$Q_2(\alpha^{(i+1)}|y|\psi^{(i)}) = \sum_{n=1}^N \sum_{k=1}^K \hat{p}_{nk} \log(\pi_k(\omega_n, \alpha^{(i+1)})) \quad (4.21)$$

Ones can maximise separately  $Q_1$  and  $Q_2$  while the maximization of  $Q_1$  produces new estimates  $\theta^{(i+1)}$  and the maximization of  $Q_2$  produces new estimates  $\alpha^{(i+1)}$ .

## 4.6 Empirical Analysis

The 1991-94 deindustrialization of the Haitian economy and the structural adjustment plans of 1986-1989 and 1996-99 spurred important changes in the macrostructure of the labour market. The market deregulation process (liberalization of the interest rate, removal of the trade barriers, elimination of import tariffs and quotas, etc.) (Aspilaire, 2014) has modified dramatically the components and the microstructure of the employment market. This poorly-regulated market is increasingly subject to changes unplanned by the policy makers.

### 4.6.1 Situation of the Labour Market and Data description

The labour market in Haiti is characterized by low regulation, paucity of jobs, wage discrimination and low wages. The main factors of the labour market are unemployment, inactivity, social exclusion and low regulation. The tough labour conditions of the market have discouraged approximately 66% of the 1.4 million people considered as unemployed in 2012 (Rodella and Scot, 2016) and half of a million of these people were actively seeking for a job. In this section we discuss the main characteristics of the Haiti's labour market and we aim a better understanding of the key characteristics through a factual analysis of its state until 2010.

Underemployment and long-term unemployment, high level of unemployment of the youth, important proportion of inactive among

the working-age population. Unemployment is not a provisional economic situation, but it is a structural fact of the economic reality.

Underemployment has an important part in the proportion of global unemployment. Some categories are more concerned with unemployment and underemployment than others, the proportion of women and rural inhabitant, which are unemployed is the double of the other categories. Labour force participation is around 61% which is below the Caribbean countries average of 65%, the participation of the men in labour force follows the same trend, but women participation rate (51%) is very close to the regional level which is 52% (Rodella and Scot, 2016). In 2001 more than 38% of the working age population were unemployed, 48% in metropolitan areas (Verner, 2008), in 2007 the proportion of unemployment fall down by 10% (Rodella and Scot, 2016) but this important decline did not imply an increase of the employment rates. This declination was the effect of abandoning and discouragement of the job seekers. But since 2007 an important rise in employment rate has been observed, reaching about 54% of the working age population. The formal market gives few opportunities to the younger workers with high level of schooling, 61% of unemployed people have attended the high school (Lamaute-Brisson, 2005), and more than 70% of informal workers attended at least the high school level, therefore we understand that informal employment is constituted by individuals excluded from the formal market. This observation can be explained by the fact that the unemployment rate is 3.1 times higher for younger workers than for other age groups (Lamaute-Brisson, 2005). The Haitian economy is not formalized and, moreover, since 1999 shadow economy is expanding in the economy regarding its size of the national GDP (Schneider et al., 2010; Aspilaire, 2014) therefore informality seems like a normal activity. The informal economy has an average size of 56.4% of the GDP over the period 1999 and 2007 (Schneider and al., 2010), over the period 2008-2010 informality has an average growth of 64.17% (Aspilaire, 2014). Henceforth the informal employment is persistent and represents a significant alternative to unemployment.

Analysing the labour markets (formal and informal) situation obliges the use of the empirical methods. Are the usual empirical models (segmented and competitive) enough adequate to highlight the complex characteristics of the Haitian labour markets? The next section brings answers to this question. In the Haitian economy, 57.1% of jobs are informal (IHSI and CELADE, 2007), while only 1.9% of jobs are provided by formal private firms (Aspilaire, 2014). Two related issues arise: the impact of the entry barriers in the regular employment market on informal employment and the efficacy of the labour market policies. Addressing the first issue will help to clarify the other one.

The Dirichlet process mixture model was elaborated previously to the Haiti labour market data for the year 2007. Our empirical analysis

is based on data pooled from a national survey regarding employment and the informal economy in Haiti, conducted by the national office of statistics Institut Haïtien de Statistique et d'Informatique over the period October 2007- March 2008 called Enquête sur l'Emploi et l'Economie Informelle. The survey encapsulates two steps: a household employment survey and a survey regarding the informal production units identified by the employment survey. The unit of observation is a given individual in a given year. The raw data consist in 30,337 observations. We pooled 14224 individuals 7579 are employed among them 5,842 individuals available for the analysis of informal employment. Our analysis includes urban and rural workers, but excludes agricultural activities. Our sample is limited to individuals aged between 14 and 65 years. We exclude workers without wages or revenues (832).

TABLE 4.1: Summary statistics on the labour market

	Total	Inactive	Active	
			formal	informal
Sample (%)	100	46.71	8.14	69.10
Monthly revenues	4606	-	7025	4306
Sex (female = 1%)	58.29	64.65	47.39	54.84
Age (years)	33.15	27.42	37.17	37.20
Geographic (%)				
Urban	80.37	76.85	93.00	81.66
Rural	19.63	24.55	7.00	18.34
Education (%)				
None	18.52	15.84	14.01	21.17
Low	34.02	33.83	23.29	35.42
Medium	44.27	49.85	45.11	39.77
High	3.19	0.48	17.59	3.63
Marital Status (%)				
Single	39.88	61.40	28.99	24.16
Married	24.56	12.66	41.21	32.00
Free union	27.68	20.44	22.48	34.01
Separated	7.88	5.50	7.33	9.83
Household Size	5.51	5.84	5.25	5.30

Source: EEEL, 2007. Computations by the author. Monthly revenues in local currency (gourde).

For the analysis of the labour market we choose the variables that the literature generally admits as these variables that globally influence the wages function. Hence we choose the following variables

**Revenues:** monthly revenues (real numbers) of each interviewed individual in the local currency unit (gourde)

**Sex:** gender of each person (man and woman)

**Age:** Age of workers (real numbers) when the survey took place

**Geographic settings:** The living place of the interviewed individual (rural and urban)

**Marital Status:** marital status divided in four modalities (single, married, free union and separated)

**Education:** educational level (illiterate, elementary, middle and high school, and university)

**Working Hours:** number of hours (real numbers) that individual workers spend at work per week

**Industry:** industry of activity (manufacture, trade, building and construction, services, and others)

**SPC:** socio-professional category (trainee, salaried, self-employed, and entrepreneur)

$Age^2$ : when we take the square of the age and we add it to the age we approximate the effect of the experience. Denoting  $\alpha$  and  $\beta$  the respective coefficients of the age and the age squared and  $\bar{x}$  mean of the age, the average effect of experience is given by the sum:  $\alpha + 2\beta * \bar{x}$ .

**Household size:** the size (entire number) of the household of each individual worker interviewed during the survey.

## 4.6.2 Testing Labour Markets Differences

In order to analyse the labour market segmentation or competitiveness, two important aspects must be eluded. Are the labour markets (formal and informal) segmented and which kind of segmentation (each segment has a unique wage function) (Gunther and Launov, 2011). Are the labour markets perfectly competitive?

According to Dickens and Lang (1985) and Basu (1997), a labour market is segmented when there are entry barriers into the upper segment to block workers' free mobility. In this section we first test the hypothesis of the labour markets differences. Hence, we estimate the model of discrete choice between sectors (formal and informal) by bivariate probit. We also estimate the labour force participation decision through a standard univariate probit.

Analysing the results in table 4.2, we find the significant negative influence on labour participation of women, experience (age squared) and living single. We find that university level and age has a highly

TABLE 4.2: Labour Market Participation and Segments Comparison

	Univariate Probit	Probit
Dependent variable	Participation	formal sector
Intercept	-2.035	0.014
<b>Gender (Reference: Men)</b>		
Women	-0.561*** (-22.021)	-0.801*** (-21.702)
Age	0.157*** (25.494)	-0.011 (-1.085)
Age <sup>2</sup>	-0.002*** (-21.428)	0.0003* (2.491)
<b>Education (Reference: Illiterate)</b>		
Elementary school	-0.039 (-1.080)	0.345*** (7.727)
Middle school	0.067* (2.315)	-0.320*** (-6.987)
University	0.834*** (7.327)	0.16... (1.837)
<b>Marital status (Reference: free union)</b>		
Married	0.038 (1.009)	0.087* (2.052)
Separated	-0.083 (-1.584)	0.016 (0.240)
Single	-0.027*** (-7.591)	0.076 (1.382)
<b>Geographic settings (reference: urban)</b>		
Rural	0.068* (2.277)	1.114*** (28.049)
Household size	0.017*** (3.795)	-
<b>Socio-professional category (reference: salaried)</b>		
Trainee	-	-0.425*** (-5.928)
Self-employed	-	-0.577*** (-13.132)
Entrepreneur	-	-0.959*** (-7.234)

Source: EEEI, 2007. Computations by the author, significance codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; ..=10%.

significant positive effect on the labour market participation while rural individuals, middle-school level people and the size of household have low positive but significant influence on the labour participation.

Being a woman, attending middle school are two characteristics which have a negative effect on the participation in the formal labour market. A trainee does not have too much chance to be on the formal



market. This can be explained by the fact that the rules and law do not oblige enterprises and institutions to give a training period to the trainee and apprentices. We also find that self-employed individuals and entrepreneurs have negative chance to participate in the formal market. This is the effect of the entry barriers to undertake activities on the formal market, such as the expensive cost to start an enterprise and the long-last time it takes to register an enterprise (World Bank 2012). Unlikely on the formal sector, trainee, self-employed and entrepreneurs have more opportunities to evolve on the informal sector because the entry barriers are less rigorous. When an individual worker has attended the university or elementary school level, he is likely most attracted by the formal market. People who attended elementary school has a low level of expected reserve salary. Workers with the university level is seeking for a normal job where promotion and social security are granted. People who live in the urban area are most likely entering the informal sector than the formal.

The coefficients of the human capital variables (school, age and experience) show that market participation behaviour across the two sectors are very different. The most remarkable and significant variable on the two different models is the education level. It seems like education is more valuable in the formal market than the informal market.

### 4.6.3 Can a unique wage function represent the global labour market?

We evaluate the assumption considers each a priori segment of the labour market is defined by a unique wage function.

We use univariate models. We evaluate a wage function for each a priori component, to which we give the following form

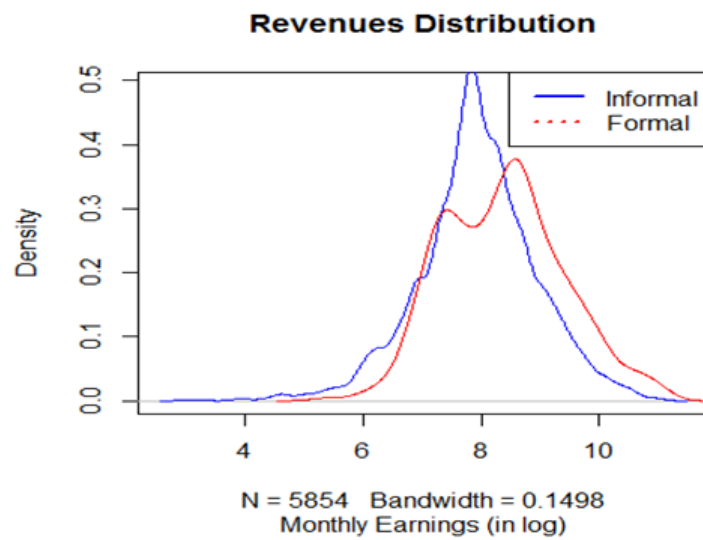
$$y_{ij} = x_i' \beta_j + u_i$$

Where  $y_{i,j}$  determine the revenues for individuals  $i$  in segment  $j$  and  $x_i'$  the individual characteristics (educational and sociodemographic) of individuals.

The kernel density estimates of monthly revenues earned by formal and informal agent ( figure 4.2) shows, despite the important discrepancy in means between formal and informal wages (4.1), the difference in densities is not much as important as it seems. The overlap of the density function of the revenues (figure 4.2) show that not all informal workers earn less than formal workers.

The analysis of the wage functions across the sectors is based on the estimates of three OLS univariate models of the structure of the wage functions. We proceed as if the income is the influenced variable, hence the results cannot be considered as due to errors of measurement in income.

FIGURE 4.2: Densities of monthly log-wages for formal and informal workers



Source: author's design from EEEI 2007 data.

TABLE 4.3: Wage functions for the sectors: OLS models

Dependent variable $\ln(\text{wages})$	Both Sectors	Formal Sector	Informal Sector
Intercept	4.423***	4.958***	6.003***
<b>Education (Reference: Illiterate)</b>			
Elementary school	-0.319*** (-4.543)	-0.365* (-2.27)	-0.217*** (-3.298)
Middle school	0.269*** (3.991)	0.529*(2.569)	0.183** (3.227)
University	1.4758*** (9.991)	1.857*** (5.161)	1.044*** (7.76)
Age	0.1114*** (7.956)	0.122*** (3.657)	0.095*** (7.37)
Informal Sector	1.633*** (27.164)	-	-
<b>Socio-professional category (reference: salaried)</b>			
Trainee	-3.435*** (-30.639)	-4.207*** (-14.981)	-3.076*** (-30.504)
Entrepreneurs	0.0305 (0.166)	-0.596 (-0.935)	0.435** (2.912)
Self-employed	-0.887*** (-13.188)	-2.089*** (-11.488)	-0.363*** (-6.21)
$Age^2$	-0.0013*** (-7.226)	-0.0013*** (-3.351)	-0.001*** (-5.498)
$R^2$	0.243	0.181	0.210
Number of observations	7579	2294	5285

Source: EEEI, 2007. Computations by the author, significance codes: \*\*\*= 0.1%; \*\*=

1%; \*= 5%; ...=10%.

Results (table 4.3) show globally all the school levels influence the wages of the formal sector more significantly positively or negatively. The classification is explained significantly through the earnings as variable explained by the workers' characteristics and socioeconomic variables. However, the values of R-squared (table 4.3) of each component show the discriminant power of the components are very low and insignificant because R-squared is less than 0.5. Therefore, we cannot use OLS models for wage functions to analyse the market segmentation because such function has a poor capacity of explanation of the segmentation structure of the labour market. Moreover, the differences between the wage functions are not important for human capital gains. All these observations are coherent with the findings of Cain (1976) who stipulates that OLS techniques are inappropriate to make the comparison of the wage functions between formal and informal sectors.

