



# Gender disparities in Africa's labour markets: An analysis of survey data from Ethiopia and Tanzania

Pablo Suarez Robles

## ► To cite this version:

Pablo Suarez Robles. Gender disparities in Africa's labour markets: An analysis of survey data from Ethiopia and Tanzania. Economics and Finance. Université Paris-Est, 2012. English. NNT: 2012PEST0057 . tel-00909494

**HAL Id: tel-00909494**

**<https://theses.hal.science/tel-00909494>**

Submitted on 26 Nov 2013

**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.

University Paris East  
Faculty of Economics and Management  
Doctoral School « Organisations, Markets, Institutions » (OMI)  
Research Team on the Use of Panel Data in Economics (ERUDITE)

**Gender Disparities in Africa's Labour Markets:  
An Analysis of Survey Data from Ethiopia and Tanzania**

Pablo Suárez Robles

PhD Supervisor: M. Alexandre Kolev

A thesis submitted for the degree of Doctor of Philosophy in Economics

July 2012

Defended on October 25<sup>th</sup> 2012

Jury members:

Philippe Adair (President), University Paris East

Alexandre Kolev (PhD Supervisor), University Paris East and ITC-ILO

Mathilde Maurel (Reviewer), University Paris 1 Panthéon-Sorbonne

Christophe Nordman (Reviewer), IRD/DIAL and University Paris-Dauphine

Pierella Paci, World Bank

Catherine Saget, ILO

L'Université Paris Est n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses: ces opinions doivent être considérées comme propres à leurs auteurs.

## **Acknowledgements**

I am deeply thankful to my PhD Supervisor, M. Alexandre Kolev, for his invaluable support and guidance throughout this research. I will always be grateful to him for granting me the opportunity to work on this PhD thesis and on other interesting and challenging projects, be it with the University Paris East, the French Development Agency, the World Bank or the International Training Centre of the International Labour Organization. His inspirational debates, insightful suggestions and continuous encouragement have been determinant in forging into me the necessary skills and will to carry through these projects. The competitive expertise and innovative spirit he provided will be a source of inspiration in my professional career and personal life.

I would further like to express my gratitude to M. Christophe Nordman for his priceless help in providing me with valuable comments and suggestions on earlier versions of this work.

I also wish to thank my PhD committee members and all the people to whom I am indebted for their helpful advices and feedbacks that contributed to the accomplishment of my thesis. I am especially grateful to Jorge Saba Arbache, Caterina Ruggeri Laderchi, Alexandru Cojocaru, Mayra Buvinic, Gizaw Molla, Cecilia Valdivieso and Emily Kallaur from the World Bank, Blandine Ledoux from UNESCO/BREDA, Philippe Adair from the University Paris East, Marc Gurgand and Denis Cogneau from the Paris School of Economics, Luigi Benfratello from the University of Naples, Assefa Admassie from the Ethiopian Economic Association / Ethiopian Economic Policy Research Institute, Marcel Fafchamps and two anonymous referees from the Journal of African Economies, and Alain de Janvry from the World Bank Economic Review.

## Table of Contents

<b>Acknowledgements .....</b>	<b>iv</b>
<b>Abstract.....</b>	<b>xii</b>
<b>Chapter 1 Introduction.....</b>	<b>1</b>
1.1 Why looking at gender disparities in Africa's labour markets?.....	2
1.2 Objective and structure of this thesis .....	7
1.3 Countries selected .....	13
1.4 Overview of main findings .....	17
<b>Chapter 2 Addressing the Gender Pay Gap in Ethiopia: How Crucial is the Quest for Education Parity? .....</b>	<b>23</b>
2.1 Introduction.....	23
2.2 Data and concepts .....	25
2.2.1 Ethiopia Labour Force Survey 2005 .....	26
2.2.2 Definitions and measurement issues .....	27
Wage employment .....	27
Formal and informal wage employment .....	27
Earnings .....	28
2.2.3 Descriptive statistics .....	29
Gender disparities in the labour force and in employment status .....	29
The unadjusted gender pay gap.....	31
Gender disparities in education characteristics among the wage employed .....	32
Gender disparities across sectors of activity and occupations .....	33
2.3 Methodology .....	35
2.3.1 Estimation of wage equations .....	36
Heckman's two-step estimation procedure.....	36
Bourguignon-Fournier-Gurgand two-step estimation procedure.....	40
Quantile regression analysis .....	44

2.3.2 Decomposition of the gender wage gap.....	47
Neumark and Cotton decomposition procedures .....	47
Treatment of the sample selection correction .....	49
2.4 Results.....	50
2.4.1 Estimations of the wage equations .....	50
2.4.2 Wage decompositions .....	54
2.5 Conclusion .....	60
Appendix A. Generalized Lorenz curves for hourly earnings .....	63
Appendix B. Earnings equations.....	65
Appendix C. Gender earnings gap decompositions .....	75
<b>Chapter 3 Analysing the Nature and Extent of Gender Inequalities in Time Use: New Insights from Ethiopia.....</b>	<b>79</b>
3.1 Introduction.....	79
3.2 Data and concepts .....	81
3.2.1 Brief theoretical literature review .....	81
3.2.2 Ethiopia Labour Force Survey 2005 .....	83
3.2.3 Definitions and measurement issues .....	85
3.3 Methodology .....	87
3.3.1 Decomposition of total work time .....	87
3.3.2 The determinants of market and household work time .....	88
A generalized Tobit model for market work time.....	90
A standard Tobit model for housework time .....	92
Decomposition of the total marginal effect.....	93
3.4 Results.....	94
3.4.1 Decomposition of total work time .....	94
3.4.2 Further disaggregations of time use estimates .....	104
3.4.3 The determinants of market and household work time .....	108
Preliminary considerations on models' specification and data quality issues .....	109
Ancillary parameters of the generalized Tobit models for market work time .....	113
Sex and area of residence.....	113

Complementarity/substitutability of time allocation decisions.....	114
Human capital and other individual characteristics .....	115
Composition of the household .....	117
3.5 Conclusion .....	119
Appendix A. Generalized Lorenz curves for the market and household work time .....	124
Appendix B. Housework time equations .....	126
Appendix C. Market work time equations.....	128
<b>Chapter 4 Assessing the Impact of Informality on Earnings in Tanzania: Is There a Penalty for Women? .....</b>	<b>134</b>
4.1 Introduction.....	134
4.2 Data, concepts and descriptive statistics .....	138
4.2.1 Tanzania Integrated Labour Force Survey 2006.....	139
4.2.2 Key concepts.....	140
Employment.....	140
Informal employment.....	140
Labour income .....	144
4.2.3 Descriptive statistics on gender-differentiated employment patterns .....	145
4.3 Methodology .....	150
4.3.1 Informal employment effects under the assumption of homogeneity.....	151
4.3.2 Informal employment effects under the assumption of partial heterogeneity.....	153
4.3.3 Informal employment effects under the assumption of full heterogeneity .....	156
4.4 Results.....	157
4.4.1 Informal employment effects under the assumption of homogeneity.....	157
4.4.2 Informal employment effects under the assumption of partial heterogeneity.....	162
4.4.3 Informal employment effects under the assumption of full heterogeneity .....	167
4.5 Conclusion .....	170
Appendix A. Methodological and estimation issues.....	172
Appendix B. Descriptive statistics.....	181
Appendix C. Informal employment effects under the assumption of homogeneity .....	184
Appendix D. Informal employment effects under the assumption of partial heterogeneity ....	185



Appendix E. Informal employment effects under the assumption of full heterogeneity .....	198
<b>Chapter 5 Conclusion .....</b>	<b>205</b>
<b>References.....</b>	<b>211</b>

## List of Tables

Table 2.1: Selected labour market indicators.....	30
Table 2.2: Nature and terms of wage employment in total wage employment.....	30
Table 2.3: Unadjusted gender hourly wage gaps .....	31
Table 2.4: Levels of education among the wage employed by gender .....	33
Table 2.5: Industries and occupations by wage employment sector and gender .....	35
Table 2.6: Summary statistics of the variables used in the earnings equations .....	65
Table 2.7: OLS and selectivity corrected (Heckman's two-step method) log hourly earnings equations in wage employment by gender.....	67
Table 2.8: OLS and selectivity corrected (Heckman's two-step method) log hourly earnings equations in wage employment by gender and age cohort .....	69
Table 2.9: Simultaneous quantile log hourly earnings regression estimates in wage employment by gender.....	71
Table 2.10: OLS and selectivity corrected (BFG method) log hourly earnings equations in wage employment by gender and sector .....	73
Table 2.11: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment (OLS estimates) .....	75
Table 2.12: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment by age cohort (OLS estimates).....	76
Table 2.13: Neumark and Cotton decompositions of the gender predicted log hourly earnings differential in wage employment by quartile .....	77
Table 2.14: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment by sector (OLS estimates).....	78
Table 3.1: Decomposition of the average total work hours per week by gender .....	96
Table 3.2: Decomposition of the average total work hours per week by place of residence and gender.....	99
Table 3.3: Marginal effects from Tobit model for housework time (evaluated at sample mean) by gender and place of residence .....	126

Table 3.4: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender.....	128
Table 3.5: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender in urban areas.....	130
Table 3.6: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender in rural areas .....	132
Table 4.1: Selected key labour markets indicators .....	145
Table 4.2: Distribution of workers by type of production unit and type of job (status in employment), men in main job .....	148
Table 4.3: Distribution of workers by type of production unit and type of job (status in employment), women in main job .....	149
Table 4.4: List and definition of covariates .....	181
Table 4.5: Mean covariates values by employment status and gender .....	182
Table 4.6: OLS regressions of log hourly income from main job by employment status and gender.....	184
Table 4.7: Propensity score Probit regression models predicting informal employment by employment status and gender.....	185
Table 4.8: Mean covariates values by propensity score strata and employment status, men in paid employment .....	187
Table 4.9: Mean covariates values by propensity score strata and employment status, women in paid employment.....	189
Table 4.10: Mean covariates values by propensity score strata and employment status, men in self-employment .....	191
Table 4.11: Mean covariates values by propensity score strata and employment status, women in self-employment .....	193
Table 4.12: Frequency counts per propensity score stratum.....	195
Table 4.13: Heterogeneous effects of informal employment on log hourly income from main job by employment status and gender (SM-HTE) .....	196
Table 4.14: Matching estimates effects of informal employment on log hourly income from main job by employment status and gender.....	198

## List of Figures

Figure 2.1: Generalized Lorenz curves for hourly earnings in main occupation by gender and age cohort. ....	63
Figure 2.2: Generalized Lorenz curves for hourly earnings in main occupation by gender and wage employment sector. ....	64
Figure 3.1: Generalized Lorenz curves for the total work time by gender in urban areas .....	103
Figure 3.2: Generalized Lorenz curves for the total work time by gender in rural areas.....	103
Figure 3.3: Generalized Lorenz curves for the market work time by gender and place of residence .....	124
Figure 3.4: Generalized Lorenz curves for the housework time by gender and place of residence .....	125
Figure 4.1: Heterogeneous effects of informal employment on log hourly income from main job by employment status and gender (SM-HTE) .....	197
Figure 4.2: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), men in paid employment .....	199
Figure 4.3: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), women in paid employment .....	200
Figure 4.4: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), men in self-employment .....	201
Figure 4.5: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), women in self-employment .....	202
Figure 4.6: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), graphs of the nonparametric smoothed curves without plotted points, men in paid employment .....	203
Figure 4.7: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), graphs of the nonparametric smoothed curves without plotted points, women in paid employment .....	204

## Abstract

### ENGLISH

The main objective of this thesis is to contribute to our better understanding of the main factors behind large and persistent gender disparities in Africa's labour markets. This work looks at three key dimensions of labour market gender inequality in Africa: (i) the gender wage gap, (ii) gender inequalities in allocating time to market and household work, and (iii) the gender-differentiated income effect of informality. Chapter 2 shows that, in Ethiopia, progress towards gender equity in education is important to improve women's wages but not enough to close most of the gender wage differential. Other interventions would be needed as, for instance, information campaigns and other awareness-raising efforts in support of the anti-discriminatory provisions of Ethiopia's own constitution and legislation, to compensate for the adverse impact of unobservable factors (discriminatory practices, social and cultural norms...), that directly contribute to the gender wage gap and indirectly, through job selection. Chapter 3 highlights the coexistence of two phenomena in Ethiopia, a strong gender-based division of labour and a double work burden on women. The country would benefit from pursuing and intensifying its efforts to ensure better access to education at all levels for women, and from providing better information and enforcement of the law in support of women's economic and social well-being, as it would help changing mentalities and attitudes that impede women to take full advantage of their abilities and that keep them subordinated to men. Finally, in Chapter 4 we observe that, in Tanzania, women face a significantly higher informal employment wage penalty than men. To explain this result, we conjecture that the exclusion hypothesis, according to which individuals are denied access to formal jobs due to the disproportionate constraints they face (burden of household responsibilities, lack of adequate infrastructure...), is more acute among women.

**JEL classification** : J16, J22, J24, J31, J42, J71

### FRENCH

Cette thèse se centre sur trois sources importantes d'inégalité de genre sur le marché du travail en Afrique : (i) les salaires, (ii) l'allocation du temps entre travail marchand et travail domestique, et (iii) les revenus de l'emploi informel. Le Chapitre 2 montre que, en Ethiopie, les progrès en matière d'égalité de genre dans l'éducation sont nécessaires pour accroître le salaire des femmes, mais pas suffisants pour enrayer l'écart de salaire avec les hommes. D'autres interventions seraient nécessaires, telles que des campagnes d'information et d'autres efforts de sensibilisation sur les dispositions antidiscriminatoires de la législation nationale, afin de compenser l'effet adverse de facteurs non-observables (pratiques discriminatoires, normes culturelles et sociales...) qui contribuent directement au différentiel de salaire entre les sexes et indirectement, à travers la sélection dans l'emploi. Le Chapitre 3 met en lumière la coexistence de deux phénomènes en Ethiopie, une forte division du travail selon le genre et une double charge de travail des femmes. Le pays gagnerait à poursuivre et intensifier ses efforts pour un meilleur accès des femmes à tous les niveaux d'éducation, et pour une meilleure diffusion et application de la loi en faveur du bien-être économique et social des femmes, car cela contribuerait au changement des mentalités et attitudes qui empêchent les femmes d'exploiter pleinement leur potentiel et les subordonnent aux hommes. Finalement, dans le Chapitre 4 nous observons que les femmes occupant un emploi informel en Tanzanie subissent une pénalité salariale bien plus élevée que celle des hommes. Pour expliquer ce résultat, nous conjecturons que l'hypothèse d'exclusion, selon laquelle les individus n'ont pas accès à l'emploi formel en raison des contraintes disproportionnées auxquelles ils font face (fardeau des tâches domestiques, manque d'infrastructures adéquates...), est plus forte parmi les femmes.

**Classification JEL** : J16, J22, J24, J31, J42, J71

# **Chapter 1**

## **Introduction**

This research aims at contributing to the scarce but much-needed empirical literature on labour market gender inequality in Africa. It is the fruit of a long period of research, which started six years ago, when, freshly graduated with a Master's degree in economics from the Paris Sorbonne University, I was given the chance to take part of a joint research project of the World Bank and the French Development Agency on gender disparities in Africa's labour markets and, at the same time, to start a Ph.D. on the subject.

While progressing in my research, it became clear to me that achieving gender equality was fundamental, not only as a human rights concern, but also for the sake of social and economic development. Going through the empirical literature on Africa, I further realized that it was probably there, in that continent, the poorest and less developed of the world, that gender inequalities in many respects were the more pervasive and deeply rooted. Paradoxically, I also discovered that relatively few studies had documented the huge disadvantaged faced by African women in their working lives. Africa appeared to be lagging behind in this field of research while it was very likely to be one of the regions, if not the region, most in need. As empirical analysis is a prerequisite for evidence-based policy-making, it became clear to me that real advances towards the ultimate goal of gender equality in Africa would not be reached unless more efforts were made in assessing and understanding the nature, extent and root causes of its multiple sources.

My motivation to carry-out this research on gender disparities in Africa's labour markets stems from this awareness and a desire to add a modest but concrete contribution to the recent but still

insufficient research community's efforts to understand what drives women disadvantages in Africa.

### **1.1 Why looking at gender disparities in Africa's labour markets?**

The rationale for this research on gender inequality in labour markets in the African context departs from the recognition that, notwithstanding the increasing attention gained in recent years, empirical research on the subject remains scarce, both in itself and compared to industrialized countries and other developing regions of the world. While relatively few studies have documented the huge disadvantages faced by African women in their working lives, the fact that empirical analysis for Africa lags behind in this field of research is all the more critical. On the one hand, it is now well established that gender equality matters for social and economic development. On the other hand, gender inequality is a matter of concern from a human rights perspective.

On the first point, research indicates that improvements in gender equality can generate gains in economic efficiency and improvements in other development outcomes (World Bank, 2011). A growing body of the literature finds that gender inequalities are detrimental to society at large, and that unlocking the full economic potential of women would importantly contribute to poverty reduction and growth stimulation. Morrison *et al.* (2007) presented a conceptual framework for understanding the links between gender equality and aggregate poverty reduction and growth. Increased gender equality through better women's access to markets (labour, credit, land...), education and health, and through mother's greater control over decision-making in the household, would lead to increased women's labour force participation, productivity and earnings, and to improved children's well-being. These improvements would translate in fine into

current poverty reduction and short-term growth thanks to higher incomes and consumption expenditures, and into future poverty reduction and long-term growth thanks to higher savings and better children's health and educational attainment, and productivity as adults.

Existing studies indicate, in particular, that education is crucial as it substantially lowers the magnitude of gender inequalities in labour income and underemployment, and substantially increases men's and women's probability of getting a paid job (see, for instance, Arbache *et al.*, 2010). Moreover, education renders the most rewarding (public and formal private) employment sectors more accessible to men and women, and has positive returns even in the informal sector (De Vreyer and Roubaud, 2012).

Education matters, not only as a positive determinant of labour market outcomes, but also because of its beneficial impact on the economy as a whole. Klasen (1999) found that gender inequality in education has a significant negative impact on economic growth, and appears to be an important factor contributing to Africa's poor growth over the past thirty years. As pointed out in Ward *et al.* (2010), the estimate of loss of growth owing to gender inequality in education rises to 0.38 per cent per annum in sub-Saharan Africa. This accounts for 11 per cent of the growth difference between this region and East Asia and the Pacific (Klasen and Lamanna, 2008).

In addition to increasing growth, greater gender equality in education promotes other important development goals, including lower fertility and lower child mortality. Better-educated girls and women are likely to have fewer children. The decline in fertility associated with greater gender equality can have profound economic impacts. A fall in fertility leads to a lower dependency ratio and tends to increase per capita output, providing a demographic dividend.

The intergenerational transmission and accumulation of human capital rely mainly upon mothers as they are the primary caregivers of children. Indeed, it is well-recognized that women directly contribute more to the rearing of children than men and they have the primary responsibility in the household for children's health, nutrition and well-being. Thus, gender equality, by giving



women more bargaining power in the domestic sphere, could improve children's health and educational opportunities, bringing clear and direct benefits for the future stock of human capital in an economy (Ward *et al.*, 2010)<sup>1</sup>.

Concerning explicitly gender inequalities in labour markets, Arbache *et al.* (2010) pointed out that, yet, still relatively little is known in many African countries, and even less is known about how to design more effective policies to reduce them. Lack of suitable data on African labour markets is a major obstacle to obtain accurate empirical evidence on the multiple sources of work-related gender disparities, as well as on their extent and drivers.

Overall, the little evidence available shows that, in Africa, women typically experience worse labour market outcomes than men, with higher levels of unemployment and underemployment, and lower access to productive and paid employment. Occupational segregation by sex is widespread and leads to allocational inefficiencies and earnings gaps. Women tend to adapt their preferences to occupations that are socially acceptable, that is they are inclined to pursue careers that are more conducive to combining work and reproductive responsibilities, which leads to their concentration in informal and precarious employment, where pay and conditions of work are worse than in public and formal jobs (World Bank, 2011).

Lower human capital endowments, the burden of household and care responsibilities, under-provision of basic public goods and lack of adequate infrastructure services, social and cultural norms, and labour market and workplace discrimination based on gender are considered as important factors explaining women's disadvantages relative to men in Africa's labour markets.

---

<sup>1</sup> Several empirical studies have documented the fact that, when women have better command over income resources, decisions on how these resources are spent tend to favour children more in terms of human capital investment (see, for example, Bourguignon and Chiappori 1992, Hoddinott and Haddad 1995, Browning and Chiappori 1998, Bussolo *et al.* 2009, Angel-Urdinola and Wodon 2010, Backiny-Yetna and Wodon 2010).

Besides the recognition of its crucial importance for the pace of development, gender inequality has also become a major issue in the global human rights agenda during the last decades.

In fact, international support for the rights of women is not new and started soon after the end of World War II. The Charter of the United Nations (UN) signed in 1945, and the Universal Declaration of Human Rights adopted in 1948 by the UN General Assembly, recognize, promote and encourage respect for human rights and for fundamental freedoms for all men and women.

A further significant milestone in women's rights is the UN Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) adopted in 1979, which explicitly defines discrimination against women and sets up an agenda for national action to end such discrimination.

The year 1975 was declared by the UN General Assembly as the International Women's Year, and the years 1976-1985 were subsequently declared as the UN Decade for Women. Since then, several World Conferences on Women took place, in particular, to review and appraise the achievements of the UN decade for Women and to adopt forward-looking strategies. The Fourth of its kind, held in Beijing in 1995, reaffirmed that women's rights are human rights and that gender equality is an issue of universal concern, benefiting all, and committed to specific actions to ensure respect for those rights.

In 2000, building upon a decade of major United Nations conferences and summits, world leaders adopted the United Nations Millennium Declaration, committing their nations to a new global partnership to reduce extreme poverty and setting out a series of time-bound targets, with a deadline of 2015, known as the Millennium Development Goals (MDGs). Gender issues were integrated in many of the subsequent MDGs, and explicitly in the 3<sup>rd</sup> Goal "Promote gender equality and empower women", and the 5<sup>th</sup> Goal "Improve maternal health". The 2010 Summit on the MDGs took stock of the progress in the achievement of objectives and concluded with the adoption of a global action plan "Keeping the Promise: United to Achieve the Millennium

Development Goals”. As regards the achievement of the 3<sup>rd</sup> Goal on gender equality and women’s empowerment, the Summit pointed out, among others, the following facts. Gender gaps in access to education have narrowed, but disparities remain high in university-level education and in some developing regions. Despite progress made, men continue to outnumber women in paid employment, and women are often relegated to vulnerable forms of employment.

Finally, it is worth mentioning that, in 2010, the United Nations General Assembly voted to create a new UN entity, called UN Women, tasked with accelerating progress in achieving gender equality and women’s empowerment.

Gender equality in the world of work is also a major objective contained in international labour standards. In this respect, it can be said that important developments have taken place since the creation of the International Labour Organization (ILO) in 1919. The four key ILO gender equality Conventions are: Equal Remuneration Convention (No. 100), Discrimination (Employment and Occupation) Convention (No. 111), Workers with Family Responsibilities Convention (No. 156), and Maternity Protection Convention (No. 183). These conventions have been ratified by a large number of countries, including in Africa. For instance, the two fundamental conventions on discrimination, which are the conventions No. 100 and No. 111, have been ratified by, respectively, 51 and 53 of the 54 ILO member countries in Africa.

At its 92<sup>nd</sup> session in 2004, the International Labour Conference adopted the Resolution concerning the Promotion of Gender Equality, Pay Equity and Maternity Protection, and the 98<sup>th</sup> session of the International Labour Conference of 2009 adopted the Resolution concerning Gender Equality at the Heart of Decent Work. The Domestic Work Convention (No. 189) adopted in 2011 is the latest important international achievement in support of women’s rights. This convention offers specific protection to domestic workers who are for the most part women

and girls. It lays down basic rights and principles, and requires countries to take a series of measures with a view to making decent work a reality for domestic workers.