According to table 4.4, where the average wages of the formal sector are greater than the average wages of the informal sector then potential wages are different across sectors. This means, labour markets (formal and informal) may be weakly competitive (Magnac, 1991). Therefore, individual workers can take benefit from comparative advantages by working in a sector or in another one with a standard education level. These individual advantages cannot be cleared from the global market, except probably a small fraction of individual comparative advantages. It could also be revealed that markets can be segmented. Nevertheless, these two fundamental assumptions cannot be tested in the discrete choice models. For this reason, we estimate a generalized bivariate Tobit model in two steps also called Heckit model.

In order to analysis an eventual segmentation of the labour market we evaluate the assumption that the Haitian labour market can be defined by a unique wage function which is the opposite of the assumption of Gunther and Launov (2011) which considers each segment of the labour market is defined by a unique wage function. We evaluate a unique wage function as follows:

$$y_{ij} = x_i' \beta + u_i$$

Where  $y_{ij}$  determine the revenues for individuals  $i$  in and  $x_i'$  the individual characteristics of individuals. OLS techniques are inappropriate for the simultaneous analysis of the wage functions across formal and informal sectors (Cain, 1976). Therefore, we use an Heckit (a two-steps Tobit) model which allows also to test the hypothesis that the labour market participation is influenced by human capital accumulation.

Due to the fact that we exclude individuals who not gain wages from the estimation can be a source of bias which we take into account in the model through the term of selectivity  $\rho$ . The non-significance of  $\rho$  shows that individuals who are employed in the labour market

TABLE 4.4: Tobit (two-step Heckman) Model for a Unique Labour Market

Dependent variable	Probit Participation	Tobit ln(wages)
Intercept	0.82	6.632
<b>Gender (Reference: Men)</b>		
Women	0.757*** (20.77)	-0.279** (-2.917)
Age	-0.0086*** (-4.808)	0.0704*** (9.975)
Age <sup>2</sup>	-	-0.0007*** (- 8.726)
<b>Education (Reference: Illiterate)</b>		
Elementary school	-0.475*** (-10.167)	-0.186** (-2.781)
Middle school	0.323*** (7.493)	0.131** (2.849)
University	-0.225** (-2.913)	0.867*** (10.902)
<b>Marital status (Reference: free union)</b>		
Married	-0.132** (-2.993)	0.037 (0.976)
Separated	0.047 (0.685)	0.0142 (0.292)
Single	-0.153** (-2.972)	-0.0738 (-1.625)
Household size	-0.0037 (-0.556)	-
<b>Socio-professional category (reference: salaried)</b>		
Trainee	-	-0.467*** (- 6.969)
Self-employed	-	-0.166*** (- 4.652)
Entrepreneur	-	0.424*** (5.525)
Number of working hours/week	-	0.0017** (3.027)

Source: EEEI, 2007. Computations by the author, significance codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; !=10%.

TABLE 4.5: Selectivity of the model

Number of observations	6734
$\sigma$	0.924*** (74.66)
$\rho$	0.068 (0.18)

Source: EEEI, 2007. Computations by the author, significance codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; !=10%.

have almost similar characteristics than those who are not. Then a mixture model for the Haitian labour market without the selectivity term  $\rho$  would give unbiased estimators. The Heckit model shows that education is the most powerful indicator of the wages level in the global market (formal and informal). The entrepreneur is the second most important indicator of the wages level. However, woman and individuals who live in the rural agglomeration are likely earning less. We observe opposite signs of the main estimates between the labour market participation and the wages function (table 4.4). Education, gender, age, and marital status show opposite sign between the two models, this means a global unique wage function cannot represent the global market.

The fact that we find a mixture model for the Haitian labour market without the selectivity term would give unbiased estimators and the result that shows the labour market cannot be represented by a unique wage equation allow us to adopt the Gunther and Launov assumption of a unique wage function for each sector (2011). Furthermore, this assumption is evaluated through a mixture model (Dirichlet Process) which can detect segmentation while it does not take into account the selectivity term. This model is also estimated to bring answers to the main questions related to the idea that informal employment is a voluntary choice of the individuals as an employment opportunity we must test the existence of entry barriers into different sectors.

#### 4.6.4 Structure of the labour market: Segmentation or Competition?

The infinite mixture model applied to the data allows the informal employment market being constituted of any number of segments. We transform all the categorical  $X_{il}$  in ordinal values and we discretize all the continuous  $X_{ij}$  we obtain that all the  $X_{ik}$  values are pooled on a 6-point scale.

In Bayesian inference results are usually sensitive to the parameters that determine the probabilities of the model as well as to prior for parameters that determine the fit of the model. Therefore, we let the data determine the value of mixing proportions as it was drawn randomly from uniform Dirichlet. For the same reason, we do not initiate the value of *theta* parameters, we let it also to be drawn randomly from uniform Dirichlet. We fix the convergence parameter as low as possible ( $1 * 10^{-8}$ ) to have significant clustering results.

We first obtain a two-segment informal labour market, although we estimated the model with four initial segments. The values of the vector *theta* are given for each component and the value of *lambda* represents the value of mixing proportion for each segment (table ??5). For each new run of the model (with a change of variables), we obtain a similar table with new values.

TABLE 4.6: Component and mixture proportion for the partitioned labour markets

	1-segment	2-segments	3-segments	4-segments
lambda	0.2059	0.0407	0.6928	0.0604
$\theta_1$	0.3374	0.3227	0.3452	0.1254
$\theta_2$	0.1907	0.2076	0.1672	0.4158
$\theta_3$	0.2546	0.2426	0.2071	0.1878
$\theta_4$	0.2165	0.2120	0.2087	0.2016
$\theta_5$	0.0000	0.0000	0.0684	0.0669
$\theta_6$	0.0000	0.0148	0.0028	0.0018
$\theta_7$	0.0005	0.0000	0.0003	0.0002

Source: EEEI, 2007. Computations by the author

The values of  $\theta_k$  (table 4.5) represent  $\theta_k \equiv P(c_{k-1} < Y_{il} \leq c_k)$ . The values of lambda (0.629) show that the greatest value of the mixing proportion is given by the 3-segments model. This is an important indicator for the three-segments model.

Four criteria are usually used to evaluate the number of segments: Akaike Information Criterion (AIC) (Akaike, 1973), Bayesian Information Criterion (BIC) (Schwarz, 1978), Consistent Akaike Information Criterion (CAIC) (Bozdogan, 1987) and Integrated Completed Likelihood (ICL) (Biernacki, Celeux, and Govaert, 2000). AIC recovers the real number of clusters when one of the segments is way smaller than others. BIC is independent of the mixture proportions. CAIC is generally more effective than BIC. ICL is more robust than BIC by respecting more mixture model assumptions. The results (table 4.6) show that the entire four criteria impose a three-segments market (formal, and two levels of informal). However, taking out any of the variable does not change the number of clusters.

TABLE 4.7: Criterion indexes for the number of segments of the global market

	Homogeneous	2-segment	3-segment	4-segment	Winner
AIC	-63203.90	-32408.10	-32319.74	-32326.50	3-segments
BIC	-63227.26	-32458.16	-32396.50	-32429.96	3-segments
CAIC	-63230.76	-32465.66	-32408.00	-32445.46	3-segments
ICL	-63227.26	-32457.60	-32395.74	-32429.08	3-segments
Loglik	-63196.90	-32393.10	-32296.74	-32295.50	4-segments

Source: EEEI, 2007. Computations by the author

When we run the model by taking out each variable separately we obtain also a three-segment labour market according to the values of AIC, BIC, CAIC, and ICL. Now we find that the global labour market (formal and informal) is heterogeneous hence for modelling the unobserved heterogeneity we use a semi-parametric method which are the finite mixture models.

TABLE 4.8: Semi-parametric mixture model determining wage functions for the global labour market

Dependent variable: wages	Formal sector	Informal1	Informal2
(Intercept)	0.1567	0.1196	0.0848**
Sex (Reference: Men)	-0.1257 (-1.124)	-0.1024*** (-5.11)	-0.0281*** (-5.22)
<i>Education (Reference: Illiterate)</i>			
Elementary school	-0.1806 (-1.144)	-0.0564* (-2.327)	-0.0195** (-3.055)
Middle school	0.1221 (0.971)	0.0874*** (3.933)	0.0217*** (3.864)
University	1.4299*** (7.497)	0.4491*** (8.831)	0.1355*** (5.749)
Age	0.0694* (7.956)	0.0187*** (3.900)	0.007*** (5.265)
<i>Socio-professional category (reference: salaried)</i>			
Trainee and apprentice	-0.9304*** (-3.395)	-0.249*** (-7.402)	-0.1501*** (-15.557)
Entrepreneurs	0.9251*** (4.810)	-0.2719*** (-6.807)	0.7850** (27.961)
Self-employed	0.1191*** (-13.188)	-0.0357 (-1.362)	-0.0401*** (-6.016)
$Age^2/100$	-0.0961** (-2.638)	-0.0218*** (-3.701)	-0.0074*** (-4.587)
<i>Geographic settings (reference: urban)</i>			
Rural	0.0028 (0.1801)	-0.0174 (-0.759)	-0.0218*** (-3.630)
<i>Marital status (Reference: free union)</i>			
Married	0.2436'' (1.954)	0.0136 (0.647)	0.0104'' (1.794)
Separated	0.0844 (0.456)	0.0197 (0.610)	0.0060 (0.741)
Single	-0.118 (-0.778)	-0.0601* (-2.478)	-0.0068 (-0.908)
Number of working hours/week	0.0005 (0.231)	0.0015*** (3.877)	0.0003** (3.259)
$\hat{\pi}_j$	0.0745	0.2761	0.7877
Prior Probability	0.096	0.331	0.573

Source: EEEI, 2007. Computations by the author; significance codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; ''=10%.

The interpretation of the results (table 4.7) shows that formal market is dominated by workers who attended the university. those who are entrepreneurs, and the married people. Trainees and apprentices have a lot of difficulty to integrate the formal labour market. The first informal labour market that we denote informal 1 is led by all

the education level, but particularly by the university level, which has a positive influence on wages and elementary school which has a negative impact on the revenues. Women, entrepreneurs and trainees had a negative impact on the earnings of the first informal market. Analysing the educational level and the gender on the second informal market, the situation is similar to the first informal market. Trainees, self-employed and experience influence wages negatively on the second informal labour market. The number of working hours per week impact positively both informal market sides. Unlike the two sides of the informal market, on the formal market the number of working hours is not a significant factor because work duration is standardized on this market.

Regarding these results, it is a fact that the labour markets (formal and informal) are subject to segmentation, but the number of segments must be considered with some technicality. Now, it is very important to explain this multi-sector characteristic of the informal employment market. Base on this explanation we analyse the mobility across segments or the entry barriers into the segments.

Now we know that the informal employment market has two different segments. Are these two segments meant the labour market is segmented? Consider the definition of the market segmentation according to Dickens and Lang (1985), and Basu (1997), we should determine whether workers are free to move from a side to another one or not. In the Dirichlet process model the importance of the variables cannot be explicitly observed. To identify which variable has the most influence over the clustering process, we run the EM (Dempster, Laird, and Rubin, 1977) algorithm in order to estimate the multinomial cut-point of our Bayesian semi-parametric model.

TABLE 4.9: Component and mixture proportion for the two-segments informal labour market

	1-segment	2-segments	3-segments	4-segments
lambda	0.2044	0.0024	0.7436	0.0344
$\theta_1$	0.3047	0.0000	0.3152	0.1091
$\theta_2$	0.2275	0.3542	0.1645	0.3721
$\theta_3$	0.2122	0.3222	0.2151	0.2375
$\theta_4$	0.2539	0.1979	0.2426	0.2168
$\theta_5$	0.0000	0.1192	0.0614	0.0639
$\theta_6$	0.0010	0.0000	0.0010	0.0003
$\theta_7$	0.0003	0.0006	0.0003	0.0000

Source: EEEL, 2007. Computations by the author

The values of lambda (0.0024) show that the lowest value of the mixing proportion is given by the 2-segments model. We analyse further the number of segments that exists in the informal labour market.



TABLE 4.10: Criterion indexes for the number of the two-segment informal market

	Homogeneous	2-segment	3-segment	4-segment	Winner
AIC	-67062.20	-29769.72	-29744.78	-29751.19	3-segments
BIC	-67085.17	-29818.95	-29820.26	-29852.93	2-segments
CAIC	-67088.67	-29826.45	-29831.76	-29868.43	2-segments
ICL	-67085.17	-29818.35	-29819.59	-29852.20	2-segments
Loglik	-67055.20	-29754.72	-29721.78	-29720.19	4-segments

Source: EEEL, 2007. Computations by the author

Analysing the results (table 4.9) criterion indexes BIC, CAIC, and ICL we obtain a two-segments informal labour market model. The AIC value indicates a three-component model, but we know this criterion gives the real number of clusters when one of them is way smaller than the others.

We also observe an important gap between the positive impact of the education levels of informal workers in the segment 1 compared to the second segment (table 4.10). In the first segment of the informal market (informal 1) only entrepreneurs, and individuals who have attended university have a positive impact on the revenues. In the second segment of the informal market (informal 2) individuals gain wages their characteristics do not matter, because the coefficient of the intercept (0.0918) is significantly positive (table 4.10). Elementary educational level and the number of working hours have a positive effect on the wages. Hence, this side (informal 2) of the informal market is the very accessible segment. The segment 1 is harder for women than the second segment of the informal market (table ??).

Therefore, the question whether informal workers are free to move from one side to another of the labour market has a negative answer. Hence, we must reject the assumption of unlimited sector mobility which implies a fully competitive market. We observe two main entry barriers across the upper segment. Education is the first entry barrier into the upper side because its high significant coefficient implies that this segment access is given only to highly educated workers. The second fundamental entry barrier is the initial capital because this segment is expressly dedicated to entrepreneurs. The non-significance of the number of working hours per period for informal workers of the first segment explains that informal workers in the upper side prefer maximizing their time utility and their revenues simultaneously when they are educated and have the financial resources. They are those who voluntarily chose to evolve in informality. We deduce income rationality and time utility are the only motivation for individuals with human capital accumulation to become voluntary informal.

That also means informal work provides some important non-monetary advantages, which is not the case in the regular labour market. This result is also consistent with our initial assumption that

TABLE 4.11: Semi-parametric mixture model determining wage functions for the global labour market

<b>Dependent variable: wages</b>	<b>Informal1</b>	<b>Informal2</b>
(Intercept)	0.2489	0.0918**
Sex (Reference: Men)	-0.1352* (-2.644)	-0.0417*** (- 6.632)
<b><i>Education (Reference: Illiterate)</i></b>		
Elementary school	-0.1062 (-1.314)	-0.0243*** (- 3.348)
Middle school	-0.0052 (-0.082)	0.0225*** (3.482)
University	0.9445*** (7.794)	0.2075*** (9.113)
Age	0.0432** (2.685)	0.0088*** (5.778)
<b><i>Socio-professional category (reference: salaried)</i></b>		
Trainee and apprentice	-0.5947*** (- 4.573)	-0.1484*** (- 13.472)
Entrepreneurs	0.6478*** (5.228)	0.8274*** (3.730)
Self-employed	0.0142 (0.203)	-0.0344*** (- 3.544)
$Age^2/100$	-0.0005*** (- 2.866)	-0.00005*** (-5.122)
<b><i>Geographic settings (reference: urban)</i></b>		
Rural	-0.0522 (-0.686)	-0.0197** (- 2.874)
<b><i>Marital status (Reference: free union)</i></b>		
Married	0.1174" (1.810)	0.0054 (0.813)
Separated	0.0789 (0.417)	0.0027 (0.283)
Single	-0.0405 (-0.505)	-0.0096 (-1.192)
Number of working hours/week	0.0009 (0.792)	0.0004** (4.213)
$\hat{\pi}_j$	0.171	0.829
Prior Probability	0.212	0.788

Source: EEEI, 2007. Computations by the author; significance codes: \*\*\*= 0.1%, \*\*= 1%, \*= 5%; " =10%.

individuals working in informality are also maximizing their utility.

Analysing the results of the table 4.9, we observe that the gap between the revenues of the two segments is similar to the education gap between the same two segments.

Knowing the probability for a worker with specific characteristics to choose voluntarily informal activities is very important. It could

be useful for the economic policies which aim to decrease informal employment and to improve the attractiveness of formal activities in order to reduce the number of voluntary informal workers. The following results (table 4.12) provide probabilities for workers to evolve in the two segments of informality, high-paid (segment 1) and low-paid (segment 2).

TABLE 4.12: Probability to belong to one segment for global labour market and informal labour market

	Global market			Informal market	
	Formal	Informal 1	Informal 2	Informal 1	Informal 2
$\pi_j$	0.096	0.331	0.573	0.212	0.788
$\hat{\pi}_j$	0.074	0.266	0.659	0.171	0.829

Source: EEEI, 2007. Computations by the author,  $\pi_j$  :prior probability;  $\hat{\pi}_j$  : estimated probability

In the global market and the informal market, the values of  $\pi_j$  (table 4.12) indicate the proportion of individual workers who should evolve in each part of the labour market according to their personal characteristics in order to maximize their income. This calculation includes two steps. First attribute to each worker the cluster where his revenue is maximum. Second, determine for each cluster the proportion of individual workers that this cluster is more profitable for. Hence, if the theoretical probabilities ( $\pi_j$ ) are equal to the estimated probabilities ( $\hat{\pi}_j$ ) then the global labour market is not similar to a queuing system where individual workers are in waiting lines, because that means each individual worker evolves in the sector that maximizes his revenues. That would mean there exists no entry barriers across the segments which implies a competitive labour market. In the global market, we observe that in the formal market and informal 1,  $\pi_j$  is greater than  $\hat{\pi}_j$  and in the informal 2,  $\hat{\pi}_j$  is greater than  $\pi_j$  that means there are entry barriers in the formal sector and in the informal 1. Hence the individual workers who cannot bypass the entry barriers of the formal sector switch to the first informal market and the others who are unable to bypass the barriers of the first informal market enter the second informal labour market.

In the informal labour market, the individual behaviour regarding the first segment can be assimilated to a maximization revenues behaviour. Highly educated workers who dispose of financial resources dominate the first component of the informal market. Hence, the 4.1% difference between the maximization income behaviour and the individual behaviours represent the individuals who cannot drop the entry barriers in this segment. For this reason, they (4.1% of the workers) switch to the second segment of the informal market.

We actually know that we have a two-segment informal labour

market with voluntary and involuntary (ejected from the formal market) workers. Regarding the results above (table 4.12), we can determine the part of informal workers who have individual comparative advantages (Gindling, 1991; Maloney, 1999, 2004; Perry et al., 2007) and they would not do better in the formal market.

Following the theoretical approaches aforementioned, the share of workers that would not do better out of the informal market is the sum of workers in each informal segment following the distribution of revenues-maximizing. According to the results (table 4.12) there are 17.1% of informal workers taking benefit from individual comparative advantages, so they would not do better in the formal sector than in the informal market. Therefore, approximately 83% of informal employment are involuntary the informal segment comprises 78.8% of individuals excluded from the formal market.

The entry barriers of the formal market, reduce the remuneration that individuals with adequate characteristics to evolve in the formal market would gain if the labour market was competitive. Despite the weakness of the number of individuals working in the formal market the entry barriers are still enormous.

## 4.7 Conclusion

The Dirichlet process mixture model that we propose for designing the informal labour market allows detecting unobserved sub-segments in informal employment. This Bayesian model also allows determining the causes of informal employment segmentation between free mobility across all the sectors or entry barriers into the segments. In line with Gunther and Launov (2011), our model contains an implicit approach for analysing whether the informal work is voluntary or not. Our results show that informal labour market is segmented. We obtain a two-segment market. Therefore, we conclude that none of the theoretical explanation of labour market in developing countries (expansion of informal employment) provides a sufficient way to understand labour market structure.

Besides the earnings rationality, the main factors as entry barriers into the higher-paid side of the informal labour market are education and financial resources. Moreover, the fact that we find that work duration was not significant for workers on the upper side of the informal market that confirms our initial hypothesis that informal workers tend to maximize their utility when they are highly educated and financially stable. Therefore, they obtain some non-pecuniary advantages in the informality that formal employment probably cannot provide them. Such workers are identified to be the voluntary individuals evolving in the informal labour market. In the informal market workers can maximize their revenues and their utility simultaneously.

Our model shows that 17.1% is the share of workers that have individual comparative advantages in the informal market. They represent workers that would not do better out of the informal market, which is the sum of workers in each informal segment following the distribution of revenues-maximizing. We find that the main share of informal work is involuntary. But for the workers who take benefit from comparative advantages the labour market is an integrated market that offers desirable jobs with various characteristics. Moreover, these particular workers make the choice to evolve in informality. Furthermore, the same labour market for relatively unskilled individuals (inexperienced and low educated) place entry barriers into the formal side and oblige them to subsist in the informal sector.

Our model can be applied to study the segmentation of any market (formal or informal). This model can be used as informality prevention for public policies which want to block other individuals to join the informal labour market by increasing the formal market attractiveness. With all the individual features of the workers' computing in the model we determine his probability to become voluntary or involuntary informal. The model can be used also in designing employment policies to reduce involuntary informal employment by creating more stable jobs regarding the characteristics of involuntary informal workers.

Our semi-parametric Bayesian model can be subject to further research. Estimate the expected revenues for workers by using the knowing characteristics into each segment would be a very interesting avenue. Ones could think to extend the model to a spatial Dirichlet process mixture to obtain such results.

## Conclusion

All through the schools of economic thought that have studied the informal economy, approached informality as if it were an evil in itself. Whether it is neo-Marxist thought or dualistic conception, whether it is the structuralist thought of the Keynesian and neo-Keynesian school, or the legalistic thought of the neo-classical and liberal schools, all have approached informality with this view. They all propose economic policies to adopt and the measures that must be taken to combat informality. Due to the fact that they do not approach the informal economy from a favourable angle, the different theories elaborated have a hard time shedding light on the informal in Haiti. The ever-increasing magnitude of the phenomenon of the informal economy raises the question of whether the angle under which these great theorists approached the informal economy remained valid for the countries of Latin America and the Caribbean, particularly Haiti. If the informal economy is seen as a problem of the economy, so what solutions can the policy makers proposed to a country where it represents the larger part of the economy? The informal sector deserves to be approached differently than a phenomenon with a negative effect on the global economy. The major socioeconomic problems that have affected Haitian society and which are the main reasons why people have embarked massively in the informal are still persistent. These include poverty, unemployment and poor living conditions. In 2012 70% of Haitians and 23% of the Haitian population would suffer from extreme poverty (ONPES and World Bank, 2014). So how can one think of solving a problem whose conditions of existence are reinforced daily by powerful factors? It becomes legitimate to question the validity of theories on the informal economy elaborated by the various schools of economic thought as an economic problem in Latin America and the Caribbean, particularly in Haiti, Question from that point of view without thinking in the case where situations of informality can become irreversible. Viewing the informal economy only from a problematic angle, the theoreticians of the main schools of thought who have studied the informal sector have to model these cases in order to make appropriate economic measures and policies. The dualist modelling of developing-country economies based on segmented labour market assumptions has constructed a predominantly static picture of the informal economy, which underpins that the informal economy does not have its own dynamics. We have shown that the informal labour market is segmented, but beyond that, the informal labour market has its own dynamics. The analysis of the impact of neoliberal policies on the evolution of the informal

economy shows that it was resilient to the shocks of the labour market during the 1980s, to economic and financial shocks 1986, 1991 and 2004. In spite of these economic shocks and the labour market, the informal economy has not stopped growing. The informal economy is more resilient to shocks than the structured economy, which always responds negatively. This economically interesting aspect of the informal economy does not suggest that it is an economic downturn. If it is supported and framed by the formal in this dynamic, could informal economy help to cushion the current economic shocks? As little as this question remains unanswered, but the fact is there, in a situation of shock by the demand the informal economy proves to be much better than the formal economy. Then, in the case of a shock in labour supply, the capacity of the two sectors for absorption is based on the quality of the informal market compared with the formal ones, as soon as the former do not accuse a significant inferiority gap, the two sectors may each absorb part of the unemployment. The elasticity of the supply of self-employment is a key factor in determining the informal sector as a shock absorber. The shock absorber level of the informal shifts in the same direction as the elasticity of supply Informal work. As a result, it appears that the informal economy is a better shock absorber of labour supply than the formal economy, if the conditions of minimum quality compared to the formal market are respected and if the elasticity condition of the self-employment is present. In this case, the informal economy would be the shock absorber that absorbs most of the unemployment. When we consider that the aggregate absorptive capacity of the urban system depends not only on the elasticity of output and the response of informal employment to shocks, but also on the weight of the informal sector in total urban employment. In this line the informal economy would become the best weapon against unemployment alongside the formal economy if it wants to frame it with the immense means at its disposal. On the other hand, the results of the informal dynamics raise the question of whether the image of the lethal element of the global economic system, projected by economic theorists and analysts from the informal economy as a whole would not deserve another appreciation. Apart from the fact that it is on the margins of all social, legal and economic regulations, should the informal economy deserve a consideration other than that which theory recognizes it today? Even without an answer to this question, the study of the dynamics of the informal economy shows that it has no worse sides than it has been attributed and even it has less because it can serve as an instrument of absorption to the various possible shocks that the global economy could face. This study of the dynamics of the informal economy calls into question the ideas adopted on the informal economy, which claimed that the informal economy was only a secondary part of the economy itself without any dynamics and possible perspective (Harris and Todaro, 1970). This

study exposes all the weaknesses of the theories erected on the informal economy as a factor itching the global economy, as the case of Haiti has shown that some of these theories could fit but their recommendations of economic policies are not appropriate. In addition to exhibiting the countercyclical nature of the informal economy in Haiti, this study also showed that the informal labour market is not fundamentally a market where evolve only the excluded from the formal labour market, unlike some theorists that advocate informal economy is a leftover of pre-capitalism constituted of marginalized individuals (Chen et al., 2004; Williams and Round, 2008). This work shows that deregulatory and neoliberal policies have influenced positively the informal economy growth unlike this theory that argues deregulatory can reduce informal economy growth (Nwabuzor, 2005; London and Hart, 2004). This work has also shown that informality does not necessarily imply poverty and low human capital, unlike the theory that stipulates the informal labour market is a marketplace where workers gain only subsistence earnings (Lewis, 1954; Harris and Todaro, 1970). However, the lower level of earnings of informal workers concerns workers with less human capital accumulation so there is a partial linkage between poverty and informality (Orlando, 2001; Cling et al., 2010). It is also a segmented market whose barriers to entry are the level of human capital and the financial capital of departure. The informal labour market includes rational economic agents with qualities that allow them to evolve into the formal labour market (Rosen, 1981, Heckman and Sedlack, 1985, Carneiro, Heckman, and Vytlačil, 2005). However, their position in the formal market is due to the search for comparative advantages. We have also observed that the informal economy remains what it is, that is to say, complex and difficult to apprehend completely from a theoretical point of view. For example, neo-Marxist theories appeared to be in conformity with the informal economy in Haiti, but the analysis of the recommendations and policy measures of this school of thought passed clearly through the Haitian informal debate. This analysis allowed us to go beyond theories and see how the road to modelling the informal economy was saddled and rocky, when it comes to apply a theory to the informality object. In view of the results obtained in the study on the various key factors of the informal economy, although the variables used did not reveal the formal ties between informal and informal, the results obtained provide answers Economic and political decision-makers so that they can act on the informal economy in order to regulate or curb it. The results of this study refer back to the three schools of thought of the informal sector, namely, dualism, structuralism and legalism. Insofar as each of these schools finds a favourable score in the study, but where this observation is crucial, none of them thought has come close to the goal. Analysing the economic policy recommendations and the measures stipulated by these theories to stop informality growth, none of them bring solution in the Haiti case. These results also warn



of any economic policy that claims to make the informal economy the driving force of an economic system, because it must be remembered that the informal economy exists through the weaknesses of the formal economy. For example, in the case of the Haitian informal sector, we have two glaring examples that reinforce this caveat: first, it has been observed that this informal sector has an independent size or almost the size of the money supply in circulation. Then, if we decide to make the informal sector the main pillar of the economy, the currency would no longer need to be regulated since it would no longer act on the global system. The second example always refers to the informal Haitian economy, which evolves in the same direction as inflation between 2007 and 2010, so if informality becomes the motor of the Haitian economy, what will happen in Haiti? The magnitude of the informal economy in Latin America and the Caribbean is now a delicate situation for states and governments, as the necessary solution seems to go beyond that. Should we fight the informal as theories recommend, it would be double the cost of the consequences of the crisis, the tentacles of the informal through informal employment, are already too extensive in the region. Should the informal sector be legalized to include it in the economy as a whole? Initially, it would carry too many gangue and residues within it, which would have a strong effect on the structured economy. Hence, as soon as legalized the informal economy would never again be informal, it would have passed on the other side of the barrier, so it would become formal. Nevertheless, the end of the informal economy is not for tomorrow, it is up to each country to find the economic model that suits it best in order to improve its economic situation, since there can be no panacea model, Informal economy. Economic and financial authorities must act on tax variables such as taxes and macroeconomic variables such as inflation, and government spending if they wish to slow the spread of informality in the economy. Public authorities must send a strong signal of consolidation and fight against corruption since the index of the quality of public services and the level of corruption evolve in the same direction as informal activities. The state has to invest in education and training because poor informal workers are those with the lowest accumulation of human capital. By investing in their training, the state would facilitate their accumulation of human capital, which would increase their incomes and, at the same time, reduce the poverty rate. The informal economy has demonstrated its ability to absorb economic shocks as well as shocks to the labour market. The State should develop a framework for formal and informal cohabitation while weakening the capacity of the informal economy to grow in the global economy.