While many African countries have ratified these conventions, their effective application and enforcement remain an issue. Tribal, customary and religious laws and norms, which are largely prevalent in many African countries, are serious obstacles to the incorporation and effectiveness of international labour standards and to achieving gender equality and empowering women. Empirical research on the situation of women in Africa's labour markets is thus essential.

## **1.2 Objective and structure of this thesis**

Conducting empirical research on Africa is often a challenge, as many surveys are not nationally-representative and not conducted on a regular basis, collect incomplete and limited information on respondents' labour market and job characteristics, and use conceptual and statistical definitions, and data collection methods that differ from international standards. In this thesis, we were able to overcome some of these challenges by exploiting two nationally-representative labour force surveys conducted in 2005 and 2006 respectively in Ethiopia and Tanzania. These labour force surveys offer great opportunities to analyse the essence of gender inequalities in Africa's labour markets and the mechanisms surrounding them.

The main objective of this thesis was to uncover what are the most important and pervasive sources of gender disparities in Africa's labour markets, relying on the 2005 Ethiopia Labour Force Survey and the 2006 Tanzania Integrated Labour Force Survey. This thesis consists of three empirical studies, each one filling-in an important knowledge gap in one of the following key dimensions of labour market gender inequality in Africa: i) the gender pay gap, ii) gender

inequalities in allocating time to market and household work, and iii) the gender-differentiated income effect of informality.

An analysis of the gender wage gap in Ethiopia is performed in Chapter 2. Weichselbaumer and Winter-Ebmer (2003) point out that, in contrast with the abundant literature in developed countries, and the growing number of studies for emerging countries, fewer studies have actually attempted to address this important question in the case of Africa.

Moreover, most of the studies that actually aim to explain the extent of the gender pay gap in Africa do not adopt a comprehensive approach. Little is known on the gender wage gap for different points in the wage distribution, different age cohorts or different types of wage employment. And this, despite the fact that it is well-known that paid-employed workers are highly heterogeneous – they differ not only in their background characteristics but also in how they respond to a particular situation –, that their decision-making processes and constraints vary over different stages of the life-cycle, and that paid-employment is clearly segmented along sectoral lines such as public vs. private and formal vs. informal.

According to existing empirical evidence, sub-Saharan Africa is the region of the world with the highest earnings disparities as males earn on average 48 per cent more than females (Ñopo *et al.*, 2011). The unexplained gender earnings gap that remains after controlling for differences in (individual, demographic and job-related) characteristics is often significantly lower but still huge. For instance, Nordman *et al.* (2011) found, using data from an original series of urban household surveys, that, in seven West African cities, only about 40 per cent of the raw gender earnings gap on average is explained by differences in observable characteristics. Indeed, women's disadvantages in terms of human capital and job attributes are not the sole responsible for the gender earnings differential we observe in most African countries. Discriminatory

practices, gender specific preferences, cultural and other non-observable factors are likely to be also important contributors.

Another important dimension of gender inequality, which is investigated in Chapter 3, is disparity in time use. As regards information on this dimension, the situation is even worse. In most African countries, there are almost no studies focusing on time use, and there are two main reasons for that. First, comprehensive data on time allocation that are nationally representative are really scarce, not to say almost non-existent. And when they exist, they are often imprecise and subjective. Second, the household economy, which is largely invisible and uncounted in economic data and in the system of national accounts (SNA), is most of the time neglected from labour studies, while its recognition is essential to obtain a more complete and comprehensive picture and a better understanding of employment and labour effort. The household economy is of particular importance, not only because reproductive activities are essential for family survival, but also from a gender standpoint, as it is where women work predominantly. In addition, synergies and trade-offs between market and household work are evident. Hours spent working in productive and reproductive activities, whether complementary or substitutes, do not represent separate decisions but rather are outcomes of an optimization process in which time allocation decisions are jointly determined.

The little research done so far reveals that gender differences in time use are particularly impressive in Africa. There is a strong sexual division of labour as illustrated by the fact that women bear the brunt of housework (reproductive activities) while men are mostly responsible for market work (productive activities). Women suffer from time deprivation due to the multiple roles they play, especially in the domestic realm where they are primarily responsible for household chores, which are essential for family survival but are usually low-productive, labour-intensive, and time- and energy-consuming. The housework burden on women limits their time

available to market work and allows them to engage only in productive activities compatible with their household duties. As a result, women have little choice but to take informal and low-quality jobs offering flexible work arrangements and allowing for easy entry into or exit from the labour market.

However, women experience a double work burden, as they tend to accumulate both productive and reproductive activities, assuming a dual role as workers and housekeepers, unlike men, who traditionally endorse the role of breadwinner and focus only on market work. As a consequence, the total workload on women exceeds by far that on men. This phenomenon is observed to a greater extent in rural areas, where lack of adequate physical capacities (such as roads, utility supply systems, communications systems, water and waste disposal systems) and the under-provision of services flowing from those facilities impose greater work burdens and lengthen the time it takes people, in particular women, to perform activities related to household survival, reducing the time for participating in more economically productive activities.

Because time is a limited resource, the more time an individual spends working, the more her time for rest and leisure will be reduced. When the total work time exceeds a certain threshold (the so-called time poverty line), individuals do not have enough time for rest and leisure, and thereby are considered time poor.

For instance, Bardasi and Wodon (2006) found that, in Guinea, 17 per cent of all adults are time poor. This headcount is much higher for women (24 per cent) than men (9 per cent), and higher in rural areas (19 per cent) compared with urban areas (15 per cent)<sup>2</sup>.

In Chapter 4, we devote a specific attention to the effect of informality on earnings and its gender dimension, an area which has attracted little attention in the case of Africa. Our research focuses

---

<sup>2</sup>Bardasi and Wodon (2010) proposed a new definition of time poverty as working long hours without choice because an individual's household is poor or would be at risk of falling into poverty if the individual reduced her working hours below a certain time poverty line. Using this new definition, they found that women are even more likely to be time poor than men in Guinea.

on the following important questions: What is the magnitude of informal employment and its different sub-segments? How differently do male and female informal workers fare? What drives gender differences in informal employment outcomes?

Addressing these issues is important because labour markets in Africa are highly heterogeneous and segmented. Earnings determination process significantly differs according to status, sector and type (formal/informal) of employment. There are important barriers inhibiting mobility between these employment segments. For instance, formal jobs are scarce and hardly accessible to common workers, they require more formal qualifications and marketable skills, they are occupied by more educated people, and they offer higher earnings, job stability and security, and decent work conditions. Evidence shows that women's job characteristics are systematically less favourable than men's (especially in rural areas, where agricultural and informal employment is massively predominant). Indeed, women are concentrated in the less rewarding sectors, industries and occupations, suggesting that there is also a form of sex-based segmentation in the labour markets.

Informal employment accounts for the vast majority of workers outside agriculture in sub-Saharan Africa, and involves generally more women than men, the former being in addition over-represented in the most precarious jobs within informal employment.

Analysing informal employment is thus of crucial importance. Not only informal employment is a major entry point for the poor to engage in industrial and service sector activities, contributing undoubtedly to poverty reduction<sup>3</sup>, but also, notwithstanding its beneficial overall effect on poverty alleviation and its contribution to economic growth, informal jobs are for the most part less secure and rewarded than formal ones. Informal workers are usually disadvantaged in the

---

<sup>3</sup> This is true especially in rural areas where movements from traditional agriculture to other sources of income are observed.



labour market and potentially incur many risks given that work is usually their sole source of living.

Most of the few studies that attempt to measure and analyse informal employment do not present a comprehensive picture taking into account its heterogeneous nature and multiple faces.

Moreover, many different concepts, definitions and measurement approaches are used in the literature to analyse informal employment, thus fueling controversies and debates around its importance, determinants and policy implications.

To resolve these ambiguities, recent efforts have been made to harmonize the measurement of informal employment at the international level. In 1993, the Fifteenth International Conference of labour Statisticians (15th ICLS) adopted an international statistical definition of the informal sector. In order to obtain an internationally agreed definition, the informal sector had to be defined in terms of characteristics of the production units (enterprises) in which the activities take place (enterprise approach), rather than in terms of the characteristics of the persons involved or their jobs (labour approach). One of the major criticisms made of the informal sector definition adopted by the 15th ICLS is that an enterprise-based definition of the informal sector is unable to capture all aspects of the increasing informalisation of employment, which has led to a rise in various forms of precarious employment, in parallel to the growth of the informal sector that can be observed in many countries.

The ILO report on decent work and the informal economy (ILO, 2002) developed a conceptual framework for employment in the informal economy that was submitted in 2003 to the 17th ICLS for discussion. The 17th ICLS examined the framework and adopted guidelines endorsing it as an international statistical standard. This conceptual framework relates the enterprise-based concept of employment in the informal sector in a coherent and consistent manner with a broader, job-based concept of informal employment, taking into account the total number of informal jobs, whether carried out in formal sector enterprises, informal sector enterprises or households.

Our analysis was able to use this latest improved and comprehensive internationally-agreed statistical definition, which is still rare in many studies due to the lack of adequate data.

In the remaining of this Chapter, we first present the countries selected for the case studies, and then we provide an overview of the main findings.

### **1.3 Countries selected**

The countries covered in this Thesis are two Anglophone East-African countries, namely Ethiopia and Tanzania. Both countries belong to the group of low-income economies, as classified by the World Bank. However, compared to Ethiopia, which is one of the poorest countries in the world, Tanzania performs better in many respects. In 2005, Ethiopia ranked 174<sup>th</sup> out of 187 countries in the Human Development Index, which is far below the 152<sup>nd</sup> position reached by Tanzania<sup>4</sup>.

Economic data indicate that Tanzania performed particularly well economically throughout the last decade. Within a stable macroeconomic environment, Tanzania experienced a period of relatively high growth and low inflation. This positive situation has been favoured by a series of macroeconomic and structural reforms initiated in the mid-1980s when the country began its transition to a market economy. However, while these measures have been fruitful, contributing in fine to the acceleration in economic growth, they have also had negative consequences such as the rapid growth of the informal sector. A third of the Tanzanian population still lives below the

---

<sup>4</sup> In 1990, the United Nations Development Programme (UNDP) introduced a new way of measuring development by combining indicators of life expectancy, educational attainment and income into a composite human development index (HDI). Concretely, this aggregate index serves as a frame of reference for both social and economic development.

national poverty line, and the country remains strongly dependent on external aid and vulnerable to international conjuncture.

Compared to Tanzania, the socio-economic situation in Ethiopia appears less favourable, although the country has long subscribed to the goals of human development and poverty eradication as guiding principles for its development strategy and programs, and has recently made significant progress in key human development indicators. Widespread poverty along with general low income levels, nutritional deficiencies, low education levels (especially among women), inadequate access to clean water and sanitation facilities, a high rate of migration, and poor access to health services remain among the main development challenges in the country (MoFED, 2005).

Ethiopia and Tanzania have nonetheless made remarkable progress towards *de jure* gender equality. Both countries have ratified most major international human rights instruments, including Convention of Elimination of all Forms of Discrimination Against Women (CEDAW), and have signed the African Union Solemn Declaration on Gender Equality in Africa (SDGEA).

Ethiopia's commitment to addressing gender disparities has strengthened over time, and a number of legislative measures have been adopted to ensure equality under the law.

In 1960, the Civil Code gave women in Ethiopia more rights than those available to their contemporaries in the United Kingdom and the United States (World Bank, 2009a). In 1993, an office for Women's Affairs (WAO) was established in the Prime Minister's Office with the task of implementing the National Policy on Ethiopian Women, which aimed to facilitate conditions necessary for: i) equality between men and women in political, social and economic life, including ownership of property, ii) access to basic services by rural women, and iii) elimination

of customary practices and prejudices against women, and enabling women to hold public office and participate in public decision-making processes.

In 1994, the Constitution of Ethiopia proclaimed the equality of men and women in terms of marital and family rights, labour rights, political rights, and in terms of participation in economic and social life.

The Ministry for Women Affairs (MoWA) was established in 2005, replacing the Office for Women's Affairs (WAO). It is responsible for ensuring the gender sensitivity of policies, identifying discriminatory practices, fostering adequate participation of women in various government bodies, undertaking studies and initiating recommendations on the protection of women's rights and ensuring their implementation.

With regard to Tanzania, it is first worth mentioning that, in the 2006 World Economic Forum Global Gender Gap report (WEF, 2006), the country was ranked number 1 out of 115 countries in terms of women's economic participation.

As part of its commitment to achieving the third of the Millennium Development Goals (MDGs), which calls explicitly for the achievement of gender equality and the empowerment of women, Tanzania has addressed gender issues in a number of areas.

Through a special amendment passed in 2000, discrimination on the basis of gender is prohibited under the constitution, which also protects the right of women to own land. The parliament has enacted a number of laws in support of women's economic and social well-being, including the Sexual Offences Act of 1998 and the two Land Acts of 1999, which established that women should be treated equally with men in terms of rights to acquire, hold, use and deal with land.

The Employment and Labour Relations Act of 2004 prohibited discrimination in the workplace on the basis of gender, required employers to promote equal opportunities, introduced maternity leave, and contained provisions protecting a mother's right to breastfeed and to be protected from

engaging in hazardous employment. Gender budgeting processes are being institutionalized in all ministries, as well as regional and local authorities, and affirmative action to include women in decision-making is being undertaken as reflected in a recent act which increases the number of women's special seats in local government councils and in the Union Parliament.

Finally, Tanzania adopted in 2000 a Women and Gender Development Policy (WGDP) to ensure gender mainstreaming in all government policies, programs and strategies, and in 2005 a National Strategy for Gender Development (NSGD) to specify how gender mainstreaming is to be implemented.

Despite significant advances in the legal framework, global comparisons show that Tanzania and, more particularly, Ethiopia lag behind other countries in achieving *de facto* gender equality. For example, according to the Gender Empowerment Measure (GEM), an aggregate index developed by UNDP to measure women's and men's capacities to actively participate in economic and political life, in 2005, Tanzania and Ethiopia ranked, respectively, 44<sup>th</sup> and 72<sup>nd</sup> out of 93 countries.

In Ethiopia, customary and religious laws remain largely in use. The society remains a male dominated one and gender differentials are recorded in all dimensions of well-being, including empowerment and effective access to productive assets such as land and credit. The economic role of women is largely constrained and undervalued, partly by traditional gender roles reinforced by high fertility and large time-requirements of some tasks (such as fetching water). Furthermore, several traditional practices have debilitating effects on women's physical and mental health with adverse consequences also on their productivity.

In Tanzania, a country which records higher *de facto* gender equality than Ethiopia, despite the positive legal framework and political context for gender equality, many barriers still hinder women's empowerment and contribution to the economy. Social and cultural norms strongly influence the ability of Tanzanian women to realize their potential, especially in the social and economic spheres. Indeed, mentalities and attitudes prevailing in the Tanzanian population, who lives predominantly in rural and traditional communities, impede women to take full advantage of their abilities and keep women subordinated to men.

The challenges faced by the country to achieve decent work for all men and women remain huge. Women shoulder the bulk of domestic activities, they have less access to wage employment than men, their careers seldom take them into positions of senior management, they receive far lower earnings than men, and they are penalized by a pervasive sex-based occupational segregation that is detrimental to them in terms of the quality of their employment (ILO, 2010).

## **1.4 Overview of main findings**

Chapter 2 provides new evidence on the gender wage gap and its deterministic factors in Ethiopia using the most recent and relevant data available, that is the 2005 Ethiopia Labour Force Survey. The analysis is conducted with a special focus on the role of education and sex-based job segregation, which are believed to be important contributing factors to the gender wage differential. This chapter goes beyond most of the thin related existing literature on Africa, in taking into account the heterogeneity of wage-employed workers and the segmented nature of the labour market, by adopting a comprehensive approach where the factors related to the gender wage gap are analysed for different points in the wage distribution, different age cohorts and different types of wage employment.

Our results show that female wages are far below male wages, at about 66 per cent in relative terms. Women's wage penalty is the highest among young people, low-paid workers and those employed in the informal private sector. The pay differential with men tends to lower as women get older and are better-paid, and when they hold formal jobs, especially in the public sector.

Disparities in education, along with other human capital characteristics such as potential work experience and training, significantly contribute to the gender wage gap but clearly not as much as job-related factors, in particular selection across industries and occupations, which appear to explain no less than half of the wage differential. Less than a quarter of women's wage penalty remains unexplained and may be attributable to discriminatory practices, gender specific preferences, cultural and other non-observable factors.

A deeper analysis of job characteristics, which account for the main source of the wage gap, reveals that selection across industries, occupations and types of job is not only driven by human capital endowment but also by gender status, which seems to be another significant factor that either picks up a form of sex-based discriminatory sorting or segmentation, and/or some gender specific preferences.

It is among youths and low-wage earners that the gender wage gap is the more driven by job attributes and the less by unobservable factors. Human capital explains a larger share of the wage differential among the elderly, low-paid workers and public sector employees. In the private sector, informal workers differ from formal workers in that their human capital and job characteristics are clearly more deterministic of the gender wage gap.

Overall, our findings suggest that progress towards gender equity in education is important to improve women's wages but not enough to close most of the gender wage gap. In particular, other interventions would be needed to compensate for the adverse impact of unobservables, such as discriminatory practices, gender specific preferences, and social and cultural norms, that directly contribute to the gender wage gap and indirectly, through job selection.

Chapter 3 casts new light on a key dimension of work-related gender inequality that has so far rarely been addressed in the case of Africa, which is the allocation of time between market and household work. Using again the 2005 Ethiopia Labour Force Survey, we examine how Ethiopian men and women differ in their allocation of time between market and household work, then we identify the gender-based division of labour, explore the gender disparities in total workload, and analyse the determinants of market and household work time across gender.

An important research question addressed in this chapter is whether and to what degree there is complementarity or substitutability between the two types of work. In other words, we are interested in understanding the connection between market and household work time allocation decisions.

The picture that emerges from our findings is striking. In Ethiopia, there is a strong gender-based division of labour which is characterized by both women (men) participating more and spending longer hours in household (market) work. However, the incidence and the average duration of market work for women are important and much higher than those of housework for men. Thus, compared to men, who generally focus only on market work, women tend to accumulate both types of work, and thereby are double-burdened. The situation is clearly exacerbated in rural areas where people have limited access to basic infrastructure. These findings are observed not only on average, but also at all points of the population distribution.

Our econometric results further suggest that there is substitutability in time allocation decisions in urban areas. While housework time appears to be quite insensitive and inelastic to market work hours, especially among men, the amount of hours devoted to productive activities seems to be, in turn, constrained and conditioned to some extent by the amount of hours allocated to household chores.

The same holds for men in rural areas, with the nuance that the adverse impact of housework hours on market work time is less strong in the countryside than in cities. By contrast, women's



time allocation decisions seem to be disconnected in rural areas, where gender disparities and traditional gender roles are the more deeply rooted. Indeed, the time rural women spend working at home does not affect and constrain, or it does but very marginally, their time spent working in the labour market, and vice versa.

The results also indicate that the effect of education on labour supply is strongly gendered. First, education has generally a negative influence on the time allocated to housework, especially among women. Second, education is usually a negative predictor of the time devoted to market work by men, and a positive predictor of that devoted by women.

The presence of adult women in the household negatively affects men's housework time, while the presence of adult men is associated with lower hours of market work and longer hours of housework performed by women. Besides, the presence of other adult women in the household relieves women of part of their housework burden.

Within households with infants, women are more heavily burdened because they have to take care of them. As infants grow up, they start taking part in many types of household activities, allowing men to be less involved in such activities.

Overall, these findings confirm that Ethiopia would benefit from pursuing and intensifying its efforts to ensure better access to education at all levels for women, and from providing better information and enforcement of the law in support of women's economic and social well-being, especially in remote areas where usually tribal law prevails, which would be helpful in changing mentalities and attitudes and in lessening the perpetuation of traditional gender roles.

The fourth and last Chapter is novel in that it is the first attempt to analyse the gender-differentiated income effect of informality in Tanzania using the latest internationally-agreed statistical definition of informal employment and accounting for population heterogeneity (not only in background characteristics but also in response to a particular treatment).

In this chapter, we assess the impact of informality on earnings and determine whether, compared to men, women are penalized by working informally. An important hypothesis that we try to test is whether individuals are constrained to work in informal jobs (the so-called exclusion hypothesis), and how this varies by gender.

To do so, we employ the 2006 Tanzania Integrated Labour Force Survey, which is the most recent and well-suited nationally representative household survey for our purpose, and we conduct treatment effect analysis of informal employment on earnings, separately for men and women in wage and self-employment, making three different assumptions for the treatment effect: (i) homogeneity, (ii) partial heterogeneity, and (iii) full heterogeneity of the population in the treatment response.

As regards self-employment, no conclusions can be drawn because most of the results are not statistically significant.

With regard to wage-employed workers, our findings indicate that they clearly differ in how they respond to a particular treatment, which is to work informally, thus supporting the full heterogeneity assumption over the strict homogeneity and partial heterogeneity hypotheses.

Allowing for full heterogeneity, we find that, in wage-employment, women face a significantly higher informal employment income penalty and are more affected by the exclusion hypothesis than men.

Overall, our results suggest that the decision to work informally is probably not the mere result of a rational choice in which people weigh the expected economic returns against the costs based solely on economic factors. Informal wage-employed workers, and in particular women, are affected by a range of constraints, such as, for instance, the burden of household responsibilities and the lack of adequate infrastructure, that constitute important obstacles which may explain why they participate in informal employment.

Accordingly, working informally is likely to be less exclusively and intentionally linked to economic gain and more the result of a constrained choice, which would argue in favour of the exclusion hypothesis. The fact that women face a significantly higher informal employment income penalty than men further suggests that women face disproportionate constraints and have little choice but to work in informal employment.

## **Chapter 2**

### **Addressing the Gender Pay Gap in Ethiopia: How Crucial is the Quest for Education Parity?<sup>5</sup>**

#### **2.1 Introduction**

As part of its overall objective to reach the Millennium Development Goals (MDGs)<sup>6</sup>, the Government of Ethiopia has made remarkable efforts towards universal primary education, gender equality and women empowerment. While there are still large gender disparities in education, Ethiopia has seen an enormous and rapid increase in enrolment in primary education that has contributed to reduce the gender imbalance (MoFED, 2005). The emphasis given to education and gender equality reflects also its instrumental importance in fostering progress towards other goals, such as raising labour compensation and supporting women's progress in the labour market. Research shows that women's earnings can influence their status and decision-making power within the family, as well as their choices about labour force participation and fertility. Women's wages are especially important for children, as they tend to spend their earnings directly on their needs (UNICEF, 1999). This raises important policy questions for Ethiopia, a country that has ratified the UN Convention on the Elimination of All Forms of Discrimination against Women. How significant is the gender pay gap? What lies behind the pay differentials between men and women? Are discrimination and other non-observable factors important and similar across the wage distribution and the types of employment? How likely will the achievement of the education MDGs translate into a significant reduction in wage disparities

---

<sup>5</sup> This Chapter draws on Kolev and Suárez Robles (2010a and 2010b).

<sup>6</sup> See <http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators%2fofficialList.htm>

across gender? In contrast with the abundant literature of the gender pay gap in developed countries, and the growing number of studies for emerging countries, fewer studies have actually attempted to address these important questions in the case of Africa (Weichselbaumer and Winter-Ebmer, 2003).

Available evidence based on survey data confirms the presence of large gender pay gaps in several African countries. Some earlier studies estimate, for instance, that the ratio of female earnings to male earnings could range from 40 per cent in Kenya (Kabubo-Mariara, 2003), to 70 per cent in Cameroun (Lachaud, 1997), 80 per cent in Botswana (Siphambe and Thokweng-Bakwena, 2001) and 90 per cent in Burkina-Faso (Lachaud, 1997). In the case of Ethiopia, Temesgen (2006) finds that in the manufacturing sector in 2002 female hourly wages stood at 73 per cent of male wages. Similarly, in a study on the size and the determinants of the gender wage gaps in three African countries, Appleton *et al.* (1999) find that in urban Ethiopia in 1990, female earnings represented on average 78 per cent of male earnings. In most of these studies that attempt to explain the extent of the gender wage gap, the unexplained term, which is likely the result of discriminatory practices, gender specific preferences, cultural and other non-observable factors, along with differences in educational endowments, account for a non-negligible share of the pay gap.