# Appendix A

## Calculation of the Estimates of the PLS Model

### A.1 Estimates of the coefficients

Cronbach's alpha

$$Var(\sum_{h=1}^p x_h) = p + \sum_{h \neq h'} cor(x_h, x_{h'})$$

$$\alpha' = \frac{\sum_{h \neq h'} cor(x_h, x_{h'})}{p + \sum_{h \neq h'} cor(x_h, x_{h'})}$$

$$\alpha = \frac{\sum_{h \neq h'} cor(x_h, x_{h'})}{p + \sum_{h \neq h'} cor(x_h, x_{h'})} \times \frac{P}{P-1}$$

$$\alpha = \frac{\sum_{h \neq h'} cov(x_h, x_{h'})}{Var(\sum_h x_h)} \times \frac{P}{P-1}$$

$$Var(\sum_{h=1}^p x_h) = Var(\sum_{h=1}^P (\pi_{h0} + \pi_h \xi + \varepsilon_h)) = (\sum_{h=1}^P \pi_h)^2 Var(\xi) + \sum_{h=1}^P Var(\varepsilon_h)$$

Dillon-Goldstein's  $\rho$

$$\rho = \frac{(\sum_{h=1}^P \pi_h)^2 Var(\xi)}{(\sum_{h=1}^P \pi_h)^2 Var(\xi) + \sum_{h=1}^P Var(\varepsilon_h)}$$

$$\hat{\rho} = \frac{[\sum_{h=1}^P cor(x_h, t_1)]^2}{[\sum_{h=1}^P cor(x_h, t_1)]^2 + \sum_{h=1}^P (1 - cor^2(x_h, t_1))}$$

The outer estimate  $y_j$  of the standardized LV ( $\xi_j - m_j$ ) (mean=0 and standard deviation=1) are given by

$$Y_j \propto \pm [\sum_h \omega_{jh}(x_{jh}, \bar{x}_{jh})]$$

$$Y_j = \sum_h \tilde{\omega}_{jh}(x_{jh}, \bar{x}_{jh})$$

We denote the coefficients  $\omega_{jh}$  and  $\tilde{\omega}_{jh}(x_{jh})$  the outer weights.

We estimate the mean  $m_j$  by

$$\hat{m}_j = \sum \tilde{\omega}_{jh}(x_{jh} \bar{x}_{jh})$$

and the LV  $\xi_j$  by

$$\hat{\xi}_j = \sum_h \tilde{\omega}_{jh} x_{jh} = y_j + \hat{m}_j$$

$$\hat{\xi}_j^* = \frac{\sum_h \tilde{\omega}_{jh} x_{jh}}{\sum_h \tilde{\omega}_{jh}}$$

$$\hat{\xi}_{ij}^{0-100} = 100 \times \frac{(\hat{\xi}_{ij}^* - x_{min})}{(x_{max} - x_{min})}$$

The inner estimates  $z_j$  of the standardized  $LV(\xi_j - m_j)$  are given by

$$Z_j \propto \sum_{j': \xi_j' \text{ is connected with } \xi_j} e_{jj'} y_j'$$

In the simple regression the estimation for the weights  $\omega_{jh}$  is given by

$$\omega_{jh} = cov(x_{jh}, z_j)$$

In the multiple regression the estimation for the vector  $\omega_j$  of the weights  $\omega_{jh}$  is given by

$$\omega_j = (X_j' X_j)^{-1} X_j' z_j$$

$$sign(\omega_{jh}) = sign(cor(x_{jh}, \xi_j))$$

$$\omega_{ij} = \begin{cases} cov(x_{jh}, z_j) & \text{if } sign[cov(x_{jh}, z_j)] = sign[cor(x_{jh}, \xi_j)] \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.1})$$

When, in absolute value, the weights are all equal and reflect the signs of the correlation between the variables (MVs and their LVs), we have

$$\omega_{jh} = sign(cor(x_{jh}, z_j))$$

$$y_{ji} = \sum_{jh: x_{jhi} \text{ exists}} \tilde{\omega}_{jh} (x_{jhi} - \bar{x}_{jh})$$

$$z_{ji} = \sum_{k: \xi_k \text{ is connected with } \xi_j \text{ and } y_{ki} \text{ exists}} e_{jk} y_{ki}$$

$$\omega_j = (X_j' X_j)^{-1} X_j' z_j$$

$$\omega_j = [Var(X_j)]^{-1} cov(X_j, z_j)$$

$$communality_j = \frac{1}{p_j} \sum_{h=1}^{p_j} cor^2(x_{jh}, y_j)$$

$$communality = \frac{1}{p_j} \sum_{h=1}^j p_j communality_{j'}$$

$$\text{redundancy}_j = \text{communality}_j \times R^2(y'_j, \{y_{j'}' \text{ s explaining } y_i\})$$

$$\text{GoF} = \sqrt{\text{communality} \times \overline{R^2}}$$

$$\text{communality}_j = \frac{1}{p_j} \sum_{h=1}^{p_j} \left(1 - \frac{\sum_{i=1}^n (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh} y_{ji})^2}{\sum_{i=1}^n (x_{jhi} - \bar{x}_{jh})^2}\right)$$

$$\text{communality}_j = 1 - \frac{\sum_{h=1}^{p_j} \sum_{i=1}^n (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh} y_{ji})^2}{\sum_{h=1}^{p_j} \sum_{i=1}^n (x_{jhi} - \bar{x}_{jh})^2}$$

$$\begin{aligned} \text{redundancy}_{jh} &= \frac{\hat{\pi}_{jh}^2 \times \text{var}(y_j)}{\text{var}(x_{jh})} \times \frac{[\text{Pred}(y_j)]}{\text{var}(y_j)} \\ &= \frac{\hat{\pi}_{jh}^2 \times \text{var}[\text{Pred}(y_j)]}{\text{var}(x_{jh})} \\ &= \frac{\text{var}[\hat{\pi}_{jh} \times \text{Pred}(y_j)]}{\text{var}(x_{jh})} \end{aligned}$$

Where  $\text{Pred}(y_j)$  is the prediction of the LV  $y_i$  from the structural model, that is

$$\text{Pred}(y_j) = \sum_{j': \xi_{j'} \text{ explaining } \xi_j} \hat{\beta}_{j'} y_{j'}$$

If we assume that the regression coefficient of  $\text{Pred}(y_j)$  in the regression of  $(x_{jhi} - \bar{x}_{jh})$  on  $\text{Pred}(y_j)$  is close to  $\hat{\pi}_{jh}$ , then we get

$$\text{redundancy}_{jh} = 1 - \frac{\sum_{i=1}^n (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh} \times \text{Pred}(y_{ji}))^2}{\sum_{i=1}^n (x_{jhi} - \bar{x}_{jh})^2}$$

And then, if we assume that the variances of the variables  $x_{jh}$  are close to each other:

$$\text{redundancy}_{jh} = 1 - \frac{\sum_{h=1}^{p_j} \sum_{i=1}^n (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh} \times \text{Pred}(y_{ji}))^2}{\sum_{h=1}^{p_j} \sum_{i=1}^n (x_{jhi} - \bar{x}_{jh})^2}$$

The quality of each structural equation is measured by the cv-redundancy index (i.e. Stone-Geisser's  $Q^2$ ). It is a kind of cross-validated  $R^2$  between the MVs of an endogenous LV and all the MVs associated with the LVs explaining the endogenous LV, using the estimated structural model.

- Sum of squares of observations for one MV:  $SSO_{jh} = \sum_i (x_{jhi} - \bar{x}_{jh})^2$
- Sum of squared prediction errors for one MV:  $SSE_{jh} = \sum_i (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh}(-i)y_j(-i))^2$
- Sum of squares of observations for block j:  $SSO_j = \sum_h SSO_{jh}$
- Sum of squared prediction errors for block j:  $SSE_j = \sum_h SSE_{jh}$

- Cross validated-communality measure for block J:  $H_j^2 = 1 - \frac{SSE_j}{SSO_j}$

$$Pred(x_{jhi} - \bar{x}_{jh}) = \hat{\pi}_{jh}(-i)Pred(y_j(-i))$$

- Sum of squared prediction errors for one MV:

$$SSE'_{jh} = \sum_i (x_{jhi} - \bar{x}_{jh} - \hat{\pi}_{jh} \times Pred(y_{ji}))^2$$

- Sum of squared prediction errors for block j:

$$SSE'_j = \sum_h SSE'_{jh}$$

- Cross validated-redundancy measure for an endogenous block J:

$$F_j^2 = 1 - \frac{SSE'_j}{SSO_j}$$

## A.2 Tables of PLS models

TABLE A.1: Correlations between variables blocks and components factors

Block1	F1	F2	F3	F4
TMo	0.980	-0.057	-0.181	-0.053
logPIB	0.820	-0.488	0.299	-0.001
Inf	0.432	0.887	0.165	-0.009
Tchon	0.986	0.074	-0.141	0.057

Source: Computations by the author

TABLE A.2: Correlations between variables blocks and components factors

Block2	F1	F2	F3	F4	F5
DG	0.909	-0.163	0.383	0.000	0.001
Tax	0.899	-0.241	-0.365	-0.025	0.000
CC	0.675	-0.024	-0.247	0.689	-0.090
IG	0.486	0.614	-0.157	0.350	0.490
CPIB	0.391	0.920	-0.026	-0.010	-0.006

Source: Computations by the author

TABLE A.3: Correlations between variables blocks and components factors

Block 3	F1	F2
M1/M2	0.964	0.265
logMo	0.964	-0.265

Source: Computations by the author

## Appendix B

# Mathematical Validity of Bootstrapping

The bootstrap method in the one-sample case is mathematically conceived as follows. A random sample of size  $n$  is considered but the probability distribution of that sample is not known. We denote  $X = (X_1, X_2, \dots, X_n)$  the random sample,  $x = (x_1, x_2, \dots, x_n)$  its observed realization and  $L$  the unspecified probability distribution of  $X_i, i = 1, 2, \dots, n$ . The main purpose of the bootstrap method is to estimate the sampling distribution of a specified random variable  $R(X, L)$  which probably depending on both random sample  $X$  and the unspecified distribution  $L$  on the basis of observed data  $X$ . Consider  $\theta(L)$  as some parameter of interest of  $L$  (mean, standard deviation or correlation) and  $t(X)$  an estimator of  $\theta(L)$ , such as the sample mean, the sample correlation or a multiple of the sample range, two particular choices of  $R$  are specified in traditional jackknife theory. The first one is the estimation of the mean and variance of the sampling distribution of:

$$R(X, L) = t(X) - \theta(L) \quad (\text{B.1})$$

are calculated by the standard jackknife theory. Recomputing  $t()$   $n$  times, removing one component of  $X$  each time, give the possibility to obtain the bias and the variance estimates called  $\hat{\beta}(t)$  and  $\widehat{Var}(t)$  which are functions of  $X$ . The second choice of  $R$  concerns the following formulation:

$$R(X, L) = \frac{t(X) - \hat{\beta}(t) - \theta(L)}{(\widehat{Var}(t))^{\frac{1}{2}}} \quad (\text{B.2})$$

But none of the random variables of the traditional choices play a significant role in the bootstrap theory. There are three stages in the bootstrap process for the one-sample problem (Efron, 1979). The first stage is about constructing the sample probability distribution  $\hat{L}$ , at each point  $x_1, x_2, \dots, x_n$  we put mass  $\frac{1}{n}$ . The second step refers to draw a random sample of size  $n$  from  $\hat{L}$ , with the fixation of  $\hat{L}$  that is mean  $X_i^* = x_i^* \sim \hat{L}, i = 1, \dots, n$ . We denote  $X^* = (X_1^*, X_2^*, \dots, X_n^*)$  the bootstrap sample and  $x^* = (x_1^*, x_2^*, \dots, x_n^*)$  its realization. The

last point is approximate the sampling distribution of  $R(X, L)$  by the bootstrap distribution of  $R^*(X^*, \hat{L})$  which means the distribution of  $R^*$  is induced by the random mechanism with  $\hat{L}$  held fixed at its observed value. What makes bootstrap difficulties is the calculation of the bootstrap distribution, but three methods are recognized to solve that problem. The three methods are the direct theoretical calculation method, the Monte Carlo approximation, and the Taylor series expansion methods.

What makes bootstrap difficulties is the calculation of the bootstrap distribution, but three methods are recognized to solve that problem. The three methods are the direct theoretical calculation method, the Monte Carlo approximation, and the Taylor series expansion methods. Effron who developed the bootstrapping process, supposes that data are from an unknown distribution, where  $X_1, X_2, \dots, X_n \sim F$  with replicates of  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$  has an expected mean:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $\bar{x}$  is the approximation of the expected mean of the sample, that we denote  $E_L[X]$ . Answering the question how accurate is  $\bar{x}$  as an estimator of the true mean  $\theta = E_L[X]$ . The traditional measure of  $\bar{x}$ 's accuracy is the standard error. If we denote  $\mu_2(L) = E_L[X^2] - (E_L[X])^2$  the second central moment of  $L$ , then for a sample of size  $n$  from distribution  $L$  the standard error is:

$$\sigma(L) = \left( \frac{\mu_2(L)}{n} \right)^{\frac{1}{2}} \quad (\text{B.3})$$

The fact that we ignore the distribution  $L$  implies that we also ignore  $\mu_2(L)$ . To measure the  $\bar{x}$ 's accuracy, since we ignore  $\mu_2(L)$ , it is possible to use the estimated standard error:

$$\hat{\sigma}(L) = \left( \frac{\hat{\mu}_2}{n} \right)^{\frac{1}{2}} \quad (\text{B.4})$$

Where  $\hat{\mu}_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}$  the unbiased estimate of  $\mu_2(L)$ . But  $\mu_2(L)$  can be estimated more easily, if we denote  $\hat{L}$  the empirical probability distribution.

$\hat{L}$ : at each point  $x_1, x_2, \dots, x_n$  we put mass  $\frac{1}{n}$ . We can replace  $L$  by  $\hat{L}$  in (1.2) we obtain:

$$\hat{\sigma} = \sigma(\hat{L}) = \left[ \frac{\mu_2(\hat{L})}{n} \right]^{\frac{1}{2}} \quad (\text{B.5})$$

This is the estimated standard error for  $\bar{x}$ , then the bootstrap estimate. Returning to the one-sample situation, define  $P_i^* = \frac{N_i^*}{n}$ , where  $N_i^* = \#\{X_i^* = x_i\}$  as at (3.2), and the corresponding vector  $P^* = (P_1^*, P_2^*, \dots, P_n^*)$

By the properties of the multinomial distribution,  $P^*$  has mean vector and covariance matrix

$$E_* P^* = \frac{e}{n}$$

$$Cov_* P^* = \frac{I}{n^2} - \frac{e'e}{n^3}$$

Under the bootstrap sampling procedure, where  $I$  is the identify matrix and  $e = (1, \dots, 1)$ .

Given the observed data vector  $X = x$ , and therefore  $\hat{F}$ , we can use the abbreviated notation

$$R(P^*) = R(X^*, \hat{L})$$

For the bootstrap realization of  $R$  corresponding to  $P^*$ . In making this definition we assume that the random variable of interest,  $R(X, L)$ , is symmetrically defined in the sense that its value is invariant under any permutation of the components of  $X$ , so that it is sufficient to know  $N^* = nP^*$  in order to evaluate  $R(X^*, \hat{L})$ . This is always the case in standard applications of the jackknife.

The bootstrap distribution of  $R(X^*, \hat{L})$  can be approximated by expanding  $R(P^*)$  in a Taylor series about the value  $P^* = \frac{e}{n}$ , we find:

$$R(P^*) = R\left(\frac{e}{n}\right) + \left(P^* - \frac{e}{n}\right)U + \frac{1}{2}\left(P^* - \frac{e}{n}\right)V\left(P^* - \frac{e}{n}\right)' \quad (\text{B.6})$$

here

$$U = \left[ \begin{array}{c} \vdots \\ \frac{\partial R(P^*)}{\partial P_i^*} \\ \vdots \end{array} \right]_{P^* = \frac{e}{n}} \quad (\text{B.7})$$

$$V = \left[ \begin{array}{ccc} \cdots & \cdots & \cdots \\ \cdots & \frac{\partial^2 R(P^*)}{\partial P_i^* \partial P_j^*} & \cdots \\ \cdots & \cdots & \cdots \end{array} \right]_{P^* = \frac{e}{n}} \quad (\text{B.8})$$

Expansion (5.4), and definitions (5.5), assume that the definition of  $R(P^*)$  can be smoothly interpolated between the lattice point values originally contemplated for  $P^*$ . How to do so will be obvious in most specific cases, but a general recipe is difficult to provide.





# Appendix C

## Mincer model and the theorem of separation

### C.1 Theoretical Aspect

Regarding the theorem of separation, the optimal schooling decision is necessary a solution to the following maximization problem:

$$\max_s \int_S^{+\infty} e^{-(r+d)t} \omega(t) h(t) dt \quad (\text{C.1})$$

Assuming that individuals have finite expected lives so finite planning horizons we denote  $d$  the rate of death of workers. Initially Mincer supposed that the individuals should spend an interval  $S$  in full-time schooling, which means  $s(t) = 1$  and  $s(t) = 0$  thereafter. We take that assumption out in our model, because we assume that people can be illiterate ( $s(t) = 0$ ) or fully educated ( $s(t) = 1$ ) so  $s(t) \in S(t) \subset [0, 1]$ , therefore the workers will have a schooling level of:

$$h(S) = \zeta(S) \quad (\text{C.2})$$

where  $\zeta()$  is an increasing, continuously differentiable and concave function. If the individual works, the growth of his human capital stock over time, for some  $g_h \geq 0$ , is given by the following differential equation:

$$\dot{h}(S) = g_h h(t) \quad (\text{C.3})$$

This equation can be solved as follows:

$$\frac{\partial h(t)}{\partial t} = g_h h(t) \Leftrightarrow \frac{1}{h(t)} \frac{\partial h(t)}{\partial t} = g_h \Leftrightarrow \frac{1}{h(t)} \frac{\partial h(t)}{\partial t}$$

$$\frac{\partial \ln h(t)}{\partial t} = g_h \Leftrightarrow \int \frac{\partial \ln h(t)}{\partial t} = g_h \int dt$$

$$\ln h(t) = g_h \int dt$$

For the specific solution, we have:

$$\begin{aligned} \ln h(t) &= K + g_h \int_0^t dx \\ h(t) &= e^{g_h t} + K \end{aligned}$$

The growth of the wage over time, for some  $g_w \geq 0$ , is given by the following differential equation:

$$\dot{w}() = e^{g_w t} + C \tag{C.4}$$

There's a boundary condition  $w(0) > 0$   
 It is supposed that  $g_w + g_h < r + d$

Using equation C.3 and C.4 the maximization problem (equation C.1) is equivalent to:

$$\max_s \frac{\zeta(S)\omega_0 e^{-(r+d-g_w)S}}{r+d-g_w-g_h} \tag{C.5}$$

Since  $\zeta(S)$  is concave, the function in equation C.5 is strictly concave. Therefore, the maximization problem admits a unique solution which is characterized by the first-order condition:

$$\begin{aligned} \frac{\zeta'(S^*)}{\zeta(S^*)} = r + d - g_w &\Leftrightarrow \frac{\partial \zeta(S^*)}{\partial S^*} \frac{1}{\zeta(S^*)} = r + d - g_w \\ \Leftrightarrow \ln \zeta(S^*) &= \text{constant} + (r + d - g_w)S^* \end{aligned}$$

In the labour market the wage earnings of the worker of age greater than  $S^*$  is determined by:

$$Y(S, t) = e^{g_w t} e^{g_h(t-S)} \zeta(S) \Leftrightarrow \ln Y(S^*, t) = \text{constant} + (r + d)S^* + (g_w + g_h)(t - S^*)$$

That equation is equivalent to the equation (3.36) in chapter 3, with:

$$\beta_0 = \text{constant} \quad \beta_1 = r + d \quad \beta_2 = g_w + g_h$$

The equation representing the earnings of the worker can be written as follows:

$$\ln Y(S^*, t) = \text{constant} + \beta_1 S^* + \beta_2 (t - S)$$

This is the standard Mincer equation where  $t - S$  is considered as the experience accumulation of the worker after schooling time. In the equation 3.20, when a cross-sectional comparison across workers is made, the term  $[r_t k_0 + \frac{K_0}{T} + \frac{(k_0)^2}{T}]t$  turn into a constant, so that: is identical to the canonical Mincer equation also, where:  $[\ln E(0) - k_0 - \frac{1}{2}(k_0)^2] + rS + [r_t k_0 + \frac{K_0}{T} + \frac{(k_0)^2}{T}]t = \text{constant} = \beta_0 \quad \beta_1 = r \quad \beta_2 = -\frac{1}{2}(\frac{k_0}{T})^2 - \frac{r_t k_0}{2T}$

Hence:

$$\ln Y(S, t) = \text{constant} + \beta_1 S + \beta_2 t^2 \tag{C.6}$$

The separation theorem has two important roles for the Mincer model. First it ensures the concavity to the experience term of the equation as Mincer assumed it, and this theorem showed that the

Mincer model can consider human-capital decisions separately from consumption decisions. By the separation theorem, Mincer model contains the fact that schooling decisions maximizes the net present discounted value of people, therefore when an individual invests in human capital by the way he maximizes the present value of his earnings.

## C.2 Ordinal multinomial logit estimates of the Mincer wages model by category of workers

TABLE C.1: Estimates of the Mincer wages model by category of workers for other variables

Explanatory Variables	Variables levels	formal salaried		informal salaried		independent informal	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Intercept</b> Ref: [0, 1265]	(1265, 2485]	1.5185* (0.7287)	4.5656	0.2676 (0.345)	1.3068	1.5062*** (0.2168)	4.5095
	(2485, 5000]	-0.4995 (0.7237)	0.6068	- 1.3382*** (0.3448)	0.2623	0.2519 (0.2155)	1.2865
	(5000, 10000]	-2.577*** (0.7305)	0.076	- 3.0862*** (0.3528)	0.0456	- 1.2501*** (0.2164)	0.2864
	(10000, 30000]	- 4.2843*** (0.7436)	0.0137	- 4.5195*** (0.3629)	0.0108	- 2.4766*** (0.2215)	0.084
	(30000, 65000]	- 6.5751*** (0.7826)	0.0014	- 7.5029*** (0.4583)	0.0005	- 4.7661*** (0.2786)	0.0085
<b>Household size</b> Ref: [1, 3]	(3,6]	0.1083 (0.1998)	1.1144	-0.0941 (0.1217)	0.9102	0.0624 (0.0755)	1.0644
	(6,10]	-0.343 (0.2306)	0.7096	-0.2908* (0.144)	0.7476	0.0859 (0.0902)	1.0897
	(10,29]	-0.3025 (0.5481)	0.7389	-0.5299* (0.2842)	0.5886	-0.0073 (0.1822)	0.9926
<b>Workplace</b> Ref: premise	<i>Without premise</i>	-1.845 (0.2745)	0.158	-0.1614 (0.1515)	0.8509	- 0.5239*** (0.069)	0.5921

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%;

\*\*= 1%; \*= 5%; "=10%.

TABLE C.2: Multinomial logit estimates of the extended Mincer wage model by category of workers - Global fitting model

Model	Log likelihood	Chi-square	p-value
<b>formal salaried</b>	-716.641	16.61	.0000***
<b>informal salaried</b>	-1929.13	12.186	.00000***
<b>independent informal</b>	-5174.591	5.457	.000000***

Source: author's calculations using the data of the survey EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; "=10%.

TABLE C.3: Estimates of the Mincer wages model of poor workers for other variables

Explanatory Variables	Variables levels	formal poor workers		informal poor workers	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Marital status</b> Ref: Free union	<i>Married</i>	-1.0021* (0.4938)	0.3671	0.0644 (0.1115)	1.0665
	<i>Separated</i>	0.6423 (0.6224)	1.9008	0.0635 (0.1572)	1.0656
	<i>Single</i>	0.1643 (0.4550)	1.1785	-0.0709 (0.1367)	0.9315
<b>Industry</b> Ref: Manufacturing	<i>Building and Construction</i>	—	—	0.2560 (0.2734)	1.2917
	<i>Trade</i>	—	—	-0.1223 (0.1447)	0.8848
	<i>Services</i>	-0.8938 (1.2046)	0.4091	-0.0775 (0.1532)	0.9254
	<i>Others</i>	-0.6602 (1.7459)	0.5167	0.1417 (0.108)	1.1523
<b>Household size</b> Ref: [1, 3]	<i>(3,6]</i>	-0.2635 (0.4017)	0.7683	0.0724 (0.1067)	1.0751
	<i>(6,10]</i>	-0.7404* (0.4438)	0.4769	0.0511 (0.1256)	1.0524
	<i>(10,29]</i>	0.292 (1.0499)	1.339	-0.0164 (0.2383)	0.9837
<b>Workplace</b> Ref: premise	<i>employer's local</i>	-35.7504 (2850.211)	0.0000	-0.0871 (0.1049)	0.9165
	<i>Without premise</i>	-0.5037 (0.4865)	0.6042	—	—

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; "=10%.

TABLE C.4: Estimates of the Mincer wages model of non-poor workers for other variables

Explanatory Variables	Variables levels	formal poor workers		informal poor workers	
		$\beta$	$exp(\beta)$	$\beta$	$exp(\beta)$
<b>Marital status</b> Ref: Free union	<i>Married</i>	0.7899* (0.38782)	2.2032	0.2463* (0.1343)	1.2792
	<i>Separated</i>	-0.045 (0.6825)	0.9559	0.0543 (0.2139)	1.0558
	<i>Single</i>	0.4776 (0.4531)	1.6122	0.0359 (0.1745)	1.0366
<b>Industry</b> Ref: Manufacturing	<i>Building and Construction</i>	1.6889 (1.3751)	5.414	0.5118* (0.2417)	1.6683
	<i>Trade</i>	1.7431 (2.0132)	5.7152	-0.0442 (0.1799)	0.9567
	<i>Services</i>	-0.3469 (0.6731)	0.7068	0.1458 (0.1763)	1.1570
	<i>Others</i>	0.6402 (0.9227)	1.8968	-0.0098 (0.4312)	0.9902
<b>Household size</b> Ref: [1, 3]	<i>(3,6]</i>	-0.202 (0.3032)	0.817	-0.2402* (0.1303)	0.7865
	<i>(6,10]</i>	-0.6106 (0.3648)	0.5429	-0.2918* (0.1573)	0.7468
	<i>(10,29]</i>	0.5097 (0.9143)	0.601	-1.0994* (0.4718)	0.3331
<b>Workplace</b> Ref: premise	<i>without premise</i>	0.3993 (0.6568)	0.6707	-0.0872 (0.1334)	0.9164

Source: author's calculations from EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; "=10%.