Other more recent studies on Africa using matched employer-employee data indicate that the relative importance of the unexplained component decreases when other factors such as job tenure and job characteristics are included as controlled variables in the wage equations, and that much of the wage gap correlated with education can be explained by selection across occupations and firms (Fafchamps *et al.*, 2006; Nordman and Wolff, 2008, 2009). In the case of Madagascar, using a better measure of female work experience obtained from matching a labour force survey with a biological survey also contributes to reduce substantially the size of the unexplained wage gap in the decompositions analysis (Nordman and Roubaud, 2005).

The aim of this Chapter is to cast new light on the gender pay gap in Ethiopia using the 2005 Ethiopian Labour Force Survey. A particular attention is drawn on the relative importance of education parity to mitigate the most pressing wage inequality and the role of job segregation. This Chapter complements the few available studies for Africa by adopting a comprehensive approach where the factors related to the gender pay gap in Ethiopia are analysed for different points in the wage distribution, different age cohorts and different type of wage employment. To this end, we start to estimate wage equations using two specifications and three different models, separately for men and women. We then apply decomposition procedures proposed by Neumark (1988) and Cotton (1988) to disentangle the effects on the pay gap of human capital and job characteristics from an unexplained component that captures the effect of discrimination and other non-observable factors.

The Chapter is organised as follows. Section 2 presents the data set, the concepts and some detailed summary statistics on gender disparities in employment, education and pay for different age cohorts, segments of the labour market and wage levels. The different methods chosen for estimating wage equations and decomposing gender wage gaps are explained in section 3. Section 4 presents the main results and section 5 concludes.

## **2.2 Data and concepts**

In this section, we start off by presenting the data used for the analysis of the gender wage gap. We further provide an explanation of the definitions and measures of key relevant labour market indicators. We finally present some basic descriptive statistics on employment and education broken down by gender.

### 2.2.1 Ethiopia Labour Force Survey 2005

To explain the difference in earnings by gender and to analyse the factors related to the gender pay gap in Ethiopia, we draw upon the Labour Force Survey (LFS) collected in Ethiopia by the Central Statistics Agency (CSA) in March 2005. The LFS is a nationally representative household survey containing information on a large number of individuals. It is designed to monitor the social and economic situation of the economically active population<sup>7</sup>. Out of the total 230 680 individuals who were interviewed in the LFS, 50.5 per cent were located in urban areas<sup>8</sup>. Accounting for sampling weights, this figure declines to 14.2 per cent of overall population. The individual record includes a broad range of information about age, gender, education, employment status, wage and non-wage activities, job characteristics, and earnings, and thus represents a good opportunity for our study. However, it is important to emphasize two weaknesses of this data. Poor economies like Ethiopia are generally characterized by a preponderance of seasonal activities, especially in agriculture but also in small manufacturing<sup>9</sup>. Yet, the fact that data collection took place in a short period of time makes the data particularly sensitive to seasonality issues. Another weakness of the data is related with the survey questionnaire which does not collect interesting information like ethnicity, religion and language which may potentially affect labour outcomes<sup>10</sup>.

---

<sup>7</sup> The LFS 2005 covers all parts of the country except the Gambela region (including Gambela town), and the non-sedentary population of three zones of Afar and six zones of Somali regions. For the purpose of the survey, the country was divided into three broad categories: rural areas, major urban centres and other urban centres. A stratified two-stage cluster sample design was used in the first two categories to select samples. The primary sampling units (PSUs) were enumeration areas (EAs). Households per sample EA were then selected as a second-stage sampling unit (SSU). As regards the third category, a stratified three-stage cluster sample design was adopted to select samples. PSUs were urban centres and SSUs were EAs. Households from each EA were selected at the third stage.

<sup>8</sup> Urban households are oversampled.

<sup>9</sup> According to Wodon and Beegle (2006), evidence for Malawi and other developing countries suggests the existence of labour shortages at the peak of the cropping season, and substantial underemployment for most of the year, especially in rural areas. As regards Ethiopia, Dercon and Krishnan (2000) use data from rural areas to show high levels of seasonal and year-to-year variability in consumption and poverty, with households also responding to changes in labour demand and prices.

<sup>10</sup> According to CSA (2006), these variables were not included in the LFS because the number of cases obtained in this specific sample size was not found to be sufficient in order to provide reliable information.

### 2.2.2 Definitions and measurement issues

#### *Wage employment*

The 2005 LFS contains self-classification information on productive activities such as work for payment, family gain or profit for own consumption performed in the last 7 days by individuals of age 5 and above. In this study, our measure of the labour force refers to all persons aged 15 or above either engaged in, or available to undertake, productive activities<sup>11</sup>. We further include under the label *wage employment* all individuals engaged in productive activities that worked as paid employees at least four hours in the last seven days. We also include all those who were working less than four hours or were not working the last seven days, and who were paid while on temporary leave or who had an assurance or an agreement for returning to work.

Wage employment, which represents a minor share of total employment in Ethiopia (only 8.7 per cent), is the basis for our analysis of the gender pay gap. Our coverage of wage employment is broader than what is usually found in wage studies for Ethiopia as it includes rural labour markets. Rural wage employment represents a significant share of total wage employment (37.8 per cent).

#### *Formal and informal wage employment*

Our ability to measure informality is very much driven by the quality of the survey questionnaire and it is important to clarify the concepts used in this study to define formal and informal wage

---

This is unfortunate because, for instance, ethnicity is likely to play an important role in wage determination and may shed light on gender differences in the labour market. Barr and Oduro (2002) provide evidence that the Ghanaian labour market is ethnically fractionalized and that this is leading to quite considerable earnings differentials between ethnic groups. Based on an analysis of data from a survey of owner-managed manufacturing businesses in Ethiopia, Mengistae (2001) find that an indigenous minority group, namely, the Gurage, happens to have a far higher rate of business ownership than other major ethnic groups as well as minorities, partly because Gurage-run businesses perform better.

<sup>11</sup> See the World Bank (2007) Ethiopia's report for a detailed discussion on the various concepts used by the CSA and the World Bank to classify employment. Our definition of the labour force is that of the World Bank, while the CSA uses a different age threshold (10 and above).



employment. In this Chapter, we follow the recommendation of the World Bank (2007) and use a broad concept of the informal sector. In the data, individuals who work for a wage or a salary are asked to describe the employment status of their main occupation. Those who report working in the public sector as government or parastatal employees, as well as NGO employees and other employees working in private organizations which have ten or more employees, or which have a license or a bank account, are classified as *formally wage employed*.

In contrast, *informal wage employment* includes paid employees who are domestics or who work in a private organization which has less than ten employees, is not licensed and has not a bank account. It also includes employees for which this information is missing and who are only paid in kind or doing casual work. The latter are in fact very likely to be located in the unregulated sector.

Following this classification, we further decompose wage employment in three components: public formal wage employment (government and parastatal employees), formal private wage employment (employees in formal private organizations and NGOs) and informal private wage employment (employees in informal private organizations and domestic employees).

### ***Earnings***

Unlike the 2001 survey, the 2005 LFS provides a good opportunity to analyse the gender pay gap as it provides information on the amount paid to wage employees in their main occupation during the last pay period and the number of times they were paid during the last month. To account for the impact of the duration of work on wages, we calculate hourly earnings from the main occupation for each worker in wage employment by dividing the monthly earnings by the monthly hours of work in the main occupation. The latter is indirectly calculated by subtracting the number of hours worked on additional activities from the total number of hours worked at all

jobs in the last seven days, and multiplying the result by four<sup>12</sup>. The use of earnings as a proxy for the returns to work is not exempt of problems, however. As earnings are available exclusively for the wage employed and from their main occupation only, this leaves aside the possibility to analyse the returns from self-employment. Moreover, it does not allow taking into account the returns of secondary employment<sup>13</sup>. Finally, non-wage benefits may be important in some cases (in particular, the government and the parastatal might offer additional benefits in terms of pension benefits or job security) and since they are not imputed this may underestimate the true level of earnings. Notwithstanding these issues, earnings data remain essential to understand the gender pay gap in Ethiopia.

### **2.2.3 Descriptive statistics**

#### ***Gender disparities in the labour force and in employment status***

Basic labour market indicators are reported in Table 2.1 separately for males and females. According to all these indicators, the situation of women appears less favourable than that of men. The participation rate and the employment ratio are lower for women, while female unemployment is similar to male unemployment. In addition, a higher proportion of the male population is in wage employment.

---

<sup>12</sup> Hours of work include overtime and exclude lunch and journey time. Here, we multiply weekly hours of work by 4 to obtain monthly hours of work, and not by 4.2 or 4.3 as it is usually done, because the last month preceding the interview was February 2005 and included only 28 days.

<sup>13</sup> This should not be a serious problem, however, since the LFS 2005 reveals that less than 9 per cent of Ethiopian wage employees hold multiple jobs. Moreover, according to data from the Addis Labour Market Survey (ALMS), wages from secondary jobs do not appear to be an important element of overall earnings (World Bank, 2007).

**Table 2.1: Selected labour market indicators**

	<b>All (%)</b>	<b>Male (%)</b>	<b>Female (%)</b>
Activity rate	82	90	75
Employment ratio	80	88	72
Wage employment ratio	7	9	5
Unemployment rate	3	2	3
Inactivity rate	18	10	25

*Source:* LFS, 2005.

*Note:* Individuals aged 15 and above. Weighted data.

The characteristics of wage employment are shown in Table 2.2. There are large gender variations in the nature and the terms of wage employment. For men, public formal wage employment and private formal wage employment constitute altogether the biggest share of the wage employed. Only 15 per cent of wage employed men are in informal private jobs. For women, however, the proportion of the wage employed in private informal jobs represents the second most frequent form of wage employment (32 per cent) after public employment (40 per cent). In relative terms, women are more likely than men to work in informal jobs, as temporary or casual employees, and less likely to work in permanent or contract employment, suggesting that the conditions of work among wage employed women are less favourable than for men.

**Table 2.2: Nature and terms of wage employment in total wage employment**

	<b>All (%)</b>	<b>Male (%)</b>	<b>Female (%)</b>
Public formal wage employment	43	44	40
Private formal wage employment	36	41	28
Private informal wage employment	21	15	32
Permanent employment	37	40	32
Temporary employment	42	39	48
Contract employment	10	12	8
Casual or other employment	11	9	12

*Source:* LFS, 2005.

*Note:* Individuals aged 15 and above. Weighted data.

### *The unadjusted gender pay gap*

Table 2.3 provides information on the crude gender pay gap for different wage levels, age cohorts and segments of the labour market. On average, female wages represent about 66 per cent of male wages which is remarkably low, even from an African perspective. The wage relative disadvantage of women is more pronounced for younger women. As women get older, the pay differential with men tends to lower. Table 2.3 further indicates that there are large variations in the gender pay gap at different points in the wage distribution. While female wages represent nearly 70 per cent of male wages at the top quartile, they account only for 55 per cent of male wages at the median and 49 per cent at the bottom quartile. These findings are at odds with the results for developed countries which show that, in line with the glass ceiling hypothesis, the gender wage gap tends to be more pronounced at the upper tail of the wage distribution. The gender disadvantage in terms of wages appears to be more pronounced in the formal private sector than in the public sector. In the latter female wages do not exceed 80 per cent of male wages while in the former female wages barely reach 74 per cent of male wages. In the informal private sector the gender wage gap is very high: female wages account for just slightly more than one-third of male wages.

**Table 2.3: Unadjusted gender hourly wage gaps**

15+	Age groups (%)			Wage percentiles (%)			Sector of wage employment (%)		
	15-24	25-34	35+	P25	P50	P75	Public	Formal private	Informal private
66	64.9	72	80.1	48.5	54.7	69.9	79.5	73.7	34.6

*Source:* Ethiopia LFS, 2005.

*Note:* Female earnings in terms of male earnings. Weighted data.

Besides looking at the average hourly earnings for men and women, it is also interesting to compare the distribution of hourly wages for men and women. This information is depicted through the generalized Lorenz curves represented in Figures 2.1 and 2.2 in Appendix A, for the

different age cohorts and the different types of wage employment, respectively. These curves show on the horizontal axis the cumulative proportion of the population and on the vertical axis the cumulative average hourly earnings. It is worth noticing first that in all age cohorts and all types of wage employment, the generalized Lorenz curves of men always lie above those of women, indicating that women earn less than men not only on average, but also at all points of the distribution of the population. Second, while the gender gaps in the average accumulated hourly earnings are quite insignificant at the first deciles of the distribution, they continuously increase as we move to higher shares of the population. Third, the generalized Lorenz curves points to large gender disparities in the distribution of hourly wages in informal private wage employment.

### ***Gender disparities in education characteristics among the wage employed***

Table 2.4 presents the educational characteristics of the wage employed separately for men and women. It reveals a strong disadvantage among women. The incidence of illiteracy is 14 per cent higher among wage employed women. More wage employed males than females are educated<sup>14</sup>. It is also worth noticing that the gender educational gap in favour of wage employed men, which is observed in all age groups, tends to decrease for younger cohorts. For both men and women, workers with more education tend to be disproportionately concentrated in better paid jobs. Public wage employment is predominantly composed of workers with general or beyond general education. In private wage employment, there is an important dichotomy between the formal and the informal sector. While formal private wage employment includes workers from all levels of education (from illiteracy to beyond general education), informal private wage employment is

---

<sup>14</sup> General education includes grades 9-12 in the new system (general secondary education, grades 9-10; preparatory secondary education, grades 11-12) and grades 9-12 in the old system. Beyond general education includes new vocational education (grades 11-12), certificate, diploma (grades 11-13), degree completed or not and above degree. Primary education includes primary education in the new system (basic education cycle, grades 1-4; general primary cycle, grades 5-8), non-formal education and literacy campaign.

predominantly composed of uneducated or low-educated workers, especially among women. Yet, it is interesting to note that in private formal employment, there are more highly educated women than men, indicating that access to private formal jobs may be more competitive for women.

**Table 2.4: Levels of education among the wage employed by gender**

	Illiterate (%)		Primary education (%)		General education (%)		Beyond general education (%)	
	Men	Women	Men	Women	Men	Women	Men	Women
All wage employed (15+)	27	41	32	23	21	18	20	18
<i>Age cohorts</i>								
15-24	36	43	32	28	16	14	16	15
25-34	20	32	32	20	27	27	21	21
35+	27	49	32	18	19	15	22	18
<i>Wage quartiles</i>								
Q1	57	67	37	29	6	3	0	1
Q2	29	51	46	34	22	13	3	2
Q3	13	32	33	25	34	31	20	12
Q4	7	7	11	7	21	27	61	59
<i>Wage employment</i>								
Public formal wage employment	13	26	23	13	26	27	38	34
Private formal wage employment	36	39	37	27	19	22	8	12
Private informal wage employment	46	62	44	33	9	4	1	1

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above. Weighted data.

### ***Gender disparities across sectors of activity and occupations***

A form of gender segmentation across occupations and industries is also evident in the data. Overall, almost all wage employees in Ethiopia are engaged in tertiary sectors activities, few work in agriculture and even less work in the manufacturing sector. Women in wage employment are more involved in tertiary sector activities and less in agriculture than men. More than one

woman out of two in wage employment is a low-skilled blue collar, working in elementary occupations. A large share of wage employed men is also working in elementary occupations. However, men are more likely than women to be in high-skilled occupations working, for instance, as professionals, technicians and associate professionals, or managers.

While a form of gender selection across sector of activities and occupations is observed among wage employees in Ethiopia, this also varies by employment types. What is worth noticing is that in informal private wage employment, almost all women in that sector work in tertiary sector activities, while more than a quarter of men work in agriculture and more than ten per cent work in the manufacturing sector. Nearly three quarters of women occupying informal jobs work in private households with employed persons (tertiary sector 6), while men occupying informal jobs in the tertiary sector are mainly involved in construction (tertiary sector 1) and other community, social and personal service activities (tertiary sector 5). The vast majority of informal private wage employees, especially among women, are blue collars, most of them being low-skilled. A much higher proportion of women than men in informal private wage employment works in elementary occupations. In formal private wage employment, more women are working in agriculture and in the manufacturing sector, but the overwhelming majority is still working in the tertiary sector. Like in the informal sector, most of the wage employees in the formal private sector are low-skilled blue collars, and the gender disparities in terms of occupation are important but relatively smaller than in the informal sector. Lastly, a majority of workers in public wage employment are white collars. Although gender differences in terms of occupation in that sector are less marked than in the formal and informal private sectors, they exist and they are significant (Table 2.5).

All in all, the descriptive statistics point to large gender disparities in education characteristics of the wage employed combined with strong gender segmentation across occupations and sectors of activity.

**Table 2.5: Industries and occupations by wage employment sector and gender**

	All wage employment (%)		Public sector (%)		Formal private sector (%)		Informal private sector (%)	
	Men	Women	Men	Women	Men	Women	Men	Women
<i>Occupations</i>								
High-skilled white collars	25	17	47	35	9	9	4	1
Low-skilled white collars	12	20	14	23	10	28	12	10
High-skilled blue collars	19	10	8	6	27	11	29	12
Low-skilled blue collars	44	53	31	36	54	52	56	77
<i>Industries</i>								
Primary sector activity	18	6	6	6	28	10	26	3
Manufacturing	10	9	4	6	15	14	13	8
Tertiary sector 1	14	9	10	8	18	18	16	2
Tertiary sector 2	7	11	1	3	12	29	11	7
Tertiary sector 3	7	3	8	6	7	2	7	0
Tertiary sector 4	18	16	40	37	1	2	0	0
Tertiary sector 5	21	20	30	34	12	15	17	6
Tertiary sector 6	1	23	0	0	0	0	10	73

*Source:* Ethiopia LFS, 2005

*Note:* Individuals aged 15 and above. Weighted data. See footnote of Table 2.6 in Appendix B for definitions of the four occupational groups and the six tertiary sectors.

The extent to which such disparities can explain the observed gender pay gap will be the focus of the remaining of this Chapter.

## 2.3 Methodology

In this section, we first introduce the three estimation methods used to analyse the determinants of wages, for the overall sample of the wage employed and for different age cohorts, for different wage employment sectors and for different wage levels, respectively. We then present the decomposition techniques of the gender wage gap implemented for our wage equations<sup>15</sup>.

<sup>15</sup> For a comprehensive review of the various methods related to the analysis of the gender wage gap, see Beblo *et al.* (2003).



### 2.3.1 Estimation of wage equations

In this study, we use the following augmented version of the traditional human capital semi-logarithmic earnings equation developed by Mincer (1974):

$$y_i = \beta_i X_i + \varepsilon_i \quad (1)$$

Where  $y$  is the log hourly earnings in the main occupation of individual  $i$ , which are observed only for paid employees,  $X$  is a vector of explanatory variables including human capital variables (dummies for three levels of educational attainment, potential work experience which is defined by age minus years of schooling minus six, and its square to take into account its possible decreasing returns, and a dummy for training), another individual variable (a dummy for marital status), job characteristics (dummies for various sectors of activity, sectors of employment, terms of employment and occupations), urban and regional dummies, and none, one or several selectivity correction terms depending on the estimation procedure and the selection model used.  $\varepsilon$  is the error term with an expected value of zero and  $\beta$  is a set of coefficients to be estimated (including a constant)<sup>16</sup>.

#### ***Heckman's two-step estimation procedure***

Having specified the wage equation, we now want to estimate the determinants of earnings separately by gender on the overall sample of the wage employed (15+), and for three different age cohorts (15-24, 25-34 and 35+). To do so, we employ Heckman's two-step estimation procedure. This method allows correcting for a possible sample selection bias, which may arise

---

<sup>16</sup> Disability status usually plays an important role in determining wages. Although this information is *a priori* available in the survey based on self-reporting, we do not introduce this variable in the regression analysis because the number of wage employees who report some kind of disability in our sample is very small. Self-reported measures of disability are also frequently attached with large measurement errors and cannot be easily interpreted in econometric analysis.

because our sample of wage employed may not be a random sample and may have specific characteristics.

According to Heckman (1979), selection bias can be thought as a form of an omitted variable bias. The omitted variables problem, which indeed Heckman has shown to be equivalent to the use of non-random samples, can be solved by including in the log hourly earnings equation (1) a sample selection term constructed from an equation modelling the probability of being in wage employment.

In the first stage, maximum likelihood estimates of the probit model used to estimate the probability of selecting wage employment (the selection equation) are separately performed for women and men. A selection correction term (the inverse of the Mills' ratio) is then included into the log hourly earnings equations, for women and men, respectively. The OLS estimates of these augmented earnings equations are unbiased and consistent.

Formally, let  $Z_i^*$  be a latent variable which is the underlying propensity of being wage employed:

$$Z_i^* = \gamma_i W_i + u_i \quad (2a)$$

Log hourly earnings (the dependent variable of the outcome equation) are observed only for  $Z_i^*$  bigger than zero. Since  $Z_i^*$  is unobserved, we use an indicator variable  $Z_i$  which is equal to one if the individual  $i$  is a paid employee and zero otherwise.  $Z_i$  is then the dependent variable of the following selection equation:

$$\Pr(Z_i^* > 0/W_i) = \Pr(Z_i = 1/W_i) = \Phi(\gamma_i W_i) \quad (2b)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution,  $Z$  is a binary outcome variable,  $W$  is a vector of regressors,  $u$  is the disturbance term with an expected value of

zero, and  $\gamma$  is the coefficient to be estimated for each explanatory variable in vector  $W$  (including a constant term). The set of explanatory variables included in  $X$  (equation 1) are also included in  $W$ , except job variables.

Rewriting the outcome equation in terms of expectations, we obtain:

$$E(y_i / X, Z_i^* > 0) = \beta_i X_i + \alpha_i + E(\varepsilon_i / u_i > -\gamma_i W_i). \quad (3)$$

If the sample of wage employees is non-randomly selected,  $\varepsilon_i$  and  $u_i$  are dependent and the conditional mean of  $\varepsilon_i$  is not equal to zero. OLS estimates of the coefficient of the explanatory variables of the log hourly earnings equation are then biased.

Since  $(\varepsilon_i, u_i)$  has a normal distribution, by using well known results for conditional distribution for bivariate normal case, we can get:

$$E(\varepsilon_i / u_i > -\gamma_i W_i) = \frac{\sigma_{\varepsilon u}}{\sigma_u} \lambda_i \quad (4)$$

Where:

$$\lambda_i = \frac{\phi(H_i)}{1 - \Phi(H_i)}, \quad H_i = -\frac{\gamma_i W_i}{\sigma_u}.$$

The inverse of the Mills' ratio  $\lambda_i$  is a monotone decreasing function of the probability  $1 - \Phi(H_i)$  that a working-age population observation with characteristics  $W_i$  is selected into the observed sample of wage employees.  $H_i$  is the negative of the predicted value from equation (2b),  $\phi$  and  $\Phi$  are respectively the density and the distribution function for a standard normal

variable, and  $\sigma_{\varepsilon u} / \sigma_u$  is the ratio of the covariance between the errors in the selection and the outcome equations, and the standard deviation of the errors in the selection equation.

If the inverse of the Mill's ratio is zero then it means that there is no sample selection bias and the OLS estimates of the coefficients of the independent variable of log hourly earnings function (equation 1) are unbiased.

If  $\sigma_{\varepsilon u} / \sigma_u$  is positive then there is one or several unmeasured variables that both increase the probability of selection into wage employment and the log hourly earnings. If it is negative then there is one or several unobserved variables that increase the probability of selection into wage employment but decrease the log hourly earnings.

Finally, instead of using the earnings function defined in equation (1) we estimate for both men and women the following equation on the overall sample of the wage employed and for three different age cohorts of wage employees:

$$y_i = \beta_i X_i + \frac{\sigma_{\varepsilon u}}{\sigma_u} \lambda_i + v_i \quad (5)$$

Where:

$$E(v_i / X, \lambda, u_i > -\gamma_i W_i) = 0 .$$

The OLS estimates of the log hourly earnings equations with the additional selection term (equation 5) are now unbiased and consistent even in the presence of a sample selection bias.

### ***Bourguignon-Fournier-Gurgand two-step estimation procedure***

In order to take into account the fact that the labour market may be segmented and that the determinants of wages may vary across different types of sectors, we further estimate earnings equations for three different types of wage employment: public formal, formal private and informal private. Heckman's two-step estimation procedure cannot be used this time since it does not allow dealing with another possible sample selection bias which arises when the labour market is segmented. In the data, men are indeed over-represented in public and formal private wage employment while women are more concentrated in informal private jobs, suggesting that the decision to work in a particular sector may not be determined randomly.

To deal with the possible sample selection bias related with the decision to work in different types of wage employment, we use a variant of the Dubin and McFadden (1984) two-step estimation procedure proposed by Bourguignon *et al.* (2007)<sup>17</sup>. The Bourguignon-Fournier-Gurgand (BFG) procedure consists of two phases. In the first stage, maximum likelihood estimates of multinomial logit sectoral choice models (non-participation in wage employment is the base category) are separately performed for men and women. Then we include in each earnings equation four selection correction terms (corresponding to the four alternatives) computed from the estimated

---

<sup>17</sup> Bourguignon *et al.* (2007) show that the widely used Lee's (1983) two-step approach as well as a more recent method proposed by Dahl (2002) tend to perform very poorly in Monte-Carlo experiments as compared with the Dubin-McFadden method. In addition, they argue that Lee's method and Dahl's method are adapted to very small samples and would not be recommended in most contemporary empirical contexts. Accordingly, given that we use large samples in our study, the Dubin-McFadden method seems to be the most appropriate one. However, Bourguignon *et al.* (2007) show that the performance of this method depends on the constraint that all correlation coefficients between the disturbance terms from each earnings equation and the disturbance terms from each multinomial logit equation sum-up zero. It greatly improves the correction performance when the latter condition holds, but the method appears sensitive to it when it is violated. They thus propose a variant of the Dubin-McFadden method, relaxing this restrictive assumption and allowing for normal error terms in the outcome equation. They show that, in Monte-Carlo experiments, this alternative method is generally less robust than the original one but it performs better when the Dubin-McFadden's assumption is violated and provides very similar results when it holds.

parameters of the multinomial logit sectoral choice models<sup>18</sup>. The augmented segment log hourly earnings equations are then estimated by OLS.

Formally, we consider the following model:

$$y_{ij} = \beta_{ij} X_i + \varepsilon_{ij} \quad (6)$$

$$Z_{ij}^* = \gamma_{ij} W_i + u_{ij} \quad (7)$$

Where  $j$  is a categorical variable, taking values in the range 0 to 3, which describes the choice of a working-age individual  $i$  among the four alternatives based on “utilities”  $Z_{ij}^*$ . The log hourly earnings are not observed for  $j=0$ , that is for all working-age individuals who choose not to participate in wage employment.

The outcome variable  $y_{il}$ , for example, is observed only if the alternative 1 is chosen by the working-age individual  $i$  which happens when:

$$Z_{i1}^* > \max_{j \neq 1} (Z_{ij}^*). \quad (8)$$

We define:

$$\mu_{i1} = \max_{j \neq 1} (Z_{ij}^* - Z_{i1}^*) = \max_{j \neq 1} (\gamma_{ij} W_i + u_{ij} - \gamma_{i1} W_i - u_{i1}). \quad (9)$$

---

<sup>18</sup> According to Bourguignon *et al.* (2007), selection bias correction based on the multinomial logit model provides fairly good correction for the outcome equation even when the IIA hypothesis is violated.

Under condition 8, equation 9 is equivalent to  $\mu_{i1} < 0$ . According to the independence of irrelevant alternatives hypothesis, we assume that the  $(u_j)$ 's are independent and identically Gumbel distributed. Their cumulative and density functions are respectively  $G(u) = \exp(-e^{-u})$  and  $g(u) = \exp(-u - e^{-u})$ . As shown by McFadden (1973), this specification leads to the multinomial logit model with:

$$P(\mu_{i1} < 0/W_i) = \frac{\exp(\gamma_{i1}W_i)}{\sum_{j=0}^3 (\gamma_{ij}W_i)} \quad (10)$$

Based on this expression, consistent maximum likelihood estimates of  $(\gamma_j)$ 's can be easily obtained.

If the sample of the working-age population who has chosen the alternative 1 is non-randomly selected, then the disturbance term of the outcome equation (equation 6)  $\varepsilon_1$  is not independent of all  $(u_j)$ 's. This would introduce some correlation between the explanatory variables and the disturbance term in the outcome equation. Because of this, OLS estimates of  $\beta_{i1}$  would not be consistent.

With the BFG two-step estimation procedure, the log hourly earnings function conditional on choosing alternative 1 is now defined as follows:

$$y_{i1} = \beta_{i1}X_i + \sigma_\varepsilon \left( \sum_{j=0}^3 r_j^* \lambda_{ij} \right) + v_{i1} \quad (11)$$

Where  $r_j^*$  is the correlation between  $\varepsilon_1$  and  $u_j^*$ , with  $u_j^* = \Phi^{-1}(G(u_j))$  and where the selection correction terms are defined as follows:

$$\begin{cases} \lambda_{i1} = m(P_{i1}) \\ \lambda_{ij} = m(P_{ij}) \frac{P_{ij}}{(P_{ij} - 1)}, \forall j \neq 1 \end{cases} \quad (12)$$

In equation 12,  $P_{ij}$  is the probability that category  $j$  is chosen by the working-age individual  $i$  (as defined in equation 10 for alternative 1), and  $m(P_{ij})$ , which is computed numerically after the multinomial logit estimation, is the expected value of  $u_j^*$ , conditional on alternative 1 being made. In equation 11, the OLS estimates of the log hourly earnings function are unbiased and consistent even in the presence of sample selection bias<sup>19</sup>.

In our Heckman and BFG two-step regressions, the same independent variables of the outcome equation are used in the first stage (except job variables), for the probit and multinomial logit maximum likelihood estimation of the selection equation. Accordingly, model identification is based solely upon the non-linearity in the functional form. Because the selection equation is non-linear, it potentially allows the higher powers of regressors to affect the selection variable. In the linear outcome equation the higher powers do not appear. Therefore, the non-linearity of the selection regression automatically generates exclusion restrictions. That is, it allows for independent source of variation in the probability of a positive outcome.

However, if the non-linearity implied by the probit or the multinomial logit model is slight, then the identification will be fragile. For this reason, it is common in applied work to look for exclusion restrictions by seeking one or several variables that can generate non-trivial variation in

---

<sup>19</sup> As recommended by Bourguignon *et al.* (2007), we use the bootstrap method to obtain consistent standard errors.



the selection variable but do not affect the outcome variable directly. Because it is often hard to come up with a variable with these characteristics, a strong justification is necessary for imposing the exclusion restriction.

In our case, we test several variables and combinations without success. Wald tests results of significance in the first stage and insignificance in the second stage in the regression analysis for males and females in overall wage employment, and in different age cohorts and wage employment sectors, were not convincing<sup>20</sup>. This is not very much surprising since the LFS 2005 was obtained from a short questionnaire, and thus offers little opportunity to find valid instruments.

Heckman and BFG estimation procedures are interesting in that they are less restrictive than OLS which does not differentiate the participation decision and the wage determination. However, results obtained with these methods are sensitive to the specification of the selection equation, the choice and the number (or, like in our case, the lack) of identifying instruments. We therefore decide to perform OLS regressions in addition to Heckman and BFG regressions.

### ***Quantile regression analysis***

Recent studies have suggested that the gender earnings differential may not be constant along the earnings distribution. As stated earlier in this Chapter, in developed countries the evidence shows that in line with the glass ceiling hypothesis, the gender pay gap tends to be more pronounced at the upper tail of the earnings distribution. In developing countries and in Africa in particular, the evidence remains very scarce. Few studies analyse the gender earnings gap and even less studies use distributional approaches (Montenegro, 2001; Nordman and Wolff, 2008).

---

<sup>20</sup> It is worth mentioning here that assessing whether an instrument is a good identifier in a two-stage regression is not straightforward and no unquestionable test has been proposed so far in the literature.

We use the quantile regression method (QRM) to estimate earnings functions at three different percentiles of the earnings distribution: the first quartile, the median and the third quartile. QRM has the potential of generating different responses of the dependent variable to changes in the regressors at different percentiles of the conditional distribution of the dependent variable. In addition, contrary to OLS method, the QRM is robust to the presence of outliers.

However, the use of the QRM is not exempt of problem. Correcting for possible sample selection bias is particularly cumbersome when using QRM and there is currently no consensus regarding the most appropriate correction procedure for selectivity bias in quantile regression models. The traditional Heckman selectivity correction procedure cannot be used when using the QRM (Buchinsky, 2001), and the more recent methods which attempt to correct for selectivity bias in quantile regression models, like the one proposed by Buchinsky (2001), who uses the work of Newey (1999), are not exempt from criticism (Montenegro, 2001; Hyder and Reilly, 2005). We then adopt a conservative approach and decide not to control for a possible sample selection bias.

We consider the log hourly earnings function defined in equation (1). The  $\theta^{\text{th}}$  quantile of the conditional distribution of  $y_i$  given  $X_i$  is defined as:

$$Q_{\theta}(y_i/X_i) = \beta_{\theta} X_i. \quad (13)$$

When the proportion of outliers is relatively large, the traditional OLS estimation method is not robust. This has led researchers to look for alternative methods that does not minimize the sum of the squared errors but minimizes instead the sum of the absolute value of the errors, such as the Least Absolute Deviation (LAD) technique. The LAD implies the minimization of:

$$\psi_{LAD} = \sum_{i=1}^n |y_i - \hat{y}_i| = \sum_{i=1}^n |y_i - \beta_i X_i| = \sum_{i=1}^n (y_i - \beta_i X_i) \text{sign}(y_i - \beta_i X_i) \quad (14)$$

Where  $\text{sign}(R)$  is 1 if  $R$  is non-negative and -1 if  $R$  is negative.

The  $\theta^{\text{th}}$  conditional quantile estimation of the log hourly earnings equation using the QRM is achieved by weighting differently positive and negative errors in equation (14) and by minimizing the new expression:

$$\psi_{QRM,\theta} = \theta \sum_{i|y_i \geq \beta_\theta X_i - \alpha_\theta}^n |y_i - \beta_\theta X_i| + (1-\theta) \sum_{i|y_i < \beta_\theta X_i - \alpha_\theta}^n |y_i - \beta_\theta X_i|. \quad (15)$$

The QRM fits a hyperplane among the observations so that a certain proportion  $\theta$  of the observations will be below of the hyperplane and the rest above it. What is important in the QRM is the number of errors that are positive or negative but not their magnitude. The vector  $\hat{\beta}_\theta$  that minimizes equation (15) is then invariant to the presence of outliers.

For the purpose of our study, we perform three simultaneous QRM estimations (first quartile, median and third quartile) of the log hourly earnings equation separately for men and women who are wage employed. Bootstrapped standard errors are computed in order to have consistent estimates of the variance-covariance matrices.

### 2.3.2 Decomposition of the gender wage gap

#### *Neumark and Cotton decomposition procedures*

The Oaxaca-Blinder decomposition procedure is commonly used in the literature on the gender pay gap (Blinder, 1973; Oaxaca, 1973). It implies the following assumption: in the absence of discrimination, the men's earnings structure would also apply to women and the women's earnings structure would also apply to men. Discrimination takes the form of men (women) receiving more (less) than what a non-discriminating labour market would have awarded them. There are inherent interpretation problems in the Oaxaca-Blinder decomposition procedure. The explained component of the gender wage gap is based on estimates of what a woman would receive if she had faced the male wage structure or what a man would receive if he had faced the female wage structure, and does not take into account a wage structure that would prevail in the absence of discrimination. Furthermore, the decomposition can be quite sensitive to the earnings structure used (whether male's or female's): none is preferable to the other *a priori*. This is why estimates of both specifications are usually calculated and used to establish a range within which the true values of the component lie.

To deal with these interpretation problems, Neumark (1988) proposed an augmented version of the Oaxaca-Blinder decomposition procedure, which is also widely used in the economics literature. According to Neumark, employers may practice nepotism toward men or discrimination against women. Under nepotism, women are paid the competitive wage but men are overpaid. The coefficients from the women's earnings functions then provide an estimate of the non-discriminatory wage structure. Under discrimination, men are paid the competitive wage but women are underpaid. The coefficients from the men's earnings functions then provide an estimate of the non-discriminatory wage structure. In reality, employers may practice both nepotism and discrimination. With the restriction that employers only care about the proportion of

men and women employed (employers preferences are homogeneous of degree zero), Neumark shows that the non-discriminatory wage structure  $\beta^*$ , which is a weighted average of the male and female wage structures, can be estimated from an earnings function estimated over the pooled sample (that is, including both men and women).

Contrary to the Oaxaca-Blinder decomposition procedure, Neumark's method assumes that similar wage structures exist for both men and women in the absence of discrimination instead of assuming varying wage structures. This procedure thus overcomes the index number problem. Furthermore, in the Neumark decomposition procedure, the unexplained component which is usually called the discrimination or treatment component is decomposed into two components that reflect, respectively, the advantage of the favoured group and the disadvantage of the discriminated group compared with the condition that would have prevailed in the absence of discrimination.

The Neumark decomposition procedure is precisely defined as follows:

$$\ln(\bar{Y}_m) - \ln(\bar{Y}_f) = \hat{\beta}^* (\bar{X}_m - \bar{X}_f) + \bar{X}_m (\hat{\beta}_m - \hat{\beta}^*) + \bar{X}_f (\hat{\beta}^* - \hat{\beta}_f) \quad (16)$$

The first component in the right hand side represents the part of the gender average earnings gap attributable to differences in characteristics evaluated at the hypothetical market that would prevail in a non-discrimination case. The second and the third component constitute the treatment or discrimination component and represent, respectively, the amount by which men's characteristics are over-valuated (men's treatment advantage) and the amount by which women's characteristics are under-valuated (women's treatment disadvantage) in the labour market.

While the improvement proposed by Neumark's decomposition is attractive, it is not exempt from criticisms. Without evidence that employers care only about the proportion of men and women employed, it is not clear that the pooled coefficient is a good estimator of the non-discriminatory wage structure (Appleton *et al.*, 1999). Other approaches have also been proposed in the literature to estimate the non-discriminatory wage structure. Reimers (1983) implements a methodology that is equivalent to assigning identical weights to men and women wage structures. A better approach, suggested by Cotton (1988), is to compute the non-discriminatory wage structure by weighting the men and women wage structures by the respective proportions of men and women in the sample.

In our decomposition analysis, we then rely on the decompositions proposed by Neumark and Cotton in order to check the sensitivity of our results to the alternative hypothesis made regarding the choice of the non-discriminatory wage structure<sup>21</sup>. These decomposition procedures are applied to decompose the gender wage gap for the overall sample of wage employed, for three cohorts of wage earners (15-24, 25-34 and 35+), for three types of wage employed (public formal, formal private and informal private) and for wage earners located at three conditional percentiles of the log hourly earnings (first quartile, median and third quartile).

### ***Treatment of the sample selection correction***

There is no obvious way to handle the selectivity bias correction within the decomposition of the raw gender gap. Out of our three wage regression models, two methods (Heckman and BFG two-step estimation models) correct for a possible sample selection bias. As a result, selection correction terms appear in the corresponding wage decompositions and these are generally treated in two different manners. A first set of studies treat the correction terms as a separate component

---

<sup>21</sup> Few studies focus on gender earnings disparities issues in Ethiopia. The choice to perform Cotton decompositions is also motivated by the fact that one of the most recent studies on this topic (Temesgen, 2006) uses this technique.

in the decomposition and dissociate the wage gap into an explained component, an unexplained (or discriminatory) component and a selection effect. A second set of studies subtract the selection effect from the observed wage gap in order to obtain a wage differential that is corrected for sample selection and that can be decomposed in two components only (an explained and a discriminatory component). The latter wage gap is then often interpreted as the differential in potential or offered wages, as opposed to observed wages. In this study, we focus on the observed wage gap and consider the impact of the selectivity correction term as a third component of the decomposition.

## **2.4 Results**

### **2.4.1 Estimations of the wage equations**

We estimate log earnings functions separately for men and women (i) in overall wage employment (ii) for three different age cohorts using Heckman's two-step estimation procedure, (iii) at three different percentiles of the earnings distribution (first quartile, median and third quartile) using the QRM and (iv) for three different types of wage employment (public, formal and informal private) with the BFG method; results are reported in Tables 2.7 to 2.10 in Appendix B, respectively. With the exception of the QRM, we focus the discussion on the wage estimates corrected for sample selectivity, which is also the most common approach used in recent studies. However, we note that the results from the two stage analysis are usually sensitive to changes in specifications and modelling, as well as in the presence, number and choice of identifying variables. Accordingly, we also estimate wage equations without correcting for selectivity and report the results in tables in Appendix B. These OLS estimates are used to

perform new gender earnings gap decompositions. Summary statistics of the various variables used in these econometric analyses are reported in Table 2.6 in Appendix B<sup>22</sup>.

The traditional Mincerian wage equation includes measures of schooling, work experience, and some other human capital factors that may affect earnings such as training. One shortcoming of the 2005 LFS is that it is not possible to account for the workers' actual experience in the labour market. The use of potential experience as a proxy is probably a good approximation of the true experience for men who usually have a higher labour force attachment, but may lead to overestimate the amount of experience for women who participate discontinuously in the labour market, for instance, because they are involved in childcare.

For each earnings function, we use two specifications. The first one is an augmented version of the traditional Mincerian wage equation and includes variables related to job characteristics (occupations, sector of activity) in addition to human capital and other individual characteristics and location variables. The second specification follows the traditional approach by excluding job characteristics. When using the first specification the impact of education on wages is likely to be underestimated since job characteristics are partly determined by education and therefore captures part of the effect of education on earnings. However, the inclusion of job characteristics is increasingly recognised as essential. They are usually important determinants of wages and job selection can be an important source of the gender wage gap. To ease the presentation, we report only the results of the wage equations accounting for job characteristics in tables in Appendix B<sup>23</sup>.

For both men and women, we note that the significance of the selection variables remains sensitive to changes in model specification. As a result, the value added of using Heckman and

---

<sup>22</sup> Sampling weights are neither used in the regression models nor in the decomposition procedures.

<sup>23</sup> Results of the wage equations not accounting for job characteristics are not reported in order to save space. They are available upon request.



BFG sample selectivity corrected earnings equations instead of OLS estimation is not obvious and in what follows, we prefer to focus our discussion on the OLS estimates<sup>24</sup>.

The results displayed in Appendix B show that the private returns to education are large and highly significant in all wage equations, and are systematically larger for higher levels of education. Moreover, in most specifications, education has clearly a greater impact on earnings for women, with the exception of the public sector, the first quartile of the wage distribution, and among youth where the returns to education tend to be greater for men than for women. Although not directly comparable, these findings contrast somehow with the results of other studies for urban Ethiopia which find that the private returns to primary education are often insignificant (Appleton *et al.*, 1995; Krishnan *et al.*, 1998), especially when wages estimates are corrected for sample selection (Appleton *et al.*, 1999). Our results also show that in Ethiopia, education is generally more rewarded in private wage employment, especially in the informal sector. This also contrasts with findings from other African countries. A study on West African cities by Kuepie *et al.* (2009) found that the returns to education were usually higher in the public sector.

Looking at the role of potential work experience, our results indicate that the impact is positive and concave for both men and women, with slightly greater returns for women in overall wage employment, in the public and the informal private sectors, and at the third quartile of the wage distribution. Women's potential work experience is thus less valued than men's in the formal private sector, and at the first quartile and the median of the wage distribution. The latter results

---

<sup>24</sup> For women, sample selection variables does not seem to be significant except in public wage employment where the selectivity coefficient related to participation in formal private wage employment is also positive. This indicates higher earnings in the public sector than those of randomly chosen individuals on account of the allocation of individuals with worse unobserved characteristics out of the public sector and into the formal private sector. However, this coefficient is significant at only 10 per cent level. As regards men, the selection correction term is statistically significant and negative in overall wage employment, and for the three different age cohorts, meaning that unobserved characteristics increase the probability of selection into wage employment but decrease the log hourly earnings. In the sectoral earnings equations, selection corrections terms provide mixed results. While they are in most cases not statistically significant in formal and informal private wage equations, in public wage employment they are significant.

might be explained in part by an overestimation of the true experience for women. Lastly, it is worth noticing that potential work experience only affects earnings among the oldest cohort. As regards training, coefficients are in all cases statistically significant, positive and higher for men. The coefficients of the job characteristics variables provide other interesting results. For both men and women, working in the public sector, followed by the formal private sector is attached with the highest wage premium. The wage premiums associated with participating in the public and the formal private sectors are however clearly higher for women. Unsurprisingly, permanent employees and high-skilled white collars are those who can expect the highest pay-off. Moreover, wages premiums associated to high-skilled white collar occupations are always greater for men. The results further point to a significant impact of the sector of activity on wages. This effect is usually different for men and women and varies by age cohorts and wage levels. The sectoral effects tend to be stronger among the youngest cohort and at the bottom of the wage distribution. The place and region of residence are also important determinants of wages in Ethiopia. Being located in urban areas is associated with higher wages, especially among men. To a large extent, these findings mirror the large dichotomy prevailing in Ethiopia between urban and rural areas. The large majority of the Ethiopian population lives in rural areas and is mostly employed in agriculture. Formal unemployment is more prevalent in urban areas but people there can enjoy greater educational and work opportunities as compared to rural areas, where the lack of opportunities encourage people to migrate to urban centres or to accept low paid jobs. The effect of regional location is also important for both men and women. This is not surprising since Ethiopia is characterized by a diversity of climates and geographies, cultures, languages and ethnic groups that influence job opportunities and choice. Moreover, regions in Ethiopia are increasingly playing a substantial role in determining their own economic policies, and this is expected to increase the disparities across local labour markets.

All in all, the previous results indicate that education and job characteristics are important determinants of wages in Ethiopia. Yet, the way these variables affect wages is somewhat different between men and women and varies across age cohorts, types of wage employment and level of wages, suggesting that wage determination may not follow a single process. Moreover, since gender selection across occupations and sectors of activity tend to be important in Ethiopia, this may constitute an important source of the explained gender wage gap. In what follows, we look at the results of the wage decompositions, drawing a particular attention to the role of education and job selection.

#### **2.4.2 Wage decompositions**

We use the OLS estimates of the previous wage equations to implement the decomposition procedures proposed by Neumark and Cotton. What lies behind the gender pay gap in Ethiopia? Are discrimination and other non-observable factors important and similar across the age cohorts, the wage distribution and the types of employment? Could education parity contribute to a significant reduction in wage disparities across gender? Is selection across occupations and sectors of activity an important source of the gender wage gap? Tables 2.11 to 2.14 in Appendix C provide the results of the gender earnings gap decompositions for two different specifications (accounting for job characteristics and not) in overall wage employment, in three different age cohorts, at three quartiles of the wage distribution and in three different types of employment, respectively.

As regards overall wage employment, the results indicate that a non-negligible proportion of the gender wage gap (between 13 and 14 per cent when accounting for job characteristics and between 26 and 29 per cent when not) can be explained by the differences in education endowments between men and women. Together with the differences in educational background, the differences between men and women in potential work experience and training contribute to

an increase in the wage gap. Overall, from at least 24 per cent to at most 49 per cent can be attributed to differences in human capital characteristics (education, potential experience and training).