TABLE C.5: Multinomial logit estimates of the extended Mincer wage model of non-poor workers - Global fitting model

Model	Log likelihood	Chi-square	p-value
<b>formal non-poor</b>	-247.8725	2.8310	.000000***
<b>informal non-poor</b>	-1254.637	5.7201	.000000***

Source: author's calculations using the data of the survey EEEI 2007; significance level codes: \*\*\*= 0.1%; \*\*= 1%; \*= 5%; "=10%.



## Appendix D

# Markov Chain Monte Carlo in the Dirichlet Process

### D.1 Describing of the Markov Chain process into the clusters

Given all these considerations that we stated earlier, the unobserved data will be treated as a realization of a stochastic process  $\{\theta_t\}_{t=1}^T$ . In this process the marginal distribution and the dependence structure are subject to a hierarchical and semi-parametric specification (Yau et al., 2011). Let  $\{s_t\}_{t=1}^T$  be a Markov chain with discrete state space  $S = 1, \dots, K$  with corresponding observation sequence  $y = (y_1, \dots, y_T)$ . Transitions between the hidden states induced by Markov dynamics are represented by the transition matrix  $q = [q_{i,j}]_{i,j \in S}$  where  $q(i, j) = p(s_t = j / s_{t-1} = i)$ , for  $i, j \in S$  and  $q_{0,i} = p(s_1 = i)$  represent initial state probabilities, and the first DPM becomes. For each state  $s_t$  it exists a parameter  $\theta_{s_t} \sim^{iid} H$ , which controls the observation likelihood for that state:  $y_t / s_t \sim F(\theta_{s_t})$ , where  $H$  is a prior distribution. The conditional distribution of  $y_t$  given  $s_t$  is called emission distribution and the latent stochastic process  $k_t$  determines which component of the emission mixture distribution  $y_t$  should be allocated to. Consider the HMM parameters  $q_0, q, \theta, K$ , in the hidden states  $s$  and the observations  $y$ , with  $s_0 = 0$  the joint distribution is given by:

$$p(s, y / q_0, q, \theta, K) = \prod_{t=1}^T p(s_t / s_{t-1}) p(y_t / s_t) \quad (\text{D.1})$$

Without other prior information on the state sequence, the prior for the transition probabilities are symmetric Dirichlet distributions. The introduction of coupling across transitions can be realized through the Bayesian formalism by sharing the parameters among the Dirichlet priors with a higher level prior. If we denote  $q_k \sim \text{Dirichlet}(\alpha\beta)$  the transition probabilities out of state  $k$  and  $\beta \sim \text{Dirichlet}(\gamma/K, \dots, \gamma/K)$  the parameters vector of the shared prior, as  $K \rightarrow \infty$ , the above hierarchical prior approaches a hierarchical Dirichlet process (Teh et al., 2006).



The shared base measure is given by:  $G_0 \sim (\gamma, H)$  where  $H$  represent a global base measure. The stick breaking construction for HDP expresses the random measures as follows:  $G_0 = \text{sum}_{k'=1}^+ \infty \nu_{k'} \delta_{\theta_{k'}}$  and  $G_k = \text{sum}_{k'=1}^+ \infty q_{kk'} \delta_{\theta_{k'}}$  where  $\beta \sim \text{Beta}(1, \gamma)$  is the stick breaking construction for DP,  $q_k \sim \text{DP}(\alpha, \beta)$  and  $\theta'_k \sim \text{iid } H$ . Either the transition probabilities  $q(kk')$  from state  $k$  to  $k'$  or the emission distributions parameterized by  $\theta_{k'}$  are described by each  $G_k$ ,  $s_t/s_{t-1} \sim \text{Multinomial}(q_{s_{t-1}})$  that permits to define globally the infinite HMM.

## D.2 Describing of the Markov Chain Monte Carlo in the Dirichlet Process

Through the Gibbs sampling process above each  $\theta_{kr}$  is updated. During the hierarchical Dirichlet process, the elements of the data set are supposed to be independent at each time of the process and therefore structural changes over time imply a sequence of hidden state with the Markov dynamics. This hidden state sequence defines the Hidden Markov Model (HMM), hence sampling for local clustering is induced by the HMM.

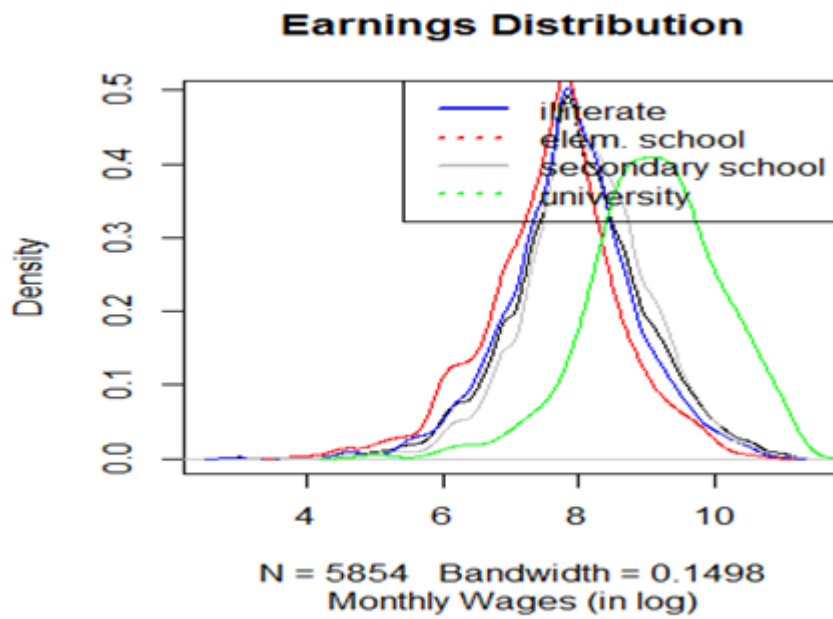
In our model the parameter vector  $\theta$  and the auxiliary variable  $u$  are the unobservable and the observed data are  $x = (x_{i1}, \dots, x_{im})$ . Therefore the solution of the inference problem is the posterior distribution  $p(\theta, u/x)$ . Nevertheless, it is uneasy to characterize the posterior distribution because  $p(\theta, u/x)$  is complex and high-dimensional; standard sampling methods cannot be applied. But Markov Chain Monte Carlo methods propose a way for sampling from such high-dimensional model, which is a simulation based method. We are able to sample draws slightly dependent from  $p(\theta/x)$ , let  $\Theta$  be the space parameter. Consecutive draws lead to a time component indexed by  $t$ . Consider a draw of  $\theta^{(t)}$  to be a state at iteration  $t$ , the next draw  $\theta^{(t+1)}$  depends only on the current draw  $\theta^{(t)}$  but not on the other past draws. Such process satisfies the Markov property

$$p(\theta^{(t+1)}/\theta^{(t)}) = p(\theta^{(t+1)}/\theta^{(1)}, \dots, \theta^{(t)}) \quad (\text{D.2})$$

Therefore our Markov chain is constituted by a lot of draws of  $\theta$  that are each slightly dependent of their immediate past. The chain runs through the parameter space having in memory only its position during the last period. The manner that our Markov chain jumps from one state to another at each period is leaded by a transition kernel. A transition kernel refers to a mechanism that describes the probability of moving to some other state based on its current state. For a defined stationary distribution  $\pi$ , the Markov chain will converge to  $\pi$  regardless the starting points. Running the conceived chain, we can obtain draws that are approximately from our posterior distribution once the chain has converged. Since the convergence of the

chain is set, we can use Monte Carlo integration methods. Thanks to the Ergodic theorem, we can calculate quantities of interest from our draws neglecting the dependence between draws. Hence, we can consider MCMC as a class of methods allowing simulation of draws that are slightly dependent and are approximately from the posterior distribution.

FIGURE D.1: Earnings distribution according to educational level



Source: Author's design from EEEI2007 data.



# Bibliography

- [1] ACCION (2003). Action poverty Outreach Findings, SOGESOL, Haiti. *Insight*, N<sup>o</sup>8.
- [2] Acemoglu, D., and Autor, D. (2011). *Lectures in Labor Economics, manuscript*. MIT.
- [3] Adair, P. (2002). L'emploi informel en Algérie: évolution et segmentation du marché du travail. *Cahiers du GRATICE*, 22, 95-126.
- [4] Adair, P. (2009). Économie non observée et emploi informel dans les pays de l'Union Européenne : une comparaison des estimations et des déterminants. *Revue Economique*, 60 (5), 1117-1153.
- [5] Adair, P., and Ballache, Y. (2012). Emploi et secteur informels en Algérie : Déterminants, Segmentation et Mobilité de la Main-d'œuvre. *Région et Développement*, 35, 121-49.
- [6] Adam, M. C., and Ginsburgh, V. (1985). The effects of irregular markets on. *European Economic Review*, 29 (1): 15-33.
- [7] Agarwal, M. K., and Dhakal, R. (2010). Impact of informal sector on poverty and employment in Nepal – Micro-level study of Chitwan District". *International conference, "The informal sector and Informal Employment: Statistical Measurements, economic implications and public policies*. Hanoi Vietnam.
- [8] Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In *Petrov, B., N., Csaki, F. 2nd International Symposium on Information Theory*, Tsahkadzor, Armenia, USSR, pp. 267-281. Budapest: Akademiai Kiado.
- [9] Allingham, G., and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1, 323-338.
- [10] Amaral, P. S., and Quintin, E. (2006). A competitive model of the informal sector. *Journal of Monetary Economics*, 53 (7), 1541-1553.
- [11] Antoniak, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems . *The Annals of Statistics*, 2, 1152–1174.

- [12] Aspilaire, R. (2012). L'économie informelle en Haïti de 1990 à 2010 : une étude à partir des modèles EMIMIC et SCVAR. *Non-published master dissertation*. Université des Antilles et de la Guyane.
- [13] Aspilaire, R. (2014). Economie informelle en Haïti : Un impact contracyclique sur le PIB ? . *Mondes en Développement* 166 (2), 101-12.
- [14] Bajada, C., and Schneider, F. (2007). Size, Causes and Consequences of the Underground Economy: An International Perspective. Aldershot (GB). *Ashgate Publishing Company*.
- [15] The role of the informal sector in the migration process: a test of probabilistic migration models and labour market segmentation for India. *Oxford Economic Papers*, 35, 399–422.
- [16] Banfield, J. D., and Raftery, A. (1993). Model-based Gaussian and non-Gaussian clustering., . *Biometrics*, 49, 803–21.
- [17] World Bank. (2005). *Ukraine jobs study. Fostering productivity and job creation*. Report No. 32721-UA, Human Development Sector Unit, Volume I.
- [18] Barro, R. J., and Sala-i-Martin, X. (2004). *Economic Growth*. 2th ed. Cambridge, MA: MIT Press.
- [19] Basu, A. (1997). *Analytical Development Economics*. MIT Press, Cambridge, MA.
- [20] Becker, G. S. (1962). Investment in Human Capital: A Theoretic Analysis. *Journal of Political Economy*, 70(5): 9-49.
- [21] Becker, G. S. (1964). *Human Capital*. New York: Columbia University Press, National Bureau of Economic Research.
- [22] Becker, G. S. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York: National Bureau of Economic Research.
- [23] Bellanger, L., and Tomassone, R. (2014). *Exploration de données et méthodes statistiques*. Ellipses.
- [24] Ben-Porath, Y. (1967). Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, 75(4): 352-65.
- [25] Bentler, P. M. (2004). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- [26] Bentler, P. M. (2003). *EQS (Version 6) [Computer software]*. Encino, CA: Multivariate Software.

- [27] Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin*, 107 (2): 238-46.
- [28] Bentler, P. M. (1983). Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika*, 48, 493-517.
- [29] Bentler, P. M., and Bonnet, D. (1980). Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychological Bulletin*, 88 (3): 588-606.
- [30] Biernacki, C., G. Celeux, and G. Govaert (2003). Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics and Data Analysis*, 41: 561-575
- [31] Blackwell, D. (1973). The discreteness of Ferguson selections. *The Annals of Statistics*, 1: 356-358.
- [32] Blackwell, D., and MacQueen, J. (1973). Ferguson distributions via Polya urn schemes. *The Annals of Statistics*, 1: 353-355.
- [33] Blinder, A., and Solow, R. (1974). Analytical Foundations of Fiscal Policy, in *The Economics of Public Finance*, (Washington, D.C: The Brookings Institution).
- [34] Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- [35] Botman, D., Laxton, D., Muir, D., and Romanov, A. (2006). A New-Open-Economy Macro Model for Fiscal Policy Evaluation, IMF Working Paper No 06/45, Washington, D.C.
- [36] Bowles, S., and Gintis, H. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.
- [37] Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3): 345-370.
- [38] BRH. (2012). *Indicateurs Macroéconomiques*. Retrieved from <http://www.brh.net/m3annuel.pdf>.
- [39] BRH. (2014). Taux de change. Retrieved from [www.brh.net/taux-de-change.html](http://www.brh.net/taux-de-change.html).
- [40] Bullard, J., and Singh, A. (2007). Worldwide macroeconomic stability and monetary policy rules. *Federal Reserve Bank of St. Louis Working paper* 2006-040.

- [41] Cagan, P. (1958). The Demand for Currency Relative to Total Money Supply. *Journal of Political Economics*, 1-37.
- [42] Cain, G. C. (1976). The Challenge of Segmented Labor Market Theories to Orthodox Theory. *Journal of Economic Literature*, 14, 1215-1257.
- [43] Carneiro, P., Heckman, J., and Vytlacil, E. (2005). *Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education*. Working Paper, Stanford University.
- [44] Cartaya, V. (1994). *Informality and Poverty: Causal Relationship or Coincidence?*. InnContrapunto. The Informal Sector Debate in Latin America. Cathy A. Rakowski ed. Albany: State University of New York Press, pp. 223-49.
- [45] Castells, M., and Portes, A. (1986). The origins, the dynamics, and effects on the informal economy, *World Underneath*. John Hopkins University Press.
- [46] Charmes, J. (1987). *Debat actuel sur le secteur informel*, Tiers Monde, tome 28.
- [47] Chen, M. A. (2005). *Rethinking the informal economy: Linkages with the Formal Economy and the Formal Regulatory Environment*. World Institute for Development Economic Research, 2005/10.
- [48] Chen, M. A., Vanek, J., and Carr, M. (2004). *Mainstreaming informal employment and gender in poverty reduction*. London: Commonwealth Secretariat.
- [49] Choi, J., and Thum, M. (2004). Corruption and the shadow economy. *International Economic Review*, 12(4), 308-342.
- [50] CIA. (2011). *World Factbook*. Retrieved from <http://www.indexmundi.com/g/g.aspx?c=ha&v=74&l=fr>.
- [51] . Cling, J. P., Razafindrakoto, M., and Roubaud, F. (2010). *The informal economy in Vietnam*. International Labour Organization.
- [52] CNSA. (2002). *Haïti : Insécurité alimentaire 2001-2002*. Port-au-Prince.
- [53] Crowley, S. L., and Fan, X. (1997). Structural Equation Modeling: Basic Concepts and Applications in Personality Assessment Research. *Journal of Personality Assessment*, 68 (3), 508-31.