When job characteristics are taken into account (sector of activity, types of wage employment, terms of employment and occupation), the contribution of the differences in human capital variables is sharply reduced, and the contribution of the differences in job characteristics between men and women rises to between 50 and 56 per cent of the wage gap, suggesting that much of the education wage gap operates through job selection. Among the differences in job characteristics, selection across sectors of activity accounts for between 24 and 30 per cent of the gender wage gap, and selection across occupations may explain up to 10 per cent. Gender differences in types and terms of employment also contribute to the wage gap but to a lesser extent.

Accounting for job characteristics of wage employees further reduces from 38-45 per cent to 15-23 per cent the part of the gender wage gap that is unexplained and thus attributable to discrimination and other non-observable factors. We observe that this unexplained component is clearly more due to women's treatment disadvantage than to men's treatment advantage in the labour market.

Considering the importance of the effect of differences in job characteristics on the gender wage gap, it is worth exploring in more details the determinants of these disparities. Is gender selection across occupations, industries and types of employment mostly driven by the differences in education? Is job discrimination against women playing any role? Can the differences in individual work preferences between men and women explain some of the differences in job characteristics? Additional multivariate analysis performed for this Chapter showed that besides the significant effect of education on job characteristics, gender is another significant factor that either picks up a form of sex-based segmentation and/or some gender specific preferences. Thus, our results indicate that gender selection across job characteristics is not only driven by disparities

in education but also by some form of discrimination and/or differences in individual preferences. As a result, one might expect the gender wage gap to be affected by some form of discriminatory practices directly through the unexplained component, but also indirectly through job selection. The fact that job characteristics are systematically less favourable for women in Ethiopia, with more women in the informal sector and less in formal public and private jobs, and that these can be only partly explained by difference in education characteristics, argues in favour of an implicit form of job discrimination that plays in-fine against women's wages. Studies on the Ethiopia's labour market show indeed that the informal sector is mostly a residual, where activities are being pursued in the absence of other options. Most new participants in the informal sector are coming from the pool of the unemployed and first job seekers (World Bank, 2007). When regular, full-time jobs that provide clear career prospects exist and are accessible to women, they usually contribute to women empowerment and offer alternative interests and achievements to domestic work or motherhood (Lim, 2002).

Yet, differences in job characteristics may also reflect some gender specific preferences. In the absence of flexible work arrangements in formal and better-paid jobs, the burden of women's household responsibilities such as housework and childcare could provide strong incentives for women to engage in the informal sector, which offer less protection but potentially more flexibility. The 2006 ALMS finds, for instance, that among a sample of unemployed people located in Addis Ababa, women were more inclined than men to look for an independent work closer to their home, probably as the result of their household responsibilities (World Bank, 2007).

Our results also echo the findings of previous studies that looked at the role of the differences in job-related characteristics on the gender wage gap. In developed countries, Gannon *et al.* (2007) find that industry effects explain around 29 per cent of the gender wage gap in Ireland, 14 per cent and 16 per cent, respectively, in Denmark and Italy, around 7 per cent in the UK and almost

nothing in Belgium and Spain. As regards Africa, using matched employer-employee data from eleven African countries, Fafchamps *et al.* (2006) find that the gender wage gap is due in large part to sorting among firms and that most of the education wage gap is operating through job selection. Another study of gender disparities in the Malagasy labour market by Nordman *et al.* (2010) provide evidence that gender specific sectoral location explains a significant share of the gender wage gap in both 2001 and 2005, arguing that this result is mainly driven by the fact that the proportion of women is higher in the self-employed sector where earnings are lower.

In the specific case of Ethiopia, Appleton *et al.* (1999) show, using data from the 1990 Survey of Adolescent Fertility, Reproductive Behaviour and Employment Status of the Youth Population in Urban Ethiopia, that male wages exceed female wages by approximately one-quarter. In addition, their findings indicate that women employees have more favourable characteristics than their male counterparts, and that the substantial gender differential in earnings is largely due to differences in returns to wage-generating characteristics, namely discrimination and other factors which are not observed. However, these results are misleading in that they fail to account for job characteristics. Another study by Bigsten *et al.* (2007) find, using the Ethiopia Urban Household Socio Economic Survey over the period 1994-2004, large, sustained, and unexplained earnings gaps between public and private, and formal and informal sectors. The authors do not provide formal evidence whether these gaps reflect segmentation of the labour market along either of these divides, but argue that, if segmentation explains any part of the observed earnings gaps, then it could only have weakened over the survey decade. Indeed, they show that both the rate of mobility between sectors and the sensitivity of sector choice to earnings gaps increased over the same period. A more explicit reference of the impact of job selection on the gender wage gap is provided by Temesgen (2006) using a sample of workers from the Ethiopian manufacturing sector. He finds that men earn on average 30 per cent more than women in that sector. However, once he controls for a number of individual and establishment level characteristics, the level of

wage premium for men over women is reduced to 5 per cent. Then he concludes that ignoring establishment characteristics in decomposition exercises would result into a biased estimation, and in his case it would have underestimated the level of discrimination by nearly 50 per cent.

Looking at our results for the different subgroups brings other interesting results. Given the importance of differences in job attributes in explaining the gender wage gap, we further comment only the decomposition results accounting for job characteristics. As regards the decompositions by age cohorts, we see that the share of the gender wage gap attributable to the differences in job characteristics is the highest among youth and then decreases significantly for older cohorts. Within these job characteristics, industry effects are the main contributors to the gender wage gap, while the contribution of other variables (terms of employment, occupation and types of wage employment) is quite modest. A greater influence of job characteristics on the gender wage gap among young people in Ethiopia is probably illustrative of the recent country's social and economic transformation. In the early 1990's, Ethiopia undertook a programme of major reforms that initiated the transition from a highly regulated and egalitarian system to a market economy. This transition process was accompanied by a rise in gender inequality. This further contributed to the emergence of new forms of gender segmentation in the labour market with a strong impact on new entrants. What is also interesting is the fact that both the education wage gap and the unexplained discriminatory component are lower among youth. This result is encouraging and provides additional evidence of the benefits for the Government of Ethiopia to pursue its efforts towards gender equity in education and society at large.

Turning to the decomposition results across the wage distribution, it first appears that the explained component reflecting observed differences in attributes, either human capital and job characteristics, steadily decreases along the wage distribution, while the part of the gender wage gap attributable to discrimination and other non-observable factors is more marked for top wage earners. Among the human capital characteristics, potential work experience and training effects

remain fairly stable, whereas the contribution of education disparities is less important at higher quartiles. Regarding job characteristics, we observe that the share of the gender wage gap attributable to differences in the types of employment and terms of employment is more important in the lower end of the wage distribution, while selection across sectors of activity tends to contribute more to the wage gap at the top of the wage distribution. These results suggest that, compared to low-paid workers, the gender pay gap among highly paid workers is more driven by selection across industries as well as by discriminatory and other non-observable factors, and less so by the consequence of differences in educational background, types and terms of employment.

We last turn to a discussion of the decomposition results across wage employment sectors. Our results show substantial differences in the sources of the wage gap between the types of employment. Observed differences between men and women characteristics contribute to explain about 70 to 80 per cent of the wage gap in the public sector and the informal private sector, but only 46 to 57 per cent in the formal private sector, where the unexplained component is particularly important. In the formal private sector, moreover, disparities in education tend to reduce the gender wage gap (-5 per cent) while the reverse is true in the public sector (15 per cent) and in informal private employment (12 per cent). If this is due to the fact that women have a higher educational background than men, it also suggests that women must be better educated than men in order to compete with men for better paid jobs in the formal private sector. Differences in work experience further explain a large share of the wage gap in the public sector (18 per cent), and to a lower extent, in the formal private sector (10 to 11 per cent) but not in the informal private sector (only 3 to 4 per cent). As regards job characteristics, what is worth noting is that selection across industries has a large effect on the wage gap in the private sector, both in the formal (20 to 23 per cent) and informal sectors (25 to 33 per cent), but a rather modest effect in the public sector (6 per cent) where there are no substantial gender differences by industries. In



turn, selection across occupations appears as an important determinant of the wage gap in the public sector and in the formal private sector, but not in the informal private sector. Differences in the terms of employment constitute an important share of the wage gap only in the public sector. In other words, these latest results indicate that in the public sector much of the gender wage gap can be explained, and it is driven mostly by the selection across occupations and the disparities in education and work experience. In the informal private sector, most of the wage gap can also be explained by differences in observed characteristics, above all the selection across industries and to a lower extent the differences in education. Explaining the gender wage gap prevailing in the formal private sector turns out to be more difficult. There, the explained component can be attributed mostly to selection across industries and differences in education. Yet, about half of the wage gap remains unexplained and this suggests that discriminatory practices may be more a matter of concern in the formal private sector.

## **2.5 Conclusion**

This Chapter has sought to contribute to a better understanding of the factors related to the gender wage gap in Ethiopia. We drew a special attention on the relative effect of education parity to mitigate the most pressing wage inequalities. We further investigated how this effect may vary across different age cohorts, different types of wage employment and at different levels in the wage distribution. Using different estimation models for the wage equations and the Cotton-Neumark decomposition procedure for the estimated wage gaps, we were able to isolate the determinants of the pay differentials and to examine the way the contributing factors may vary by age, at different points in the wage distribution, and for different types of employment. Despite the acknowledged shortcomings of the decomposition, which is fairly sensitive to the quality of the information available and the estimation models used, the results provide some interesting insights. The main findings can be summarized as follows.

First, the data indicate that on average female hourly wages represented in 2005 only about 66 per cent of male wages. There were also large variations across sub-groups. The raw gender pay differential was found to be more pronounced among youth, low-wage earners and for wage employed in the informal sector.

Second, the decomposition results showed that accounting for job characteristics substantially reduced the size of the unexplained component. The fraction of the gap that remains unexplained is generally found to be rather modest, except in the formal private sector where it was particularly high, suggesting that discriminatory practices may be more a matter of concern in the formal private sector.

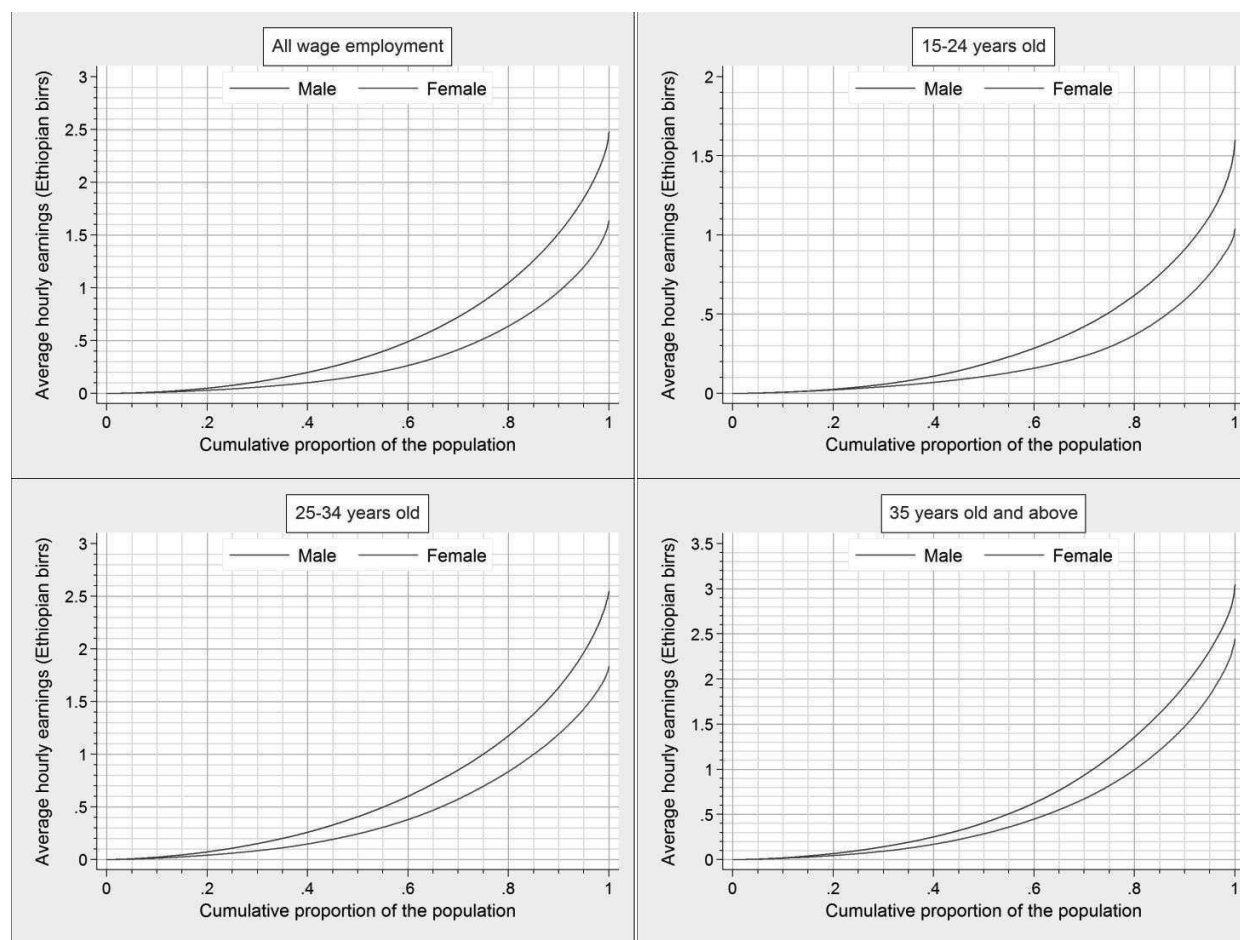
Third, a non-negligible proportion of the gender wage gap, at least an average 14 per cent once controlling for jobs characteristics but no more than 28 per cent otherwise, was explained by the differences in education endowments between men and women. This further indicates that on average about half of the education wage gap was driven by job selection, in particular selection across industries and occupations. Interestingly, the education gender wage gap appeared to be less important for younger workers.

Last, our results suggest that some form of discriminatory practices may contribute to the gender wage gap both directly through the unexplained component and indirectly through job selection.

Overall, these findings indicate that progress towards gender equity in education is important to improve women's relative wages both directly, as a consequence of the private returns of education, but also indirectly through a more gender-balanced job allocation. The benefits of such policies are indeed already visible for younger workers. Our result also suggests that other interventions may be needed to compensate for the adverse impact of possible discriminatory practices, gender specific preferences, cultural and other non-observable factors. While education is an essential asset for changing attitudes and making work pay, formal education alone is unlikely to bring about change. Information campaigns and other awareness-raising efforts to

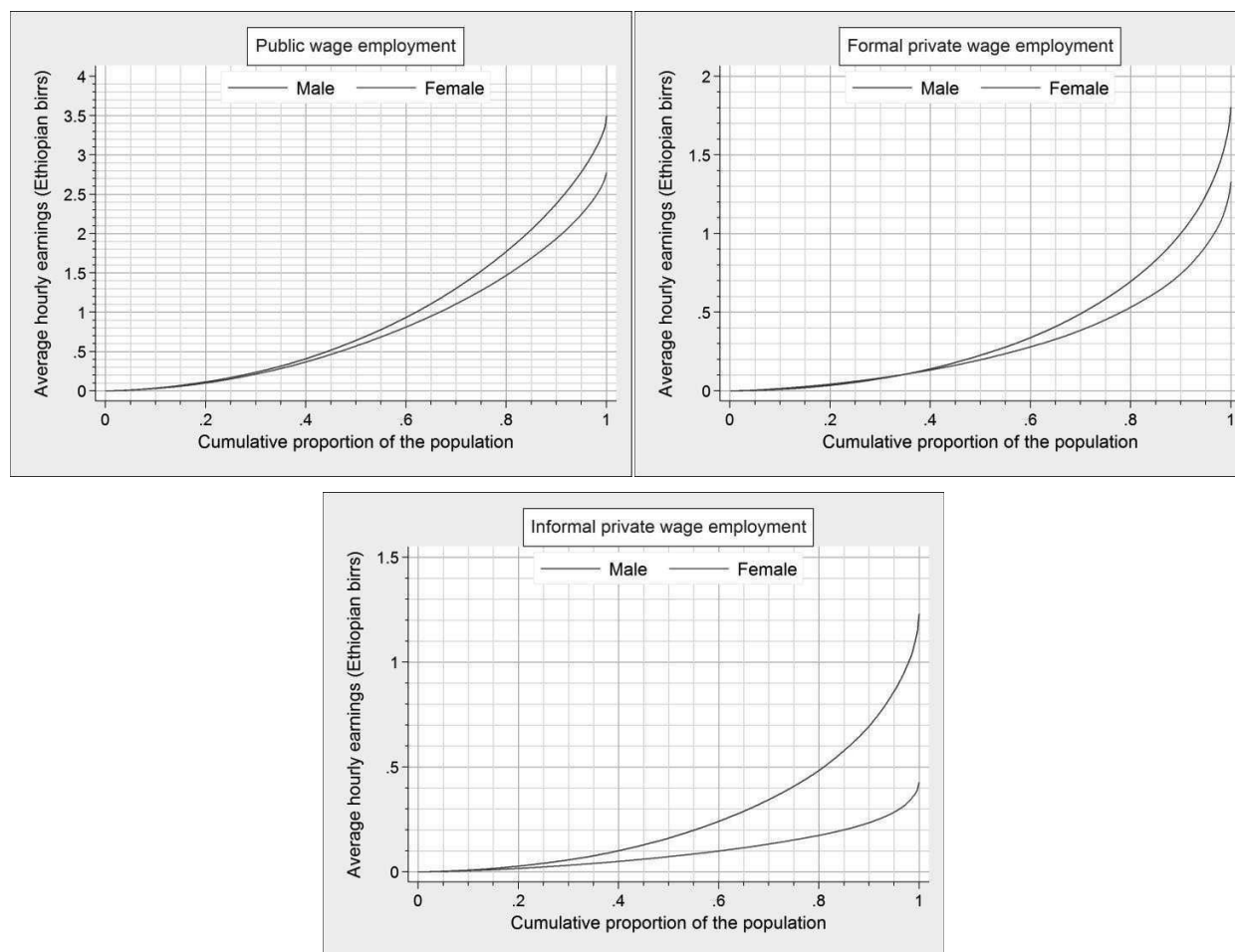
make Ethiopia's population aware of the anti-discriminatory provisions of its own constitution and legislation could be also an important way to support women wages while making initial investment in girls' education more effective.

## Appendix A. Generalized Lorenz curves for hourly earnings



**Figure 2.1: Generalized Lorenz curves for hourly earnings in main occupation by gender and age cohort**

Source: *Ethiopia LFS, 2005*. Note: *Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*



**Figure 2.2: Generalized Lorenz curves for hourly earnings in main occupation by gender and wage employment sector**  
Source: *Ethiopia LFS, 2005*. Note: *Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*

## Appendix B. Earnings equations

**Table 2.6: Summary statistics of the variables used in the earnings equations**

	All wage workers		15-24		25-34		35+		Public		Formal private		Informal private	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Illiterate*	14.4%	29.7%	19.2%	35.2%	9.7%	21.0%	15.3%	31.7%	7.2%	13.3%	17.9%	21.9%	30.6%	54.7%
Primary education	33.7%	28.6%	41.1%	35.6%	31.1%	23.7%	32.0%	23.2%	23.0%	18.0%	41.4%	31.8%	49.7%	37.8%
General education	27.8%	23.4%	25.2%	17.6%	33.7%	32.8%	24.9%	21.1%	30.3%	33.7%	28.0%	29.7%	17.3%	6.5%
Beyond general education	24.1%	18.3%	14.5%	11.6%	25.6%	22.6%	27.9%	23.9%	39.6%	35.0%	12.7%	16.6%	2.4%	1.0%
Potential experience	20.2	15.9	7.9	8.7	14.1	14.6	30.9	29.4	21.1	17.7	19.3	14.0	20.0	15.4
Training	41.8%	26.9%	25.7%	17.0%	45.1%	34.6%	47.5%	33.5%	59.2%	49.6%	29.5%	25.8%	15.1%	2.3%
Married	58.4%	31.9%	11.9%	11.8%	53.7%	42.5%	85.2%	51.9%	70.8%	55.3%	48.4%	27.8%	43.4%	9.1%
Primary sector activity*	8.6%	3.7%	14.0%	3.4%	6.5%	2.9%	7.5%	5.4%	4.9%	4.5%	12.4%	5.9%	10.8%	1.1%
Manufacturing	12.7%	9.7%	14.8%	6.2%	11.8%	10.0%	12.3%	15.4%	7.7%	9.3%	18.9%	16.8%	11.3%	4.3%
Tertiary sector 1	14.3%	6.4%	16.1%	6.1%	15.6%	8.1%	12.4%	4.8%	8.6%	5.9%	18.7%	12.8%	21.7%	1.6%
Tertiary sector 2	9.2%	13.2%	14.6%	14.2%	9.2%	13.7%	6.5%	10.9%	2.1%	2.9%	16.4%	33.8%	12.9%	7.4%
Tertiary sector 3	10.4%	3.9%	9.7%	2.0%	11.6%	4.8%	9.8%	5.8%	9.4%	7.7%	11.5%	3.2%	n.a.	n.a.
Tertiary sector 4	19.8%	14.8%	9.5%	7.5%	22.0%	19.6%	23.4%	20.6%	40.4%	36.7%	1.6%	2.7%	n.a.	n.a.
Tertiary sector 5	20.4%	19.5%	15.8%	12.9%	19.2%	22.8%	23.5%	26.3%	26.7%	33.0%	14.5%	19.1%	15.0%	4.8%
Tertiary sector 6	2.1%	27.1%	3.7%	46.8%	1.5%	15.7%	1.7%	9.1%	n.a.	n.a.	n.a.	n.a.	17.3%	80.2%
Public sector	47.2%	37.9%	22.3%	16.8%	45.9%	45.0%	60.6%	63.9%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Formal private sector	40.6%	28.4%	56.2%	30.0%	42.8%	33.1%	31.2%	19.6%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Informal private sector*	12.2%	33.7%	21.4%	53.2%	11.4%	21.9%	8.1%	16.5%	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Government employee	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	84.9%	84.5%	n.a.	n.a.	n.a.	n.a.
Gvt parastatal employee*	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	15.1%	15.5%	n.a.	n.a.	n.a.	n.a.
NGO employee	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	11.8%	12.6%	n.a.	n.a.
Domestic employee	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	42.5%	88.2%
Private Orgn. employee*	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	88.2%	87.4%	57.5%	11.8%

Table 2.6: *Continued*

	All wage workers		15-24		25-34		35+		Public		Formal private		Informal private	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Permanent employee	49.2%	36.9%	21.7%	16.5%	48.6%	43.8%	63.5%	61.7%	81.4%	77.3%	24.0%	20.8%	8.1%	5.0%
Temporary employee	33.9%	48.0%	54.8%	67.5%	34.0%	40.5%	23.3%	25.2%	10.8%	12.7%	53.6%	57.9%	57.4%	79.3%
Contract employee	9.7%	7.8%	11.5%	8.6%	10.4%	8.6%	8.3%	5.3%	6.1%	6.9%	13.8%	13.2%	9.7%	4.3%
Casual or other worker*	7.2%	7.4%	12.1%	7.3%	7.1%	7.1%	5.0%	7.7%	1.6%	3.2%	8.6%	8.2%	24.8%	11.4%
High-skilled white collar	26.6%	15.8%	13.3%	8.7%	26.8%	19.2%	33.1%	23.1%	44.2%	31.9%	13.0%	11.9%	3.9%	0.8%
Low-skilled white collar	14.9%	25.8%	16.0%	21.6%	17.0%	31.3%	12.8%	25.7%	18.3%	32.3%	12.2%	36.5%	10.8%	9.4%
High-skilled blue collar	18.0%	8.4%	26.5%	5.8%	19.5%	8.8%	12.6%	12.4%	8.3%	6.0%	26.4%	10.8%	27.8%	9.2%
Low skilled blue collar*	40.5%	50.0%	44.1%	64.0%	36.6%	40.7%	41.5%	38.8%	29.3%	29.8%	48.3%	40.7%	57.6%	80.6%
Urban	87.3%	92.3%	81.2%	92.1%	88.4%	92.7%	89.6%	92.0%	89.3%	89.7%	85.5%	92.1%	85.8%	95.4%
Addis Ababa*	28.6%	32.7%	29.8%	33.7%	30.2%	32.5%	26.8%	31.2%	18.6%	23.1%	39.6%	40.9%	30.5%	36.5%
Tigray	6.8%	8.6%	6.4%	8.1%	6.2%	9.3%	7.3%	8.6%	7.7%	12.3%	6.1%	7.5%	5.3%	5.3%
Affar	4.9%	3.2%	4.2%	3.4%	4.5%	2.7%	5.5%	3.6%	6.0%	3.7%	4.5%	2.9%	1.9%	2.9%
Amhara	15.2%	16.9%	16.3%	18.8%	13.6%	14.6%	15.8%	16.7%	17.6%	17.2%	13.0%	13.4%	13.3%	19.6%
Oromiya	19.4%	17.5%	19.8%	16.6%	17.5%	17.2%	20.5%	19.5%	20.4%	17.8%	16.8%	18.5%	24.1%	16.5%
Somalie	2.7%	1.9%	1.8%	1.9%	3.1%	2.3%	2.8%	1.5%	3.3%	2.3%	1.8%	0.9%	3.2%	2.5%
Benshangul Gumuz	3.3%	2.5%	3.2%	2.9%	4.6%	2.9%	2.5%	1.2%	4.8%	4.2%	2.1%	0.7%	1.9%	2.0%
SNNP	12.9%	11.2%	13.3%	10.5%	13.3%	12.6%	12.4%	10.5%	14.7%	11.9%	10.3%	11.4%	14.6%	10.2%
Harari	2.3%	2.2%	1.8%	1.4%	2.5%	2.2%	2.5%	3.5%	2.8%	3.6%	1.8%	1.1%	2.2%	1.4%
Dire Dawa	2.6%	2.3%	2.5%	1.8%	2.5%	2.5%	2.6%	2.8%	2.1%	2.1%	3.1%	2.2%	2.3%	2.6%
Observations	12776	8616	2886	3632	4144	2790	5746	2194	6030	3267	5190	2444	1556	2905

Source: Ethiopia LFS, 2005. Note: Individuals aged 15 and above.

\* Reference category in earnings equations. Wage employees are classified into four occupational groups. High-skilled white collars include legislators, senior officials and managers, professionals, technicians and associate professionals. Low-skilled white collars include clerks, service workers and shop and market sales workers. High-skilled blue collars include skilled agricultural and fishery workers, craft and related trade workers. Finally, low-skilled blue collars include plant and machine operators and assemblers, and elementary occupations. By default, armed forces which account for a very tiny minority of the wage employed are considered to be low-skilled white collars. The service industry is disaggregated into six sectors: tertiary sector 1 (electricity, gas and water supply, construction), tertiary sector 2 (wholesale and retail trade, hotels and restaurants), tertiary sector 3 (transport, storage and communications, financial intermediation), tertiary sector 4 (real estate, renting and business activities, public administration and defence, compulsory social security), sector 5 (education, health and social work, other community, social and personal service activities), and tertiary sector 6 (private households with employed persons).

Table 2.7: OLS and selectivity corrected (Heckman's two-step method) log hourly earnings equations in wage employment by gender

	Men		Women	
	OLS	Heckman	OLS	Heckman
Primary education	0.2521*** (9.42)	0.2150*** (8.73)	0.3123*** (14.04)	0.3143*** (14.18)
General education	0.5825*** (19.05)	0.4467*** (10.64)	0.7256*** (23.14)	0.7164*** (18.24)
Beyond general education	0.9499*** (25.52)	0.7262*** (11.50)	1.0089*** (24.14)	0.9794*** (10.98)
Potential experience	0.0282*** (15.68)	0.0038 (0.63)	0.0294*** (12.06)	0.0251** (2.15)
(Potential experience) <sup>2</sup>	-0.0004*** (-13.55)	0.0000 (0.08)	-0.0003*** (-6.04)	-0.0002 (-0.84)
Training	0.2475*** (14.52)	-0.0505 (-0.71)	0.1775*** (6.87)	0.1468* (1.71)
Married	0.1466*** (8.99)	0.1099*** (5.64)	0.1410*** (7.39)	0.1683** (2.25)
Manufacturing	0.0125 (0.37)	0.0033 (0.11)	-0.2495*** (-5.34)	-0.2502*** (-6.00)
Tertiary sector 1	0.2555*** (7.85)	0.2452*** (8.97)	0.0678 (1.37)	0.0673 (1.54)
Tertiary sector 2	-0.2472*** (-6.57)	-0.2556*** (-8.12)	-0.5327*** (-11.27)	-0.5334*** (-13.08)
Tertiary sector 3	0.2185*** (6.47)	0.2088*** (6.72)	0.0752 (1.42)	0.0751 (1.45)
Tertiary sector 4	0.0005 (0.01)	-0.0093 (-0.31)	-0.1090** (-2.32)	-0.1094*** (-2.66)
Tertiary sector 5	-0.0737** (-2.31)	-0.0864*** (-3.06)	-0.1888*** (-4.16)	-0.1895*** (-4.78)
Tertiary sector 6	-0.5177*** (-8.48)	-0.5378*** (-10.32)	-0.6310*** (-11.30)	-0.6320*** (-13.24)
Public sector	0.2037*** (6.41)	0.1946*** (7.28)	0.4875*** (11.28)	0.4870*** (13.15)
Formal private sector	0.1953*** (6.66)	0.1887*** (8.36)	0.3457*** (9.00)	0.3456*** (10.63)
Permanent employee	0.2522*** (7.36)	0.2499*** (8.51)	0.5214*** (12.22)	0.5208*** (14.57)
Temporary employee	-0.1586*** (-5.01)	-0.1579*** (-6.15)	-0.0416 (-1.15)	-0.0419 (-1.41)



**Table 2.7: Continued**

Contract employee	-0.0038 (-0.10)	-0.0054 (-0.17)	0.1834*** (3.95)	0.1827*** (4.68)
High-skilled white collar	0.6493*** (27.95)	0.6514*** (27.47)	0.4363*** (13.46)	0.4359*** (13.29)
Low-skilled white collar	0.1635*** (7.47)	0.1651*** (7.60)	0.0345 (1.36)	0.0346 (1.46)
High-skilled blue collar	0.2979*** (13.89)	0.2979*** (15.77)	0.0471 (1.38)	0.0472 (1.59)
Urban	0.3040*** (11.23)	-0.0782 (-0.86)	0.1389*** (3.67)	0.0905 (0.69)
Tigray	-0.0774*** (-2.81)	-0.0508* (-1.67)	-0.0284 (-0.91)	-0.0322 (-1.02)
Affar	0.1832*** (5.85)	0.0775* (1.81)	0.1295*** (2.74)	0.1282*** (2.89)
Amhara	-0.2052*** (-9.55)	-0.0978*** (-2.95)	-0.3605*** (-15.96)	-0.3520*** (-11.17)
Oromiya	-0.1668*** (-8.71)	-0.0604* (-1.89)	-0.1850*** (-8.67)	-0.1733*** (-4.57)
Somalie	0.1400*** (3.57)	0.3451*** (5.41)	0.2061*** (4.07)	0.2338** (2.56)
Benshangul Gumuz	-0.0365 (-1.04)	0.0638 (1.39)	-0.0694* (-1.66)	-0.0578 (-1.00)
SNNP	-0.2381*** (-11.00)	-0.0836** (-1.97)	-0.2711*** (-10.36)	-0.2554*** (-5.23)
Harari	-0.0205 (-0.50)	0.0832 (1.60)	0.0373 (0.83)	0.0426 (0.80)
Dire Dawa	-0.0845** (-1.96)	-0.0074 (-0.16)	-0.0833* (-1.71)	-0.0717 (-1.23)
Constant	-1.1230*** (-21.29)	0.2296 (0.73)	-1.3150*** (-19.01)	-1.1562*** (-2.72)
Mill's ratio	— —	-0.5602*** (-4.38)	— —	-0.0623 (-0.38)
Observations	12776		8616	
R2	0.5636	0.5642	0.7194	0.7194

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above; t-statistics in parenthesis.

\* Significant at 10% level, \*\* Significant at 5% level, \*\*\* Significant at 1% level.

**Table 2.8: OLS and selectivity corrected (Heckman's two-step method) log hourly earnings equations in wage employment by gender and age cohort**

	15-24				25-34				35+			
	Men		Women		Men		Women		Men		Women	
	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman
Primary education	0.3838*** (5.88)	0.4782*** (6.08)	0.2810*** (8.73)	0.3015*** (3.64)	0.2197*** (3.43)	0.1612*** (2.75)	0.2866*** (5.09)	0.2867*** (5.64)	0.1659*** (4.42)	0.0581 (1.31)	0.2235*** (4.29)	0.1768** (2.08)
General education	0.6861*** (8.47)	0.4515*** (3.32)	0.6383*** (11.46)	0.6384*** (11.78)	0.4449*** (5.87)	0.3039*** (3.20)	0.5784*** (7.66)	0.5791*** (5.25)	0.4633*** (10.34)	0.2566*** (3.51)	0.6076*** (8.52)	0.4860*** (2.50)
Beyond general education	0.9212*** (9.20)	0.5610*** (2.82)	0.9049*** (11.34)	0.8727*** (6.09)	0.7995*** (9.21)	0.5522*** (3.79)	0.8384*** (9.90)	0.8401*** (3.74)	0.8378*** (15.98)	0.4875*** (4.30)	0.9179*** (9.96)	0.6993*** (2.08)
Potential experience	-0.0097 (-0.63)	-0.2349** (-2.36)	0.0135 (1.14)	-0.0082 (-0.10)	-0.0051 (-0.40)	-0.0245 (-1.50)	-0.0209 (-1.48)	-0.0207 (-0.94)	0.0129*** (3.21)	0.0134*** (2.94)	0.0218** (2.45)	0.0187* (1.76)
(Potential experience) <sup>2</sup>	0.0024*** (2.74)	0.0102*** (2.89)	0.0001 (0.10)	0.0009 (0.28)	0.0001 (0.23)	0.0008 (1.43)	0.0009* (1.86)	0.0009 (1.33)	-0.0002*** (-4.24)	-0.0001 (-1.33)	-0.0002** (-1.97)	-0.0001 (-0.56)
Training	0.2943*** (7.08)	-0.6669 (-1.59)	0.1913*** (3.80)	0.1370 (0.66)	0.2543*** (8.84)	0.0161 (0.13)	0.1440*** (3.71)	0.1448 (1.34)	0.2427*** (9.90)	-0.0268 (-0.33)	0.2011*** (4.37)	0.1275 (1.05)
Married	0.1857*** (3.59)	0.1588** (2.57)	0.1922*** (4.74)	0.2523 (1.12)	0.1269*** (5.27)	0.0627 (1.52)	0.1729*** (5.98)	0.1721 (1.56)	0.1316*** (5.11)	0.0883*** (2.79)	0.0750** (2.30)	0.1385 (1.39)
Manufacturing	0.2713*** (3.80)	0.2663*** (4.21)	-0.3283*** (-4.27)	-0.3281*** (-4.72)	-0.0649 (-1.08)	-0.0687 (-1.35)	-0.3031*** (-3.60)	-0.3031*** (-4.21)	-0.1090** (-2.25)	-0.1170*** (-2.97)	-0.1300 (-1.46)	-0.1307* (-1.70)
Tertiary sector 1	0.4922*** (7.32)	0.4870*** (8.12)	0.1583** (2.07)	0.1585** (2.31)	0.1811*** (3.12)	0.1782*** (3.74)	-0.0560 (-0.65)	-0.0560 (-0.77)	0.1184** (2.55)	0.1103*** (2.88)	0.0067 (0.07)	0.0040 (0.04)
Tertiary sector 2	-0.0772 (-0.97)	-0.0833 (-1.25)	-0.5641*** (-7.49)	-0.5639*** (-8.72)	-0.2425*** (-3.65)	-0.2468*** (-4.48)	-0.5861*** (-6.83)	-0.5861*** (-8.44)	-0.2943*** (-5.42)	-0.3057*** (-6.59)	-0.3435*** (-3.67)	-0.3457*** (-4.29)
Tertiary sector 3	0.5520*** (7.20)	0.5509*** (7.39)	0.0637 (0.62)	0.0648 (0.67)	0.1353** (2.27)	0.1326** (2.48)	-0.0622 (-0.70)	-0.0622 (-0.73)	0.0509 (1.08)	0.0455 (1.05)	0.1876* (1.94)	0.1864** (2.01)
Tertiary sector 4	0.3073*** (3.71)	0.3052*** (3.75)	-0.0802 (-0.96)	-0.0805 (-1.11)	-0.0414 (-0.72)	-0.0425 (-0.81)	-0.2036** (-2.49)	-0.2036*** (-2.95)	-0.1516*** (-3.59)	-0.1605*** (-4.12)	-0.0601 (-0.69)	-0.0618 (-0.83)
Tertiary sector 5	0.3232*** (4.46)	0.3180*** (4.78)	-0.1684** (-2.20)	-0.1679** (-2.51)	-0.1374** (-2.39)	-0.1443*** (-2.89)	-0.2933*** (-3.74)	-0.2933*** (-4.39)	-0.2568*** (-5.97)	-0.2671*** (-7.01)	-0.0965 (-1.07)	-0.0989 (-1.33)
Tertiary sector 6	-0.5552*** (-5.23)	-0.5650*** (-5.75)	-0.6775*** (-8.08)	-0.6777*** (-9.47)	-0.2732** (-2.25)	-0.2827*** (-2.83)	-0.6535*** (-6.54)	-0.6534*** (-7.53)	-0.6174*** (-6.76)	-0.6411*** (-8.32)	-0.4514*** (-3.55)	-0.4551*** (-4.59)
Public sector	0.2695*** (4.38)	0.2631*** (4.23)	0.5570*** (7.95)	0.5567*** (9.41)	0.2373*** (4.32)	0.2329*** (5.02)	0.3914*** (5.51)	0.3914*** (6.15)	0.1184** (2.17)	0.1093*** (2.75)	0.4913*** (5.22)	0.4915*** (6.57)
Formal private sector	0.1690*** (3.49)	0.1641*** (4.02)	0.3430*** (6.06)	0.3430*** (7.39)	0.3200*** (6.38)	0.3188*** (8.23)	0.3247*** (5.13)	0.3247*** (5.63)	0.0817 (1.54)	0.0720* (1.93)	0.2998*** (3.33)	0.3002*** (4.28)
Permanent employee	0.0903 (1.43)	0.0890 (1.41)	0.3317*** (5.25)	0.3314*** (6.12)	0.2997*** (5.14)	0.2997*** (5.88)	0.6866*** (9.01)	0.6866*** (11.01)	0.3303*** (5.45)	0.3313*** (7.59)	0.6035*** (6.26)	0.6019*** (7.87)
Temporary employee	-0.1919*** (-3.76)	-0.1904*** (-4.06)	-0.0724 (-1.42)	-0.0727* (-1.70)	-0.0861 (-1.58)	-0.0845* (-1.90)	0.0557 (0.85)	0.0557 (1.05)	-0.1495** (-2.53)	-0.1481*** (-3.63)	-0.0414 (-0.51)	-0.0427 (-0.68)

Table 2.8: Continued

Contract employee	-0.0130 (-0.20)	-0.0074 (-0.12)	0.2066*** (3.18)	0.2065*** (3.71)	0.0898 (1.45)	0.0915* (1.70)	0.2598*** (3.16)	0.2598*** (3.82)	-0.0338 (-0.50)	-0.0325 (-0.68)	0.0564 (0.49)	0.0541 (0.61)
High-skilled white collar	0.5852*** (9.42)	0.5660*** (7.70)	0.4961*** (8.12)	0.4953*** (8.43)	0.5993*** (14.99)	0.5984*** (15.29)	0.4528*** (8.76)	0.4528*** (8.69)	0.7101*** (21.87)	0.7086*** (22.42)	0.4001*** (6.34)	0.3999*** (6.14)
Low-skilled white collar	0.0097 (0.19)	0.0107 (0.21)	-0.0891** (-2.13)	-0.0893** (-2.47)	0.1387*** (3.91)	0.1392*** (3.92)	0.0992** (2.39)	0.0992** (2.45)	0.2618*** (7.78)	0.2612*** (8.19)	0.1050** (1.99)	0.1060** (2.11)
High-skilled blue collar	0.2227*** (5.05)	0.2227*** (5.56)	0.1441** (2.52)	0.1443*** (2.82)	0.2326*** (6.37)	0.2335*** (7.36)	0.0407 (0.67)	0.0407 (0.80)	0.4154*** (12.62)	0.4176*** (14.66)	-0.0582 (-1.01)	-0.0568 (-1.08)
Urban	0.2396*** (4.68)	-0.5434 (-1.59)	0.1443*** (2.72)	0.0528 (0.15)	0.2197*** (4.99)	-0.1026 (-0.61)	0.2191*** (3.27)	0.2204 (1.20)	0.3430*** (7.54)	-0.2393 (-1.40)	0.0373 (0.41)	-0.1033 (-0.47)
Tigray	-0.1445** (-2.44)	0.0190 (0.18)	-0.1302*** (-2.97)	-0.1344*** (-2.89)	0.0736 (1.54)	0.0913* (1.74)	0.0968* (1.67)	0.0972 (1.37)	-0.1319*** (-3.29)	-0.1643*** (-3.67)	0.0206 (0.31)	-0.0192 (-0.22)
Affar	0.3649*** (5.07)	0.3519*** (3.50)	0.1633** (2.48)	0.1703** (2.54)	0.1157** (1.96)	0.0255 (0.33)	0.0264 (0.31)	0.0265 (0.31)	0.1220*** (2.87)	-0.0988 (-1.20)	0.1704 (1.61)	0.1126 (0.89)
Amhara	-0.2596*** (-5.35)	-0.0427 (-0.39)	-0.5665*** (-17.24)	-0.5533*** (-9.54)	-0.1126*** (-2.96)	-0.0188 (-0.31)	-0.2539*** (-6.10)	-0.2542*** (-4.76)	-0.2093*** (-7.08)	-0.0854* (-1.74)	-0.1422*** (-3.13)	-0.1296*** (-2.58)
Oromiya	-0.1613*** (-3.59)	0.0848 (0.71)	-0.2471*** (-7.62)	-0.2217** (-2.23)	-0.1295*** (-3.86)	-0.0286 (-0.46)	-0.1962*** (-5.12)	-0.1965*** (-3.84)	-0.1738*** (-6.58)	-0.0607 (-1.35)	-0.0769* (-1.83)	-0.0631 (-1.29)
Somalie	0.2161** (2.10)	0.9065*** (2.79)	0.3000*** (3.97)	0.3553 (1.62)	0.1120* (1.79)	0.2248** (2.53)	0.1853** (2.17)	0.1848* (1.66)	0.1574*** (2.77)	0.4253*** (4.30)	0.0618 (0.60)	0.1456 (0.83)
Benshangul Gumuz	-0.0616 (-0.77)	0.2457 (1.43)	-0.1694** (-2.57)	-0.1489 (-1.47)	0.0273 (0.55)	0.0651 (1.05)	-0.0247 (-0.41)	-0.0249 (-0.30)	-0.0675 (-1.15)	0.1174 (1.37)	0.0345 (0.32)	0.0864 (0.54)
SNNP	-0.2571*** (-5.28)	0.1206 (0.69)	-0.3616*** (-9.61)	-0.3277** (-2.49)	-0.1777*** (-4.92)	-0.0338 (-0.41)	-0.2142*** (-4.91)	-0.2145*** (-3.78)	-0.2497*** (-7.94)	-0.0747 (-1.19)	-0.2236*** (-3.68)	-0.1912*** (-2.62)
Harari	-0.0943 (-0.88)	0.2230 (1.14)	-0.0135 (-0.17)	0.0079 (0.07)	0.0211 (0.33)	0.1121 (1.27)	0.0930 (1.22)	0.0929 (1.04)	-0.0029 (-0.05)	0.0920 (1.26)	0.0637 (0.84)	0.0343 (0.36)
Dire Dawa	-0.2113* (-1.75)	-0.0212 (-0.15)	-0.0354 (-0.52)	-0.0037 (-0.03)	-0.0576 (-0.89)	-0.0132 (-0.17)	-0.1970** (-2.09)	-0.1972** (-2.28)	-0.0619 (-1.13)	0.0399 (0.56)	0.0255 (0.31)	0.0258 (0.28)
Constant	-1.2366*** (-10.33)	2.3381 (1.52)	-1.0552*** (-8.97)	-0.7590 (-0.69)	-0.6552*** (-5.18)	0.4031 (0.73)	-0.8914*** (-5.26)	-0.8964 (-1.35)	-0.6404*** (-5.61)	0.6714* (1.73)	-1.1453*** (-5.49)	-0.6741 (-0.93)
Mill's ratio	—	-1.2044** (-2.34)	—	-0.0935 (-0.27)	—	-0.4916** (-1.99)	—	0.0021 (0.01)	—	-0.7675*** (-3.51)	—	-0.2341 (-0.68)
Observations	2886		3632		4144		2790		5746		2194	
R2	0.4474	0.4487	0.6772	0.6772	0.4853	0.4858	0.6529	0.6529	0.5985	0.5994	0.6599	0.6599

Source: Ethiopia LFS, 2005.

Note: t-statistics in parenthesis.

\* Significant at 10% level, \*\* Significant at 5% level, \*\*\* Significant at 1% level.