- [54] Cunningham, W., and Maloney, W. (2001). Heterogeneity in the Mexican micro-enterprise sector: an application of factor and cluster analysis. *Economic Development and Cultural Change*, 50 (1), 131–156.
- [55] Dahl, D. B., and Newton, M. (2007). Multiple hypothesis testing by clustering treatment effects. *Journal of the American Statistical Association*, 102, 517–526.
- [56] De Soto, H. (1989). *The Other Path*. London: Harper and Row.
- [57] DeAlmeida, O., Luiza, A., Alves, L., and Graham, S. (1995). Poverty, Deregulation, and Employment in the Informal Sector of Mexico. Education and Social Policy Discussion Paper 54. World Bank, Washington, D.C.
- [58] Debrun, X., and Kapoor, R. (2010). Fiscal Policy and Macroeconomic Stability: Automatic Stabilizers Work, Always and Everywhere, IMF Working Paper No 10/111, Washington D.C.
- [59] Dell’Anno, R., and Schneider, F. (2006). *Estimating the Underground Economy by Using MIMIC Models: A Response to T. Breusch’s critique.* Working Paper, 0607, Johannes Kepler University of Linz.
- [60] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM-Algorithm. *Journal of the Royal Statistical Society, B*, 39, 1-38.
- [61] Despinas, J.-Y. (2008). *La pauvreté en Haïti : contextes historique et politiques d’ajustement structurel*. La Rebelle, p74.
- [62] Dickens, W. T., and Lang, K. (1985). A test of dual labour market theory. *American Economic Review*, 67 (2), 792–805.
- [63] Dougherty, S., and Escobar, O. (2013). *The Determinants of Informality in Mexico’s States*. OECD Economics Department Working Papers, No. 1043, OECD Publishing.
- [64] Dreher, A., C, K., and S, M. (2005). *Corruption around the World: Evidence from a Structural Model.* University of Konstanz and University of Exeter, mimeo.
- [65] ECLAC. (2012). *CEPALSTAT, Databases and Statistical Publications, Statistics and indicators*.
- [66] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1-26.



- [67] Efron, B., and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1 (1), 54-75
- [68] Egset, W., and Sletten, P. (2004). Poverty in Haiti. FAFO-paper, 31.
- [69] Escobar, M. D. (1994). Estimating normal means with a Dirichlet process prior. *Journal of the American Statistical Association*, 89, 268-77.
- [70] Escobar, M. D., and West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90, 577-88.
- [71] Fass, S. (1988). Political Economy in Haiti: The Drama of Survival. *Transaction Publishers, New Brunswick NJ*.
- [72] Feige, E. L. (1979). How Big is the Irregular Economy? . *Challenge*, 22(5), 5-13.
- [73] Feige, E. L. (1986). A reexamination of the underground economy in the United States: A comment on Tanzi,. *IMF Staff Papers, Palgrave Macmilan*, 33(4), 768-781.
- [74] Feige, E. L. (1990). Defining and estimating underground and informal economies: the new institutional economics approach. *World Development*, 18(7), 989-1002.
- [75] Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In H. Rizvi and J. Rustagi Ed. *Recent Advances in Statistics*, 87-302. New York: Academic Press.
- [76] Fields, G. S. (1990). Labour market modelling and the urban informal sector: theory and evidence. In: Thurnham, D., Salome, B., Schwarz, A. (Eds.), *The Informal Sector Revisited .OECD. Paris: Organisation for Economic Co-operation and Development*, 44-69.
- [77] Fields, G. S. (2005). A guide to multisector labour market models. *World Bank, Social Protection Discussion Paper Series, No. 0505*.
- [78] Gadner, H. (1983). *Frames of Mind*. New York: Basic Book Inc.
- [79] Gindling, T. (1991). Labor market segmentation and the determination of wages in the public, private-formal and informal sectors in San-Jose, Costa-Rica. *Economic Development and Cultural Change*, 39 (3), 585-603.

- [80] Goldberger, S. A. (1972). Structural equation method in the social science. *Econometrica* 40(6), 979-1001.
- [81] Gronau, R. (1973). The Effect of Children on the Housewife's Value of Time. *Journal of Political Economy*, 81, 168-199.
- [82] Grun, B., and Leisch, F. (2008). Identifiability of Finite Mixtures of Multinomial Logit Models with Varying and Fixed Effects. *Journal of Classification*, 97, 88-98.
- [83] Guadagnoli, E., and Velicer, W. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-275.
- [84] Gunther, I., and Launov, A. (2011). Informal Employment in Developing Countries. Opportunity or Last Resort. *Journal of Development Economics*, 97, 88-98.
- [85] Gutman, P. M. (1977). The Subterranean Economy. *Financial Analysts Journal*, 6 (33), 26-27, 34.
- [86] Hair, J. F., Black, W., Babin, B., Anderson, R., and Tatham, R. (2006). *Multivariate Data Analysis (6th ed.)*. Upper Saddle River, NJ: Prentice Education, Inc.
- [87] Harris, J. R., and Todaro, M. (1970). Migration, Unemployment and Development: A Two Sectors Analysis. *American Economic Review*, 60(1), 126-42.
- [88] Hayduk, L. A. (2006). A conceptually improved definition of the proportion of explained variance in models containing loops or correlated residuals. *Quality and Quantity*, 40, 629-649.
- [89] Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., and Boulianne, S. (2007). Testing! testing! one, two, three—Testing the theory in structural equation models! *Personality and Individual Differences*, 42, 841-850.
- [90] Heckman, J., and Hotz, V. (1986). An investigation of the labour market earnings of Panamanian males. *The Journal of Human Resources*, 21 (4), 507-542.
- [91] Heckman, J., and Scheinkman, J. (1987). The Importance of Bundling in a Gorman-Lancaster Model of Earnings. *Review of Economic Studies*, 54, 243-255.
- [92] Heckman, J., and Sedlacek, G. (1985). Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market. *Journal of Political Economy*, 93(6), 1077-125.

- [93] Hill, R., and Kabir, M. (1996). Tax Rates, the Tax Mix, and the Growth of the Underground Economy in Canada: What Can We Infer? . *Revue Fiscale Canadienne*, 44(6), 1552–1583.
- [94] Hindriks, J., Muthoo, A., and Keen, M. (1999). Corruption, extortion and evasion. *Journal of Public Economics*, 74, 395-430.
- [95] Hoyle, R. H. (1991). Evaluating Measurement Models in Clinical Research: Covariance Structure Analysis of Latent Variable Models of Self-conception. *Journal of Consulting and Clinical Psychology*, 59(1), 67-76.
- [96] Hsu, C.-N., Huang, H.-J., and Wong, T.-T. (2003). Implications of the Dirichlet Assumption for Discretization of Continuous Variables in Naive Bayesian Classifiers. *Machine Learning*, 53(3), 235-263.
- [97] Hudson, R. (2004). *Conceptualizing economies and their geographies: spaces, flows and circuits*, *Progress in Human Geography*, 28(4), 447-71.
- [98] Hugon, P. (1980). *Les petites activités marchandes dans les espaces urbains africains (essai de typologie)* . *Tiers-Monde*, tome 21(82), 405-426.
- [99] Hussmanns, R. (2004). *Measuring the informal economy: From employment in the informal sector to informal employment*. Geneva: International Labour Office, Working Paper No. 53.
- [100] IHSI. (2001a). *Enquête Budget-Consommation des Ménages 1999-2000 : Volume I*, Haïti. Port-au-Prince.
- [101] IHSI. (2001b).. *Enquête budget-consommation des ménages : revenus, consommation, transfert*, Vol 2. Haiti. Port-au-Prince.
- [102] IHSI. (2003). *Enquête sur les conditions de vie en Haïti. Les actes du Séminaire-Atelier de présentation des premiers résultats*. Port-au-Prince.
- [103] IHSI. (2005). *Recensement Général de la Population et de l'Habitat 2003, Haïti*.
- [104] IHSI. (2010). *Enquête sur l'Emploi et l'Economie Informelle 2007 : Premiers Résultats de l'Enquête Emploi, Haïti*.
- [105] IHSI, and CELADE. (2007). *Projections de population totale, urbaine, rurale et économiquement active, Haïti*.

- [106] ILO. (2003). *Guidelines concerning a statistical definition of informal employment, endorsed by the Seventeenth International Conference of Labour Statisticians (November-December 2003)*. In Seventeenth International Conference of Labour Statisticians (Geneva, 24 November - 3 December 2003), Report of the Conference Geneva, doc. ICLS/17/2003/R.
- [107] ILO. (2002). *Decent Work and the Informal Economy*, Report of the Director-General; International Labour Conference, 90th Session, Report VI. Geneva, ILO.
- [108] ILO. (1993). *Rapport pour la 15ème Conférence Internationale des Statisticiens du Travail, Bureau International du Travail, Genève*. Geneva.
- [109] ILO. (2002). *Travail décent et économie informelle, Conférence Internationale du Travail, 87ème session, Bureau International du Travail, Genève*. Geneva.
- [110] ILO. (2013). *Measuring informality: a statistical manual on the informal sector and the informal employment*. Geneva.
- [111] IMF. (2001). Haiti: Selected issues. IMF Staff Country Report No. 01/04. IMF Staff Country Report No. 02/18. Washington, D.C.
- [112] IMF. (2002). Haiti: Selected issues. IMF Staff Country Report No. 05/20. Washington, D.C.
- [113] IMF. (2005). Haiti: selected issues. *IMF Staff Country Report* No. 05/205, Washington, D.C.
- [114] Isachsen, A. J., and Strom, S. (1985). The Size and Growth of the Hidden Economy in Norway. *Review of Income and Wealth*, 31(1), pp. 21–38.
- [115] Ishwaran, H., and James, L. (2001). Gibbs Sampling Methods for Stick-Breaking Priors. *Journal of the American Statistical Association*, 96 (453), 161-173.
- [116] Jakobowicz, E. (2007). Contributions aux modèles structurels à variables latentes, PhD., Conservatoire National des Arts et Métiers. Paris.
- [117] Jensen, L., and Slack, T. (2009). *Informal work in rural America*. In: E. Marcelli and C. C. Williams (eds) *Informal Work in Developed Nations*. London: Routledge.
- [118] John, G., and Langley, P. (1995). *Estimating continuous distributions in Bayesian classifiers*. In *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence*, pp. 338–345. San Mateo: Morgan Kaufmann Publishers

- [119] Johnson, S., Kaufmann, D., and Zoido-Lobaton, P. (1998). Regulatory discretion and the unofficial economy. *The American Economic Review*, 88/ 2, pp. 387-392.
- [120] Joreskog, G. K., and Goldberger, S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 351(70), pp. 631-639.
- [121] Joreskog, K. G., and Sobom, D. (1993). *LISREL 8: Structural equation modeling with the Simplis command language*. Chicago: Scientific Software International.
- [122] Joreskog, K. G., and Sorbom, D. (1996). *LISREL 8: User's Reference Guides*. Chicago, IL: Scientific Software International Inc.
- [123] Kaufmann, D., and Kaliberda, A. (1996). Integrating the Unofficial Economy into the Dynamics of Post-Socialist Economies: A Framework of Analysis and Evidence . *World Bank Policy research Working Paper*, 1691.
- [124] Kenny, D. A., and Mc Coach, D. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. *Structural Equation Modeling*, 10(3), 333-51.
- [125] Kliksberg, B. (1994). *Latin American Development and Public Policy*. p. 317. The Macmillan Press Ltd, Great Britain.
- [126] Kline, R. (2005). *Principles and Practice of Structural Equation Modeling*. 2nd ed. The Guilford Press, New York.
- [127] Kline, R. (2011). *Principles and practice of Structural Equation Modeling*. 3rd ed. The Guilford Press, New York.
- [128] Kolev, A. (2000). Déterminants de l'Offre de Travail et Secteur Informel en Russie. *Revue d'Etudes Comparatives Est-Ouest*, 31(4) : 123-50.
- [129] Krusec, D. (2003). The effects of fiscal policy on output in a structural VEC model framework: The case of four EMU and four non-EMU OECD countries. *8th International Conference of Macroeconomics Analysis and International Finance*. University of Crete.
- [130] Lacko, M. (1999). Do Power Consumption Data Tell the Story? Electricity Intensity and Hidden Economy in Post-Socialist Countries. *Budapest Working Papers on the Labour Market*, No. BWP - 1999/2.
- [131] Lamaute-Brisson, N. (2002). *l'Economie informelle en Haïti: de la reproduction urbaine a Port-au-Prince*. Paris: l'Harmattan.