**Table 2.9: Simultaneous quantile log hourly earnings regression estimates in wage employment by gender**

	First Quartile		Median		Third Quartile	
	Men	Women	Men	Women	Men	Women
Primary education	0.3001*** (8.62)	0.3287*** (10.73)	0.2685*** (8.96)	0.3155*** (12.37)	0.2145*** (7.27)	0.2726*** (10.31)
General education	0.6647*** (16.59)	0.6983*** (16.14)	0.6051*** (18.38)	0.7156*** (22.01)	0.5366*** (16.36)	0.6699*** (18.86)
Beyond general education	1.0337*** (20.11)	0.9894*** (19.46)	0.9589*** (22.09)	0.9967*** (25.71)	0.8689*** (23.28)	0.9528*** (21.41)
Potential experience	0.0347*** (14.55)	0.0259*** (7.99)	0.0284*** (17.32)	0.0260*** (8.81)	0.0246*** (11.57)	0.0325*** (10.57)
(Potential experience) <sup>2</sup>	-0.0005*** (-12.63)	-0.0003*** (-3.93)	-0.0004*** (-13.53)	-0.0002*** (-4.11)	-0.0003*** (-8.41)	-0.0003*** (-4.94)
Training	0.2303*** (10.45)	0.1751*** (5.61)	0.2253*** (13.95)	0.1708*** (6.27)	0.2500*** (13.33)	0.1611*** (6.45)
Married	0.1162*** (5.12)	0.1188*** (5.08)	0.1155*** (8.62)	0.1472*** (6.40)	0.1214*** (6.19)	0.1332*** (7.67)
Manufacturing	0.0492 (1.24)	-0.2421*** (-5.15)	0.0126 (0.41)	-0.1761*** (-3.47)	-0.0643 (-1.58)	-0.2214*** (-3.65)
Tertiary sector 1	0.2715*** (6.11)	0.0478 (0.90)	0.2441*** (7.89)	0.1881*** (3.57)	0.1399*** (3.32)	0.0566 (1.00)
Tertiary sector 2	-0.3277*** (-5.19)	-0.5988*** (-10.65)	-0.2826*** (-6.81)	-0.4999*** (-9.86)	-0.2517*** (-5.48)	-0.4963*** (-8.48)
Tertiary sector 3	0.2625*** (6.49)	0.0781 (1.25)	0.2089*** (5.54)	0.1713*** (2.63)	0.1308*** (3.10)	0.0676 (1.05)
Tertiary sector 4	0.0350 (0.76)	-0.1769*** (-3.42)	0.0140 (0.48)	-0.0330 (-0.56)	-0.0539 (-1.53)	-0.0839 (-1.36)
Tertiary sector 5	-0.0308 (-0.60)	-0.2225*** (-4.70)	-0.0878*** (-2.72)	-0.1146** (-2.03)	-0.1570*** (-3.91)	-0.1597*** (-2.92)
Tertiary sector 6	-0.3806*** (-5.12)	-0.5587*** (-7.59)	-0.6364*** (-10.75)	-0.5665*** (-8.54)	-0.7029*** (-6.14)	-0.6844*** (-10.16)
Public sector	0.3306*** (8.29)	0.6654*** (9.98)	0.1951*** (6.45)	0.5078*** (12.38)	0.0456 (1.49)	0.4257*** (8.39)
Formal private sector	0.2871*** (6.57)	0.4088*** (6.73)	0.2188*** (6.96)	0.3470*** (8.16)	0.1017*** (3.42)	0.3235*** (6.49)
Permanent employee	0.3192*** (6.23)	0.6408*** (11.22)	0.3240*** (8.67)	0.5568*** (10.28)	0.2473*** (6.42)	0.4076*** (7.51)
Temporary employee	-0.1290*** (-2.73)	0.0572 (1.21)	-0.1182*** (-3.25)	-0.0108 (-0.23)	-0.1358*** (-3.94)	-0.1014** (-2.12)

Table 2.9: Continued

Contract employee	-0.0201 (-0.39)	0.2132*** (3.21)	0.0516 (1.16)	0.1903*** (2.92)	0.0644 (1.46)	0.1785*** (2.81)
High-skilled white collar	0.6685*** (19.17)	0.4660*** (13.20)	0.6277*** (21.53)	0.4159*** (14.30)	0.6609*** (24.39)	0.4697*** (14.24)
Low-skilled white collar	0.1883*** (6.00)	0.0648** (2.10)	0.1794*** (7.22)	0.0731*** (2.77)	0.2080*** (8.22)	0.1142*** (4.35)
High-skilled blue collar	0.3048*** (9.80)	0.0410 (1.01)	0.2845*** (11.97)	0.0430 (1.21)	0.3291*** (12.03)	0.0677 (1.48)
Urban	0.3504*** (9.64)	0.2961*** (5.77)	0.2711*** (11.66)	0.1335*** (3.71)	0.2169*** (9.30)	0.0368 (0.96)
Tigray	-0.0544 (-1.42)	-0.0020 (-0.04)	-0.0345 (-1.15)	0.0109 (0.35)	-0.0291 (-0.94)	0.0123 (0.43)
Affar	0.2545*** (5.91)	0.1780*** (3.56)	0.1793*** (4.42)	0.1247** (2.53)	0.0890** (2.05)	0.0738 (1.36)
Amhara	-0.1911*** (-5.46)	-0.3936*** (-17.06)	-0.1816*** (-7.79)	-0.3209*** (-13.65)	-0.1646*** (-7.20)	-0.3231*** (-13.31)
Oromiya	-0.1784*** (-9.21)	-0.1574*** (-5.89)	-0.1517*** (-7.30)	-0.1676*** (-7.79)	-0.1364*** (-5.38)	-0.1584*** (-6.86)
Somalie	0.1205 (1.50)	0.1327 (0.96)	0.1539*** (2.79)	0.2168*** (5.48)	0.1623*** (3.60)	0.2886*** (4.29)
Benshangul Gumuz	0.0088 (0.26)	-0.0594 (-1.17)	-0.0472 (-1.19)	-0.0491 (-0.80)	0.0034 (0.07)	-0.0911** (-2.10)
SNNP	-0.2212*** (-9.52)	-0.2385*** (-6.70)	-0.2010*** (-8.27)	-0.2289*** (-8.56)	-0.1962*** (-9.88)	-0.2643*** (-9.45)
Harari	-0.0987* (-1.67)	-0.0228 (-0.45)	0.0127 (0.26)	0.0550 (1.09)	0.0184 (0.36)	0.0691 (1.18)
Dire Dawa	-0.0558 (-0.98)	-0.0542 (-0.90)	-0.1139*** (-2.79)	-0.0851 (-1.45)	-0.0956 (-1.39)	-0.0938 (-1.63)
Constant	-1.8800*** (-24.26)	-1.9907*** (-20.13)	-1.1379*** (-20.02)	-1.4035*** (-16.33)	-0.4082*** (-6.54)	-0.7639*** (-9.15)
Observations	12776	8616	12776	8616	12776	8616
Pseudo R2	0.3867	0.5042	0.3824	0.5342	0.3460	0.5029

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above; t-statistics in parenthesis. Bootstrapped standard errors.

\* Significant at 10% level, \*\* Significant at 5% level, \*\*\* Significant at 1% level.

**Table 2.10: OLS and selectivity corrected (BFG method) log hourly earnings equations in wage employment by gender and sector**

	Public sector				Formal private sector				Informal private sector			
	Men		Women		Men		Women		Men		Women	
	OLS	BFG	OLS	BFG	OLS	BFG	OLS	BFG	OLS	BFG	OLS	BFG
Primary education	0.1681*** (3.82)	0.0461 (0.63)	0.2247*** (4.13)	0.3255*** (3.63)	0.2334*** (5.76)	0.3176*** (5.43)	0.2403*** (4.90)	0.1792** (2.46)	0.2572*** (4.33)	-0.1769 (-0.83)	0.2776*** (9.49)	0.1106 (0.46)
General education	0.5004*** (10.19)	0.1570 (1.07)	0.5236*** (8.42)	0.7692*** (4.07)	0.4986*** (10.75)	0.6810*** (6.62)	0.6130*** (10.40)	0.4813*** (3.35)	0.5846*** (7.06)	-0.0802 (-0.20)	0.7701*** (11.73)	0.4593 (0.98)
Beyond general education	0.8609*** (16.20)	0.3373 (1.51)	0.7640*** (11.14)	1.0787*** (4.26)	0.8788*** (13.30)	1.1715*** (7.87)	1.0521*** (12.75)	0.8685*** (4.01)	0.8498*** (4.31)	-0.4000 (-0.55)	1.0969*** (5.52)	0.8293 (1.56)
Potential experience	0.0254*** (10.77)	0.0034 (0.45)	0.0287*** (8.05)	0.0437*** (3.36)	0.0286*** (9.73)	0.0241*** (3.00)	0.0273*** (5.95)	0.0239 (1.37)	0.0238*** (3.69)	0.0977*** (4.02)	0.0259*** (5.42)	0.0510** (2.19)
(Potential experience) <sup>2</sup>	-0.0003*** (-7.86)	-0.0000 (-0.11)	-0.0003*** (-3.86)	-0.0006*** (-2.68)	-0.0005*** (-9.64)	-0.0004*** (-2.64)	-0.0004*** (-3.83)	-0.0003 (-0.87)	-0.0003*** (-3.53)	-0.0017*** (-3.97)	-0.0002** (-2.33)	-0.0009 (-1.41)
Training	0.1588*** (7.72)	-0.0871 (-1.08)	0.1424*** (4.98)	0.2690*** (3.51)	0.3287*** (10.93)	0.2997*** (3.07)	0.1401*** (2.78)	0.0923 (0.65)	0.3540*** (4.72)	1.1671*** (4.44)	0.2498* (1.86)	0.3287* (1.68)
Married	0.0443** (2.21)	-0.1086* (-1.68)	0.0340 (1.51)	0.0005 (0.01)	0.1760*** (6.63)	0.2218*** (5.68)	0.1541*** (4.45)	0.1131 (0.87)	0.2745*** (4.66)	0.1650 (1.37)	0.3504*** (5.79)	-0.0901 (-0.22)
Manufacturing	0.0328 (0.63)	0.0345 (0.65)	0.0684 (0.97)	0.0622 (0.84)	0.0100 (0.21)	-0.0078 (-0.15)	-0.3026*** (-4.13)	-0.2975*** (-4.38)	0.0310 (0.33)	0.0349 (0.34)	-0.3056** (-2.05)	-0.2995* (-1.92)
Tertiary sector 1	0.0744 (1.39)	0.0779 (1.38)	0.1507** (1.99)	0.1506** (2.11)	0.2887*** (6.20)	0.2736*** (5.83)	0.0185 (0.24)	0.0259 (0.39)	0.3213*** (3.95)	0.3424*** (3.84)	0.2287 (1.38)	0.2255 (1.33)
Tertiary sector 2	-0.2518*** (-3.83)	-0.2444*** (-3.59)	-0.1370 (-1.52)	-0.1486 (-1.58)	-0.1949*** (-3.80)	-0.2106*** (-4.30)	-0.5139*** (-7.19)	-0.5089*** (-8.18)	-0.1008 (-1.01)	-0.0790 (-0.69)	-0.4489*** (-3.26)	-0.4507*** (-3.24)
Tertiary sector 3	0.0636 (1.23)	0.0722 (1.50)	0.2295*** (3.24)	0.2299*** (3.19)	0.2525*** (4.88)	0.2355*** (4.43)	0.0641 (0.66)	0.0670 (0.70)	—	—	—	—
Tertiary sector 4	-0.1157** (-2.38)	-0.1104** (-2.25)	0.0204 (0.31)	0.0209 (0.29)	-0.0742 (-0.93)	-0.0890 (-1.18)	-0.2472** (-2.43)	-0.2381*** (-2.65)	—	—	—	—
Tertiary sector 5	-0.1926*** (-3.90)	-0.1879*** (-3.70)	-0.0016 (-0.02)	-0.0083 (-0.12)	-0.1060** (-2.31)	-0.1218*** (-2.79)	-0.2528*** (-3.49)	-0.2469*** (-3.39)	0.1259 (1.39)	0.1419 (1.63)	-0.1316 (-0.89)	-0.1268 (-0.88)
Tertiary sector 6	—	—	—	—	—	—	—	—	-0.3813*** (-4.53)	-0.3540*** (-3.57)	-0.3843*** (-2.97)	-0.3808*** (-2.95)
Government employee	-0.0178 (-0.66)	-0.0186 (-0.74)	0.0746** (2.10)	0.0724** (2.06)	—	—	—	—	—	—	—	—
NGO employee	—	—	—	—	0.2916*** (7.26)	0.2851*** (6.81)	0.2152*** (3.58)	0.2166*** (3.29)	—	—	—	—
Domestic employee	—	—	—	—	—	—	—	—	-0.1631*** (-2.69)	-0.1694** (-2.31)	-0.1990*** (-2.79)	-0.2020*** (-2.85)
Permanent employee	0.6736*** (8.50)	0.6742*** (8.57)	0.7923*** (7.66)	0.7915*** (6.53)	0.1819*** (3.93)	0.1874*** (3.67)	0.5795*** (8.46)	0.5809*** (7.80)	-0.0427 (-0.48)	-0.0358 (-0.42)	0.0425 (0.54)	0.0379 (0.49)
Temporary employee	0.1340* (1.67)	0.1331 (1.58)	0.1676* (1.70)	0.1623 (1.53)	-0.1235*** (-2.94)	-0.1202*** (-2.62)	0.0128 (0.22)	0.0195 (0.29)	-0.2026*** (-3.37)	-0.2010*** (-3.07)	-0.1019* (-1.87)	-0.1022** (-2.30)
Contract employee	0.2665*** (3.09)	0.2677*** (3.13)	0.2742** (2.52)	0.2708** (2.27)	0.0051 (0.10)	0.0094 (0.19)	0.3420*** (4.70)	0.3509*** (4.30)	0.0808 (0.90)	0.0832 (0.86)	0.0950 (1.18)	0.0924 (1.29)

Table 2.10: Continued

High-skilled white collar	0.6925*** (25.72)	0.6866*** (22.50)	0.4812*** (12.19)	0.4765*** (11.58)	0.6624*** (14.58)	0.6610*** (13.34)	0.4356*** (6.57)	0.4348*** (5.96)	0.5052*** (3.10)	0.5136*** (2.83)	0.3334* (1.74)	0.3366 (1.57)
Low-skilled white collar	0.2875*** (10.86)	0.2844*** (10.12)	0.1880*** (5.30)	0.1913*** (5.58)	0.0422 (1.07)	0.0412 (1.12)	-0.1138*** (-2.74)	-0.1131** (-2.49)	-0.0167 (-0.18)	-0.0288 (-0.28)	-0.0023 (-0.04)	-0.0041 (-0.06)
High-skilled blue collar	0.2669*** (7.66)	0.2660*** (8.82)	-0.1498*** (-2.93)	-0.1475*** (-2.86)	0.3095*** (10.71)	0.3097*** (9.83)	0.0758 (1.42)	0.0777 (1.59)	0.2462*** (3.65)	0.2483*** (3.86)	0.1329* (1.74)	0.1269 (1.62)
Urban	0.2331*** (6.75)	0.0324 (0.52)	0.1174** (2.18)	0.1817*** (3.03)	0.4074*** (9.38)	0.2437** (1.99)	0.1893*** (2.66)	0.2316 (1.05)	0.1968** (2.24)	1.9138*** (3.87)	0.1295 (1.50)	0.6622 (1.23)
Tigray	-0.0268 (-0.76)	-0.2761** (-2.37)	0.1438*** (3.18)	0.2065* (1.78)	-0.0933** (-2.15)	-0.0283 (-0.42)	-0.0407 (-0.62)	-0.0930 (-0.91)	-0.1154 (-1.01)	-0.3562* (-1.89)	-0.1255** (-2.09)	-0.2685 (-1.60)
Affar	0.1704*** (4.40)	-0.2137 (-1.42)	0.2420*** (3.10)	0.3046*** (2.72)	0.1641*** (3.03)	0.2813*** (3.59)	0.0095 (0.08)	-0.0331 (-0.31)	0.0552 (0.32)	-0.4390 (-1.23)	0.1244** (1.98)	0.0158 (0.13)
Amhara	-0.1548*** (-6.15)	-0.3479*** (-3.81)	-0.0630* (-1.94)	-0.0775 (-1.35)	-0.2400*** (-6.25)	-0.1665** (-2.52)	-0.2370*** (-5.08)	-0.2625*** (-2.96)	-0.3355*** (-4.14)	-0.7084*** (-3.42)	-0.6998*** (-18.44)	-0.8230*** (-8.59)
Oromiya	-0.1058*** (-4.47)	-0.2773*** (-3.20)	-0.0021 (-0.06)	-0.0225 (-0.48)	-0.2065*** (-6.44)	-0.1731*** (-3.10)	-0.2278*** (-5.50)	-0.2502*** (-3.64)	-0.2360*** (-3.65)	-0.3413* (-1.87)	-0.3133*** (-8.48)	-0.4661*** (-3.32)
Somalie	-0.0339 (-0.76)	-0.2640** (-2.08)	0.0210 (0.33)	-0.0188 (-0.19)	0.3950*** (5.94)	0.5044*** (4.23)	0.2651** (2.12)	0.2175 (0.93)	0.2794** (2.23)	-0.4518 (-1.36)	0.4560*** (6.82)	0.1254 (0.44)
Benshangul Gumuz	-0.0196 (-0.54)	-0.3292** (-2.39)	0.1445*** (2.94)	0.1366 (1.20)	-0.0377 (-0.46)	0.0729 (0.69)	-0.3265** (-2.46)	-0.3835* (-1.86)	-0.2301 (-1.16)	-0.9018*** (-2.69)	-0.2816*** (-3.02)	-0.5107** (-2.24)
SNNP	-0.1286*** (-5.03)	-0.2939*** (-3.32)	-0.0876** (-2.22)	-0.1166** (-2.06)	-0.3181*** (-8.09)	-0.2595*** (-3.38)	-0.3761*** (-7.80)	-0.4027*** (-4.09)	-0.4526*** (-6.04)	-0.8090*** (-3.55)	-0.3571*** (-7.27)	-0.5432*** (-3.27)
Harari	-0.0226 (-0.43)	-0.2412** (-2.27)	0.1376** (2.48)	0.1459 (1.35)	-0.0205 (-0.27)	0.0336 (0.33)	0.0661 (0.45)	0.0245 (0.13)	-0.0089 (-0.06)	-0.1800 (-0.71)	-0.0320 (-0.38)	-0.1838 (-1.42)
Dire Dawa	-0.1324* (-1.92)	-0.2274** (-2.37)	0.0657 (0.91)	0.0499 (0.62)	-0.0703 (-1.22)	0.0013 (0.02)	-0.2846*** (-3.03)	-0.3098** (-2.36)	-0.1354 (-0.90)	-0.5523*** (-2.73)	-0.0939 (-1.12)	-0.2233 (-1.45)
Constant	-0.9740*** (-10.62)	0.9057 (1.40)	-1.2123*** (-10.89)	-1.8071*** (-2.83)	-1.0176*** (-14.29)	-1.1146*** (-3.70)	-0.9171*** (-8.41)	-0.8232 (-1.20)	-0.9051*** (-6.71)	-5.9480*** (-3.65)	-1.1513*** (-7.15)	-1.9821** (-2.20)
BFG not waged employed <sup>1</sup>	—	0.3198** (2.53)	—	-0.0068 (-0.11)	—	-0.0204 (-0.24)	—	0.0095 (0.28)	—	-0.0774 (-0.26)	—	0.0010 (0.18)
BFG public <sup>1</sup>	—	-0.4427*** (-2.96)	—	0.1724 (1.16)	—	0.0754 (0.22)	—	-0.3079 (-0.70)	—	2.9627*** (3.53)	—	0.4147 (0.61)
BFG formal private <sup>1</sup>	—	0.8580** (2.04)	—	0.5480* (1.76)	—	-0.0121 (-0.10)	—	0.0085 (0.03)	—	-0.0531 (-0.04)	—	0.9109 (1.08)
BFG informal private <sup>1</sup>	—	-1.3544*** (-2.69)	—	-0.1441 (-0.58)	—	-1.6208** (-2.43)	—	0.2491 (0.72)	—	1.9096*** (3.15)	—	0.3886 (0.93)
Observations	6030		3267		5190		2444		1556		2905	
R2	0.5801	0.5822	0.5258	0.5269	0.4444	0.4453	0.5050	0.5054	0.3129	0.3198	0.3476	0.3485

Source: Ethiopia LFS, 2005. Note: Individuals aged 15 and above; t-statistics in parenthesis. Bootstrapped standard errors in BFG regressions. \* Significant at 10% level, \*\* Significant at 5% level, \*\*\* Significant at 1% level.

<sup>1</sup>These variables are the four selection correction terms corresponding to the four alternatives.

## Appendix C. Gender earnings gap decompositions

**Table 2.11: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment (OLS estimates)**

	Neumark	Cotton
Gender mean observed log hourly earnings gap	<b>0.735</b>	
<b>First specification (accounting for job characteristics)</b>		
<i>Difference due to:</i>		
<b>Explained (%)</b>	<b>85.4</b>	<b>77.5</b>
Human capital characteristics	25.0	24.1
<i>Of which:</i>		
Education	14.4	13.4
Experience	6.1	6.2
Training	4.5	4.4
Job characteristics	56.4	49.5
<i>Of which:</i>		
Sector of activity	29.9	23.7
Sector of wage employment	7.7	8.3
Terms of employment	8.3	8.4
Occupation	10.5	9.2
Other observable characteristics	4.0	3.8
<b>Unexplained (%)</b>	<b>14.6</b>	<b>22.5</b>
<i>Of which:</i>		
Men's treatment advantage	5.9	6.8
Women's treatment disadvantage	8.7	15.8
<b>Second specification (not accounting for job characteristics)</b>		
<i>Difference due to:</i>		
<b>Explained (%)</b>	<b>62.0</b>	<b>55.4</b>
Human capital characteristics	48.9	44.6
<i>Of which:</i>		
Education	28.9	25.8
Experience	11.8	11.3
Training	8.2	7.5
Other observable characteristics	13.1	10.8
<b>Unexplained (%)</b>	<b>38.0</b>	<b>44.6</b>
<i>Of which:</i>		
Men's treatment advantage	15.3	10.1
Women's treatment disadvantage	22.7	34.5

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above. Positive sign indicates advantage to males and negative sign indicates advantage to females.



**Table 2.12: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment by age cohort (OLS estimates)**

15-24				25-34		35+		
Neumark		Cotton	Neumark		Cotton	Neumark		Cotton
Gender mean observed log hourly earnings gap			0.758		0.497		0.340	
First specification (accounting for job characteristics)								
Difference due to:								
Explained (%)			83.1	76.8	66.3	53.0	83.5	73.8
Human capital characteristics			13.4	13.0	15.3	13.4	29.6	27.4
Of which:								
Education			12.9	12.5	11.2	9.5	22.6	20.1
Experience			-2.3	-2.2	-0.6	-0.5	-2.2	-2.2
Training			2.8	2.7	4.6	4.4	9.1	9.5
Job characteristics			72.8	66.7	48.9	38.1	42.6	36.8
Of which:								
Sector of activity			47.9	41.8	26.3	17.9	25.2	21.1
Sector of wage employment			10.0	12.4	6.6	6.8	3.9	2.7
Terms of employment			4.1	4.0	5.4	5.3	2.8	2.7
Occupation			10.7	8.5	10.7	8.1	10.7	10.3
Other observable characteristics			-3.0	-2.9	2.1	1.4	11.3	9.6
Unexplained (%)			16.9	23.2	33.7	47.0	16.5	26.2
Of which:								
Men's treatment advantage			9.4	14.2	13.5	18.2	4.6	7.3
Women's treatment disadvantage			7.5	9.1	20.1	28.8	12.0	18.9
Second specification (not accounting for job characteristics)								
Difference due to:								
Explained (%)			34.9	27.8	36.2	30.0	77.2	69.5
Human capital characteristics			30.5	26.1	28.1	24.0	53.6	50.5
Of which:								
Education			26.6	22.9	20.3	17.0	44.0	41.5
Experience			-2.6	-2.3	-0.5	-0.3	-2.6	-3.0
Training			6.4	5.5	8.2	7.3	12.2	12.0
Other observable characteristics			4.4	1.7	8.1	6.0	23.5	18.9
Unexplained (%)			65.1	72.2	63.8	70.0	22.8	30.5
Of which:								
Men's treatment advantage			36.3	37.4	25.7	24.9	6.3	0.9
Women's treatment disadvantage			28.8	34.8	38.1	45.1	16.5	29.6

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above. Positive sign indicates advantage to males and negative sign indicates advantage to females.

**Table 2.13: Neumark and Cotton decompositions of the gender predicted log hourly earnings differential in wage employment by quartile**

First quartile				Median		Third quartile	
Neumark		Cotton		Neumark		Cotton	
First specification (accounting for job characteristics)							
Gender predicted log hourly earnings gap	0.701		0.742		0.766		
Difference due to:							
Explained ( %)	87.8	80.4	84.4	78.6	82.6	73.0	
Human capital characteristics	26.8	25.5	24.6	23.7	23.7	22.5	
Of which:							
Education	16.4	14.9	14.2	13.4	12.7	11.8	
Experience	6.1	6.2	6.1	6.2	6.7	6.6	
Training	4.2	4.4	4.3	4.1	4.2	4.2	
Job characteristics	58.0	52.4	56.4	51.6	55.7	47.0	
Of which:							
Sector of activity	26.3	21.6	30.1	26.1	34.1	25.4	
Sector of wage employment	12.2	12.0	8.1	8.5	3.9	5.5	
Terms of employment	9.2	9.2	8.7	8.6	7.6	7.5	
Occupation	10.4	9.6	9.4	8.3	10.2	8.6	
Other observable characteristics	3.1	2.5	3.4	3.3	3.2	3.4	
Unexplained ( %)	12.2	19.6	15.6	21.4	17.4	27.0	
Of which:							
Men's treatment advantage	5.9	6.4	6.3	7.2	6.4	8.6	
Women's treatment disadvantage	6.3	13.2	9.3	14.2	11.0	18.4	
Second specification (not accounting for job characteristics)							
Gender predicted log hourly earnings gap	0.775		0.771		0.704		
Difference due to:							
Explained ( %)	64.8	56.1	61.3	53.6	56.9	55.2	
Human capital characteristics	49.8	44.0	49.4	43.8	47.9	46.8	
Of which:							
Education	29.9	26.0	29.4	25.2	28.1	26.4	
Experience	11.0	9.7	12.0	11.4	12.7	13.7	
Training	8.9	8.2	8.0	7.3	7.0	6.6	
Other observable characteristics	15.0	12.1	11.9	9.8	9.0	8.4	
Unexplained ( %)	35.2	43.9	38.7	46.4	43.1	44.8	
Of which:							
Men's treatment advantage	18.6	11.0	16.8	10.3	12.5	8.6	
Women's treatment disadvantage	16.6	32.9	21.9	36.1	30.6	36.2	

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above. Positive sign indicates advantage to males and negative sign indicates advantage to females.