- [132] Lamaute-Brisson, N. (2005). *Emploi et Pauvreté en Milieu Urbain en Haïti*, LC/MEX/R.882.
- [133] Lenoard, R. E. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics*, 9 (2): 508–23.
- [134] Leonard, M. (1994). *Informal Economic Activity in Belfast*. Aldershot: Avebury.
- [135] Leonard, M. (1998). *Invisible Work, Invisible Workers: the informal economy in Europe and the US*. London: Macmillan.
- [136] Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *Manchester School of Economic and Social Studies*, 22 (2), pp. 139–191.
- [137] Lo, A. Y. (1984). On a class of Bayesian nonparametric estimates I. Density estimates. *The Annals of Statistics*, 12, 351–357.
- [138] Loi du 17 avril 2003, No 29. (2003). Le Moniteur.
- [139] Loi du 06 Octobre 2009, No 1099. (2003). Le Moniteur.
- [140] London, T., and Hart, S. (2004). Reinventing strategies for emerging markets: beyond the transnational model. *Journal of International Business Studies*, 35 (2), 350-70.
- [141] Lopez, C. H. (1989). Le secteur informel, substitut d'un système d'assurances sociales en Colombie. *Problèmes d'Amérique latine*, 92, pp. 113-129.
- [142] Lucas, R. E. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics*, 9 (2), 508–23.
- [143] MacCallum, R. C. (1986). Specification searches in covariance structure modelling. *Psychological Bulletin*, 100, 107–20.
- [144] MacCallum, R. C., and Austin, J. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201–236.
- [145] MacCallum, R. C., Browne, M. W., Cai, L. (2006). Testing Differences Between Nested Covariance Structure Models: Power Analysis and Null Hypotheses. *Psychological Methods*, 11(1), 19–35.
- [146] MacCallum, R. C., Roznowski, M., and Necowitz, L. (1992). Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychological Bulletin* 111, 490-504.

- [147] MacEachern, S. N. (1994). Estimating normal means with a conjugate style Dirichlet process prior. *Communications in Statistics*, B 23, 727–41.
- [148] MacEachern, S. N., and Muller, P. (1998). Estimating mixtures of Dirichlet process models. *Journal of Computational and Graphical Statistics*, 7, 223–38.
- [149] Magnac, T. (1991). Segmented or Competitive Labor Market . *Econometrica*, 59(1), 165-187.
- [150] Maldonado, C. (1995). Entre l'illusion de la normalisation et le laisser-faire: vers la légalisation du secteur informel? . *Revue Internationale du Travail*, 3 (134), pp. 777-801.
- [151] Maloney, F. W. (1999). Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *The World Bank Economic Review*, 13 (2), 275-302.
- [152] Maloney, F. W. (2004). Informality revisited. *World Development*, 32(7), 1159-1178.
- [153] Marsh, H. W., Balla, J., and McDonald, R. (1988). Goodness-of-fit indices in confirmatory factor analysis: Effects of sample size. *Psychological Bulletin*, 103, 391-411.
- [154] Mazumdar, D. (1983). Segmented Labor Markets in LDCs. *American Economic Review*, 73(2): 254–59.
- [155] McArdle, J. J., and McDonald, R. (1984). Some algebraic properties of the Reticular Action Model for moment structures. *British Journal of Mathematical and Statistical Psychology*, 37, 234–251.
- [156] McDonald, R. P., and Hartmann, W. (1992). A procedure for obtaining initial values of parameters in the RAM model. *Multivariate Behavioral Research*, 27:57-76.
- [157] McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. published in Zarembka P. (ed.), *Frontiers in Econometrics*, Academic Press, New York.
- [158] McLachlan, G. J., and Basford, K. (1988). *Mixture Models: Inference and Applications to Clustering*. New York: Dekker.
- [159] MEF. (2012). Situation économique, financière, sociale en Haïti en 2012 et perspectives à court terme.
- [160] Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4): 281–302.

- [161] Mincer, J. (1962). On-the-Job Training: Costs, Returns and Some Implications. *Journal of Political Economy*, 70(5): 50-79.
- [162] Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- [163] Mogensen, G. V., Kvist, H., Kfrmendi, E., and Pedersen, S. (1995). The shadow economy in Denmark 1994: Measurement and Results, Study no. 3. The Rockwool Foundation Research Unit; Copenhagen.
- [164] Montas, R. (1995). *Dossier Evolution de l'économie haïtienne : 1980-1994*. Port-au-Prince.
- [165] Montas, R. (2002). *Evolution Macroéconomique, Pauvreté et Inégalité en Haïti :1986-2001*. Port-au-Prince.
- [166] Montas, R. (2005). *La pauvreté en Haïti: Situation, causes et politiques de sortie*. LC/MEX/R.882.
- [167] Montas, R. (2012). L'économie haïtienne 1986-2011, Ouverture par la migration et stagnation. *Rencontre*, pp. 26-27, pp. 29-35.
- [168] MPCE. (2007). *Document de Stratégie Nationale pour la Croissance et la Réduction de la Pauvreté: Pour Réussir le Saut Qualitatif*. Presses de l'imprimerie Deschamps. Port-au-Prince.
- [169] Nelson, R., and Phelps, E. (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review*, 56(1): 69-75.
- [170] Nevitt, J., and Hancock, G. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modelling. *Structural Equation Modeling*, (8) 353-377.
- [171] Nwabuzor, A. (2005). Corruption and development: new initiatives in economic openness and Strengthened Rule of Law, *Journal of Business Ethics*, 59 (1-2): 121-138.
- [172] OECD. (2009). *Is Informal Normal? Towards More and Better Jobs in Developing Countries*. OECD, Paris.
- [173] O'Neill, D. M. (1983). *Growth of the underground economy, 1950-81: some evidence from the current population survey*, Study for the Joint Economic Committee. U.S. Congress Joint Committee Print, U.S. Gov. Printing Office. Washington, DC, pp. 98-122.
- [174] ONPES and World Bank . (2014). *Investing in people to fight poverty in Haiti*. Washington, DC.



- [175] Orlando, M. (2001). The informal sector in Venezuela: Catalyst of Hindrance for poverty reduction, Catholic University Andres Bello. Instituto de Investigaciones Económicas y Sociales, *Poverty Project Paper*, vol. 2, Venezuela.
- [176] OXFAM. (2005). Kicking down the door: How upcoming WTO talks threaten farmers in poor countries. *Oxfam Briefing Paper*, 72.
- [177] Pahl, R. E. (1984). Divisions of labour. Oxford: Blackwell.
- [178] Park, T. (1979). Reconciliation between Personal Income and Taxable Income. Mimeo, Washington D.C. Bureau of Economic Analysis, pp. 1947–77.
- [179] Pedersen, J., and Lakewood, K. (2001). Determination of a poverty line for Haiti. FAFO Institute of Applied International Studies.
- [180] Perry, G. E., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., and Saavedre-Chanduvi, J. (2007). Informality: Exit and Exclusion in Latin America. World Bank, Washington DC.
- [181] PNUD, BID, BM, and SNU. (2010). Haïti : Résumé du PDNA du tremblement de terre, évaluation des dommages des pertes et des besoins.
- [182] Portes, A., Castells, M., and Benton, L. (1989). The Informal Economy. Baltimore, Md.: Johns Hopkins University Press.
- [183] Pradhan, M., and Van Soest, A. (1995). Formal and informal sector employment in urban areas of Bolivia. *Labor Economics*, 2, 275–297.
- [184] Pratap, S., and Quintin, E. (2006). Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review*, 50(7), 1817–1841.
- [185] Raykov, T., and Marcoulides, G. (2000). A method for comparing completely standardized solutions in multiple groups. *Structural Equation Modeling*, 7(2), 292–308.
- [186] Richardson, S., and Green, P. (1997). On Bayesian analysis of mixtures with an unknown number of components (with Discussion). *Journal of the Royal Statistical Society*, B 59, 731–92.
- [187] Rodella, A. S., and Scot, T. (2016). Sifting through the data, Labor Markets in Haiti through a turbulent decade (2001–2012). *Policy Research Working Paper 7562*, World Bank.

- [188] Rosen, S. (1978). Substitution and Division of Labor. *Economica*, 45, 235-250.
- [189] Rosen, S. (1981). The Economics of Superstars. *American Economic Review*, 71(5): 845–58.
- [190] Rosenzweig, M. R. (1988). *Labor Markets in Low-Income Countries*. In Hollis Chenery and T. N. Srinivasan, eds. *Handbook of Development Economics Volume 1*. New York: North-Holland.
- [191] Rossi, P. E., Gilula, Z., and Allenby, G. (2001). Overcoming scale usage heterogeneity. *Journal of the American Statistical Association*, 96, 20-31.
- [192] Sassen, S. (1997). Informalisation in Advanced Market Economies. Issues in *Development Discussion Paper 20*. Geneva: ILO.
- [193] Satorra, A., and Bentler, P. M. (1994). *Corrections to test statistics and standard errors on covariance structure analysis*. In A. von Eye and C. C. Clogg (Eds.), *Latent variables analysis* (pp. 399–419). Thousand Oaks, CA: Sage.
- [194] Schick, A. (2005). Sustainable Budget Policy: Concepts and Approaches. *OECD Journal on Budgeting*, 5(1), pp. 107-26.
- [195] Schneider, F. (2010). The Influence of Public Institutions on the Shadow Economy: An Empirical Investigation for OECD Countries, *European Journal of Law and Economics*, 6(3), 441-468.
- [196] Schneider, F. (2007). Shadow Economies and Corruption All Over the World: New Estimates for 145 Countries. *Economics*, 2007-9.
- [197] Schneider, F., Buehn, A., and Montenegro, C. E. (2010). *Shadow Economies All Over the World: New Estimates for 162 Countries from 1999 to 2007*. World Bank Policy Research Working Paper Series 5356, 54 p.
- [198] Schneider, F. (2005). Shadow Economies around the World: What Do We Really Know? *European Journal of Political Economy*, 21(4), 598-642.
- [199] Schneider F and Enste D.H. (2000). Shadow economies: Size, Causes, and Consequences, *Journal of Economic Literature*, 38(1), 73-110.
- [200] Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review*, 51(1), 117.

- [201] Schultz, T. W. (1962). Reflections on Investment in Man. *Journal of Political Economy*, LXX Supplement: 1-8.
- [202] Schwarz, G. (1978). Estimating a dimension of a model. *The Annals of Statistics*, 6(2), 461-464.
- [203] Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica*, 4, 639-650.
- [204] Sookram, S., and Watson, P. (2008). The Informal Sector and Gender in the Caribbean: The Case of Trinidad and Tobago. *Journal of Eastern Caribbean Studies*, 130, 43-68.
- [205] Spence, M. (1973). Job Market Signalling. *Quarterly Journal of Economics* 87, 355-74.
- [206] Spilimbergo, A., Symansky, S., Blanchard, O., and Cottarelli, C. (2008). *Fiscal Policy for the Crisis*. Washington, D. C.: IMF Staff Position Note No 08/01.
- [207] Stephens, M. (2000). Bayesian analysis of mixture models with an unknown number of components: an alternative to reversible jump methods. *The Annals of Statistics*, 28, 40-74.
- [208] Tadesse, M. G., Sha, N., and Vannucci, M. (2005). Bayesian variable selection in clustering high-dimensional data. *Journal of American Statistical Association*, 100, 602-17.
- [209] Tanzi, V. (1980). The Underground Economy in the United States: Estimates and Implications . *Banca Nazionale del Lavoro quarterly review*, 135(4), 427-53.
- [210] Tanzi, V. (1983). The underground economy in the united states: annual estimates, . *IMF staff papers*, 1930- 80, Palgrave Macmillan, 30(2), 283-305.
- [211] Teh, Y. W., Jordan, M., Beal, M., and Blei, D. (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101 (476), 1566-1581.
- [212] Tenenhaus, M. (1999). L'approche PLS. *Revue de Statistique appliquée*, 47(2), 5-40.
- [213] Tenenhaus, M. (2004). *PLS Regression and PLS Path Modeling for Multiple Table Analysis*. COMPSTAT 2004 Proceedings. Physica-Verlag, Springer, Wurzburg, Berlin.
- [214] Tenenhaus, M., Esposito, V., Chatelin, Y.-M., and Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis* 48(1), 159-205. 23, 25,37,41, 46.

- [215] Teo, T., Tsai, L. T., and Yang, C.-C. (2013). *Application of Structural Equation Modeling in Educational Research: An Introduction* in Myint Swe Khine, ed. Rotterdam: Sense Publishers.
- [216] Thiessen, U. (1997). The informal economy in Eastern Europe: the example of Ukraine. *textitEconomic Bulletin*, 34 (6): 19-24.
- [217] Thompson, B. (1992). Two and one-half decades of leadership in measurement and evaluation. *Journal of Counseling and Development*, 70, pp. 434–438.
- [218] UNDP. (2002). *La Bonne Gouvernance, Un Défi Majeur Pour le Développement Durable en Haïti, Rapport National sur le Développement Humain*. Port-au-Prince.
- [219] UNDP. (2003). *Les Objectifs du Millénaire pour le Développement, Rapport National Provisoire*, Port-au-Prince.
- [220] Verner, D. (2008). *Making the Poor Haitians Count Takes More Than Counting the Poor: A Poverty and Labor Market Assessment of Rural and Urban Haiti Based on the First Household Survey for Haiti*. World Bank.
- [221] Wetzstein, M. E. (2013). *Microeconomic Theory: Concepts and Connections*. Second edition Routledge, New York.
- [222] Williams, C. C. (2013). Tackling Europe's economy: A critical evaluation of the neo-liberal de-regulatory perspective. *Journal of Contemporary European Research*, 9(2), pp. 261–278.
- [223] Williams, C. C., and Windebank, J. (2002). Why do people engage in paid informal work? A comparison of affluent suburbs and deprived urban neighbourhoods in Britain. *Community, Work and Family*, 5(1), 67-83.
- [224] Williams, C. C., and Windebank, J. (2003). *Poverty and the Third Way*. London: Routledge.
- [225] Williams, C., and Round, J. (2008). Rethorizing the Nature of informal employment: Some lessons from Ukraine. *International Sociology*, 23(3) 367-388.
- [226] Wold, H. (1985). Partial Least Squares, *Encyclopedia of Statistical Sciences*,. vol. 6, Kotz S. and Johnson, N.L. (Eds), John Wiley and sons, New York, pp. 585-591.
- [227] Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. London: The MIT Press.

- [228] World Bank. (1998). Haiti: The Challenges of Poverty Reduction. Washington, DC: Report 17242-HA.
- [229] World Bank, (2007). *Fiscal Policy and Economic Growth*, the World Bank, Washington, DC.
- [230] World Bank. (2012a). Doing Business 2012. Washington DC, USA.
- [231] World bank. (2012b). Worldwide Governance Indicators: 1996-2011, Washington, DC.
- [232] World Bank and SEDLAC. (2005/06). dataset.
- [233] Yuan, K. H., and Bentler, P. M. (1998). Normal theory based test statistics in structural equation modeling. *British Journal of Mathematical and Statistical Psychology*, 51, 289-309.