**Table 2.14: Neumark and Cotton decompositions of the gender mean log hourly earnings differential in wage employment by sector (OLS estimates)**

Public sector				Formal private sector		Informal private sector		
Neumark		Cotton	Neumark		Cotton	Neumark		Cotton
Gender mean observed log hourly earnings gap			0.202		0.426		1.022	
First specification (accounting for job characteristics)								
Difference due to:								
Explained (%)		81.2	81.2	57.3		45.8	83.2	70.8
Human capital characteristics		39.6	40.5	8.3		6.6	20.6	20.1
Of which:								
Education		14.7	14.8	-4.9		-5.4	12.6	12.0
Experience		18.2	18.5	11.0		9.6	3.7	4.5
Training		6.8	7.3	2.3		2.3	4.3	3.6
Job characteristics		41.6	39.3	42.0		33.5	52.4	41.4
Of which:								
Sector of activity		6.1	5.5	22.9		20.2	33.8	25.3
Employment status		0.0	0.0	-0.5		-0.5	9.8	8.3
Terms of employment		12.6	12.4	3.3		3.3	4.0	3.5
Occupation		23.0	21.3	16.3		10.6	4.9	4.3
Other observable characteristics		0.0	1.4	7.0		5.7	10.2	9.4
Unexplained (%)		18.8	18.8	42.7		54.2	16.8	29.2
Of which:								
Men's treatment advantage		6.6	7.9	13.7		15.4	10.9	19.1
Women's treatment disadvantage		12.1	10.9	29.0		38.8	5.8	10.0
Second specification (not accounting for job characteristics)								
Difference due to:								
Explained (%)		70.4	72.3	23.7		17.5	54.7	40.4
Human capital characteristics		63.5	63.9	11.2		7.2	30.8	25.1
Of which:								
Education		27.0	27.4	-8.2		-9.3	18.7	14.8
Experience		25.0	25.0	16.0		13.5	5.7	5.4
Training		11.4	11.6	3.4		3.0	6.3	5.0
Other observable characteristics		6.9	8.3	12.5		10.2	23.9	15.3
Unexplained (%)		29.6	27.7	76.3		82.5	45.3	59.6
Of which:								
Men's treatment advantage		10.4	6.3	24.4		24.2	29.5	35.8
Women's treatment disadvantage		19.2	21.5	51.9		58.3	15.8	23.7

Source: Ethiopia LFS, 2005.

Note: Individuals aged 15 and above. Positive sign indicates advantage to males and negative sign indicates advantage to females.

## **Chapter 3**

### **Analysing the Nature and Extent of Gender Inequalities in Time Use:**

#### **New Insights from Ethiopia<sup>1</sup>**

##### **3.1 Introduction**

Work-related gender equality remains a matter of concern in Ethiopia. Indeed, despite the fact that the Civil Code, and the Constitution adopted in 1994, provide women with the same rights and protections as men, previous works on Ethiopia's labour market reveal the striking disadvantages that Ethiopian women face. For instance, a comprehensive report on Ethiopia's urban labour market (World Bank 2007) find, as a major result, that women typically experience worse outcomes than men in the labour market, with higher levels of unemployment, lower wages, and a greater concentration in the informal sector. Empirical evidence shows that women are less likely to participate in the labour market and, when employed, they are disproportionately concentrated in unpaid or flexible jobs that offer low earnings and low protection. This highlights a combination of factors, including lower skills, an increased burden of household responsibilities, and labour market discrimination. For instance, recent studies find that, even when they have the same human capital and job characteristics as men, women earn on average much less (See Chapter 2 in this thesis and, Appleton, Hoddinott, and Krishnan 1999; Temesgen 2006; Kolev and Suárez Robles 2010a).

Using the Ethiopia Labour Force Survey 2005, this Chapter sheds new light on a key dimension of work-related gender disparities in Ethiopia: time allocation between market and household work. As far as we know, there are almost no studies on these topics focusing on Ethiopia. The main aim and the novelty of the present research is to help fill the gap in current knowledge about

---

<sup>1</sup> This Chapter draws on Suárez Robles (2010).

the nature, the extent, and the determinants of these specific gender disparities in the Ethiopian context. In particular, the research question this Chapter attempts to address is whether and to what degree there is complementarity or substitutability between the two types of work. In other words, we are interested in understanding the connection between market and household work time allocation decisions. We assess the impact of each type of work on each other in order to know to what extent these decisions are interrelated. We assume that the hours spent in each type of work actually do not represent separate decisions but rather are outcomes of an optimization process in which allocations of time to market and household work are jointly determined.

The issue of gender inequalities in allocating time to market work (productive activities) and housework (reproductive activities) has been insufficiently treated in the literature on developing countries in general, and regarding Ethiopia in particular. This is because data on time use is scarce and because most gender studies focusing on labour supply ignore the trade-offs between these two types of work and neglect the dimension of housework. Besides contributing to a better understanding of the labour supply, extending the analysis to this dimension of work is important because household tasks, which are for the most part low productive, time- and labour-intensive, are essential for family survival. Empirical evidence shows that these tasks are primarily carried out by women. As pointed out by Blackden and Morris-Hugues (1993), the housework burden on women limits their time available to market work and allows them to engage only in productive activities compatible with their household duties. According to Blackden and Wodon (2006), the gender-based division of labour, which is characterized by the fact that men are engaged in productive activities while women bear the brunt of domestic tasks, is especially significant in Africa.

However, some studies show that there is not a perfect substitution between market and household work. For instance, Medeiros, Guerreiro Osorio, and Costa (2007), using data from urban Bolivia, argue that such a trade-off is only partial. Women tend to accumulate both types of work; therefore, they are double burdened and have a higher total workload than men. Because

women suffer from time deprivation as a result of their multiple roles, they enjoy less leisure and time for rest, and thereby are more likely to be “time poor”.

What is the situation in Ethiopia? The specific objective of this study is to examine the differences in how Ethiopian men and women allocate their time between market and household work, identify the gender-based division of labour, explore the gender disparities in total workload, and analyse the determinants of market and household work time across gender. As said before, our primary interest, that is the research question we seek to answer, is whether and to what degree market and household work are substitutes or complements in Ethiopia.

The structure of this Chapter is as follows. Section 2 presents a brief theoretical literature review, the dataset used in this study, as well as the definition and measurement of time use variables. The methodology used to decompose total work time and analyse the determinants of market and household work time across gender, are explained in Section 3. Section 4 presents the results and Section 5 concludes.

## **3.2 Data and concepts**

### **3.2.1 Brief theoretical literature review**

Theoretical and empirical literature on the allocation of time is dense and heterogeneous. The purpose of this section is to briefly document the main theoretical literature on the subject, but not to provide an extensive and exhaustive literature review, which is beyond the scope of this study. The earliest seminal work on the allocation of time was that of Becker in 1965, who proposed, in the same vein of Mincer (1962), a model for household production in which households, who are assumed to be producers as well as consumers, produce commodities, by combining inputs of goods and time according to the cost-minimization rules of the traditional firm theory, in quantities determined by maximizing a utility function of the commodity set subject to prices and

a constraint on resources. Resources are measured by what is called full income, which is the sum of money income and that forgone or "lost" by the use of time and goods to obtain utility, while commodity prices are measured by the sum of the costs of their goods and time inputs (Becker 1965). In other words, full income is the income that could be generated by a household devoting all its time to the objective of earning income. The time allocation chosen by households, through a utility-maximizing process, necessarily represents the best allocation, given the wage rates its members can command (Burchardt 2008).

Becker's original model is not exempt from criticisms. For example, it does not take sufficient account of the role institutions (availability of social services including childcare, cultural and social norms, etc.) have in shaping households' time allocation decisions (Folbre 2004). In addition, the assumed production technology is too simplistic, and the goods produced by households (such as warmth and nutrition) are not observable on markets (Pollack and Wachter 1975). Later work has moved away from the abstractness of the definition of household production activities in Becker, by recognizing that these activities typically have close but not perfect market substitutes and, in common with market production, have inputs and outputs that are inherently just as observable, although with varying degrees of difficulty (Gronau 1977; Apps and Rees 1988, 2002). One of the major and most criticized shortcomings of Becker's model is the treatment of the household as unit. Much of the relatively new literature on the intra-household allocation of resources adopts an approach, referred to as the "collective" model, that omits household production altogether. The "collective model" of the household originates with the papers of Manser and Brown (1980), McElroy and Horney (1981), and Chiappori (1988), who analyse the household consumption allocation as the outcome of Nash bargaining, with threat points given either by what the partners could achieve by leaving the household, or, alternatively, in some later formulations, by the non-cooperative Nash equilibrium with both partners remaining within the household.

Many specializations of Becker's household production model have seen the light over the past decades. Overall, household production models are more ambitious than conventional static labour supply models in that labour supply decisions can be connected to the technology of home production, which evolves over time. In fact, evolution of household technology ease the substitution of market goods for own time, resulting in a greater responsiveness of labour supply to wages rates. For example, if new methods for market-provided childcare become available, and these increase the opportunity for substitution of own time, we can predict from the model an increased labour supply elasticity (Juster and Stafford 1991).

The literature also includes dynamic life cycle models which synthesize the market work and training/education decision with choice of nonmarket time, with an emphasis on intertemporal links through the accumulation of marketable skills (Ghez and Becker 1975; Blinder and Weiss 1976; Ryder, Stafford and Stephan 1976). Time spent in today's skill acquisition costs lost wages and lost utility from leisure, but it raises future wages and thereby produces greater future consumption opportunities.

### **3.2.2 Ethiopia Labour Force Survey 2005**

To analyse gender inequalities in allocating time to market and household work and to address the issue of complementarity/substitutability of time allocation decisions, we draw upon the Labour Force Survey (LFS) collected in Ethiopia by the Central Statistical Agency (CSA) in March 2005<sup>2</sup>. The sample survey covered 54,443 households in all rural and urban parts of the

---

<sup>2</sup> Poorest economies like Ethiopia are generally characterized by a preponderance of seasonal activities like agriculture, and even their small manufacturing sectors tend to be more than usually seasonal. Individuals' time use can widely vary throughout the year due to seasonal temperature, rainfall and day length variations, seasonality of demand, religious observances...In general, there are two rainy seasons per year in Ethiopia: the *Meher* (June-August), which is the main rainy season, and the *Belg* (March). Planting and the main harvesting periods, which usually take place few months after the rainy season, are peak times for labour. The fact that data collection took place in a short period of time, during March 2005, is clearly a limitation of the survey. Whether individuals' time use in that period was more or less representative for the year as a whole, or was characterized by high underemployment or peaks of labor is difficult to assert at



country, except all zones of Gambela region (excluding Gambela town), and the non-sedentary population of three zones of the Afar and six zones of the Somali regions. Out of the total 230,680 individuals interviewed, about 50 per cent were located in urban areas; women account for 52 per cent of the sample.

For household members aged 5 and older, the survey records the number of hours spent, during the seven days prior to the date of interview, in collecting firewood and fetching water for own consumption and in domestic activities. In addition, for those who were engaged in any kind of productive activities during the reference period, the survey records, excluding lunch and journey time and including overtime, the total number of hours spent in all jobs.

This information gathered in the LFS 2005 on how surveyed individuals allocate their time is not as detailed as that in surveys using time-diary methods. Indeed, the time diary approach can provide information on individuals' allocation of time among large number of activities, as well as the sequence of these activities and the context in which they are performed. Moreover, as time-diary methods ask respondents to fill in activities and time used in the present or the past day in slots of a diary, the diary-based estimates contain, to a lesser extent, recall biases. Such biases are related to the difficulty for a respondent to remember and report exactly the amount of time used. In addition, time-diary methods do not produce total time used greater than the total time available in the reference period. Stylized (questionnaire-based) estimates, like those derived from the LFS 2005, are, therefore, less accurate than diary-based estimates. Moreover, the data used in this study does not include information on some time-intensive household activities (such as care of family members), and thus housework is most probably underestimated, especially for women, who bear the brunt of domestic tasks. However, given that the survey questionnaire does not provide a list of domestic activities, it is impossible to know exactly what activities were taken into account by respondents. According to Fenstermaker, Berk and Shih (1980), Press and

---

the country scale. Ethiopia is characterized by diverse climates which translate into diverse vegetation zones (according to the Koeppen-Geiger climate classification system, Ethiopia has 10 climate types).

Townsley (1998), and Warner (1986), women likely report their time spent in housework more accurately than men, since they undertake the bulk of housework and therefore should be more familiar with how much time each chore has cost them. Furthermore, stylized estimates of housework tend to be higher than diary-based estimates, which is in part attributable to the double counting of multiple activities performed simultaneously (Juster and Stafford 1991; Marini and Shelton 1993). This bias is expected to affect women more than men because women perform various domestic activities simultaneously more often (Lee and Waite 2005)<sup>3</sup>.

Nationally representative time-diary and time use surveys, as well as household surveys including a time use module, have not been carried out in Ethiopia. Notwithstanding its shortcomings, the LFS 2005 offers a unique opportunity to analyse gender-differentiated time-use patterns and to address the issue of substitutability or complementarity between market and household work time allocation decisions.

### **3.2.3 Definitions and measurement issues**

In this study, time-use estimates are measured in hours per week. Individuals allocate the *total time available* (168 hours a week) between tertiary activities, leisure, and work, which includes market and household work.

The *time devoted to tertiary activities* is the time needed to maintain a person's biological functioning (time spent for rest, personal hygiene, and nourishment)<sup>4</sup>. The LFS 2005 does not provide information on tertiary activities. According to Ting and Malhotra (2005), on average an adult needs to sleep eight hours per day. Therefore, the time devoted to tertiary activities should be higher in order to take into account other self-care activities. However, following Medeiros,

---

<sup>3</sup> For a detailed discussion about differences between stylized (questionnaire-based) estimates and diary-based estimates of housework time collected from the same respondents, see Kan (2006).

<sup>4</sup> Following Burda, Hamermesh, and Weil (2007), those activities that we cannot pay other people to do for us, but that we must do at least some of, are defined as tertiary activities.

Guerreiro Osorio, and Costa (2007), the time devoted to tertiary activities is set at 8 hours per day (56 hours a week) for all individuals to deal with extreme workload.

Subtracting the time spent for tertiary activities from the total time available, we obtain the *total time available for work* (112 hours a week), which is allocated between market and household work, and leisure.

The *time devoted to market work* corresponds to the number of hours worked at all jobs, excluding lunch and journey time and including overtime, and the *time devoted to housework* equals the total number of hours worked in fetching water and collecting firewood for own consumption, and in domestic activities. We assume that such activities are not market oriented, since market work includes all productive activities. All individuals who did not report the number of hours worked at all jobs, or at least in one of the housework activities, or who have a total workload that exceeds 112 hours a week, are dropped from the sample.

Finally, *leisure* is defined as all dispensable activities that we cannot pay somebody else to do for us, and corresponds by default to total time available for work minus *total work time* (leisure is thus treated as a residual).

The survey divided the population aged 10 years and older into economically active and inactive categories. This low age limit was fixed to allow comparisons with other countries and to incorporate information about these children, who often start taking part in many types of economic activities at young ages (CSA 2006). However, in line with other recent studies on Ethiopia's labour market, this study focuses on individuals aged 15 and older<sup>5</sup>.

---

<sup>5</sup> According to World Bank (2007), including the age-group 10-15 in the working-age population creates an implicit conflict between employment and education policy objectives. ILO definitions of employment, unemployment and labour force make reference to persons "above a specified age". According to the ILO, this minimum age limit needs to be specified by each country according to its national circumstances such as the compulsory schooling age, minimum age for admission to employment, and extent of child labour. However, The ILO acknowledges that these circumstances vary so greatly among countries that it is impossible to specify any universally applicable minimum age limit at the international level. In practice, in most ILO studies the minimum age limit for defining the working-age population is set at 15.

Out of the total 134,714 surveyed individuals aged 15 and older, 373 did not report the number of hours worked at all jobs or in, at least, one of the housework activities, and among those who completely reported their time use, 2 897 have a total workload that exceeds 112 hours a week. Excluding all these individuals reduces the sample to 131,444 individuals, which represents almost 98% of the sample population of individuals aged 15 and older.

### 3.3 Methodology

#### 3.3.1 Decomposition of total work time

The following decomposition of the average total work time is performed separately for men and women:

$$\bar{T} = \bar{T}_{mw} + \bar{T}_{hw} = \rho_{mw} \mu_{mw} + \rho_{hw} \mu_{hw} \quad (1)$$

Where:

$$\bar{T}_{hw} = \sum_{l=1}^3 \bar{T}_{hw_l} = \sum_{l=1}^3 \rho_{hw_l} \mu_{hw_l} = \rho_{hw} \mu_{hw} \quad (2)$$

The average total work time  $(\bar{T})$  corresponds to the sum of the averages of time devoted to market work  $(\bar{T}_{mw})$  and housework  $(\bar{T}_{hw})$ , and is equivalent to the sum of the products of the incidence  $(\rho)$  and the duration  $(\mu)$  of each type of work. The incidence of a given activity is defined as the share of the population engaged (who spent at least one hour) in this activity, and the duration is the average amount of hours devoted to the activity by those who perform it. The average of time devoted to housework is calculated as the sum of the averages of time spent fetching water

$(l = 1)$ , collecting firewood ( $l = 2$ ), and in domestic activities ( $l = 3$ ), and is also decomposed in terms of the incidence and duration of each type of housework.

The results of this decomposition obtained for men are compared to those obtained for women to identify the gender inequalities in allocating time to market and household activities. Gender differences in the average total work time are explored as well. This simple decomposition is interesting because it enables an assessment of gender disparities in both incidence and duration of each activity. Therefore, a comprehensive picture of the gender-based division of labour, which is characterized not only by the fact that some activities are primarily carried out by women while others are male-dominated, but also by the fact that the amount of time spent in each activity significantly varies across gender, can be provided.

### **3.3.2 The determinants of market and household work time**

Finally, we investigate the determinants of market and household work time across gender and place of residence. We estimate the hours worked in each type of work, separately for men and women, and, urban and rural areas. These equations include, as regressors, individual characteristics, including human capital characteristics, variables that reflect household composition, and location variables as controls. Because time is a limited resource and decisions on time spent in each type of work may influence each other, we also include the housework hours in the market work time equations, and the market work hours in the housework time equations. In this analysis, we are particularly interested in understanding the connection between market and household work time allocation decisions. We want to know whether and to what degree there is substitutability or complementarity between the two types of work. For this, we assess the impact of each type of work on each other in order to determine to what extent time allocation decisions are interrelated.

A significant proportion of men and women did not participate in at least one of these two types of work during the reference period. The dependent variables are therefore left-censored at value zero. Estimating a linear regression in the presence of censoring involves additional computational complications. Ordinary least-squares (OLS) regression will not yield consistent parameter estimates because the censored sample is not representative of the population. It is then preferable to use Tobit models which are relevant when the dependent variable of a linear regression is not always observed. For housework, we use a standard Tobit model to take account of the fact that some individuals do not report housework. And for market work, in order to correct for potential selection bias, and to discriminate between the impact of covariates on participation in market work during the reference period and working hours given participation, we use a Generalized Tobit model<sup>6</sup>. Given that in these models the estimated coefficients have no natural interpretation, we report marginal effects evaluated at sample means.

If our primary interest is to understand the connection between market and household work time allocation decisions, the two types of models proposed above may be misleading to some extent. Indeed, the hours spent in each of these activities actually do not represent separate decisions but rather are outcomes of an optimization process in which allocations of time to market and household work are jointly determined. This calls for some sensitivity analysis and the use of a more appropriate econometric model. To account for the fact that decisions are taken simultaneously and to deal with endogeneity issues, we further model hours of market work and hours of housework as a simultaneous system using the simultaneous Tobit model, following Amemiya (1974, 1979). Concretely, we use Amemiya's generalized least squares (AGLS) method, which corresponds to the estimation of an instrumental variable Tobit model, that is a

---

<sup>6</sup> For a comparison and evaluation of models and estimators appropriate for time-use data using Monte Carlo simulations, see Flood and Grasjo (2001), and Nawata (2007). Basically, the most commonly used models are the followings: the standard Tobit (type I) model, which is a special case of the Heckman's (1978) generalized Tobit model (Tobit type II), and the double-hurdle model derived by Jones (1992), which can be regarded as an extension of the Tobit type II. Heckman (1979) also suggested a two-stage method (Heckit), instead of using maximum-likelihood estimation.

Tobit model in which one or several regressors are endogenously determined. Unfortunately, the LFS 2005 provides no variables which could reasonably be used as valid instruments for market and household work hours. As a consequence, we are unable to produce accurate and reliable results using this econometric technique. Therefore, we rely only on the results obtained with the standard and the generalized Tobit methods.

In the following of this section, we describe these two models, and then we present, for market work, the method chosen to decompose the total marginal impact into an effect on participation and an effect on working time given participation.

### ***A generalized Tobit model for market work time***

Heckman's (1978) generalized Tobit model (Tobit type II), consists of a structural equation (preferred market work time function), an index equation (market work participation), a threshold equation linking preferred and observed market work time, and, finally, a stochastic specification.

$$\text{Structural equation:} \quad y_i^* = x_{1i}\beta_1 + \varepsilon_i \quad (3)$$

$$\text{Index equation:} \quad d_i^* = x_{2i}\beta_2 + v_i \quad (4)$$

$$\text{Threshold index equation:} \quad d_i = \begin{cases} 1 & \text{if } d_i^* > 0 \\ 0 & \text{if } d_i^* \leq 0 \end{cases} \quad (5)$$

$$\text{Threshold structural equation:} \quad y_i = \begin{cases} y_i^* & \text{if } d_i = 1 \\ 0 & \text{else} \end{cases} \quad (6)$$

$$\text{Stochastic specification:} \quad \varepsilon_i v_i \sim N(0, 0, \sigma^2, 1, \rho) \quad (7)$$

$y_i^*$  denotes the latent (non-observed) endogenous variable, the preferred hours of market work, and  $y_i$  denotes the corresponding observed variable (measured hours of market work).  $x_{1i}$  and  $x_{2i}$

are vectors of explanatory variables, which are assumed to be uncorrelated with the error terms  $\varepsilon_i$  och  $v_i$ .  $\beta_1$  and  $\beta_2$  are vectors of parameters.  $d_i^*$  is a latent variable that represents binary censoring and  $d_i$  is the observed value (1 if the individual reports market work, else 0).

Given the stochastic specification, the likelihood function can be derived as:

$$L = \prod_{y=0} \Phi(-X_2 \beta_2) \prod_{y>0} \left\{ \Phi \left( \frac{X_2 \beta_2 + \frac{\rho}{\sigma}(y - X_1 \beta_1)}{\sqrt{1-\rho^2}} \right) \frac{1}{\sigma} \phi((y - X_1 \beta_1)/\sigma) \right\} \quad (8)$$

Where  $y = 0$  denotes the individuals with zero hours of market work, and  $y > 0$  the individuals with positive hour,  $\Phi$  and  $\phi$  denote the univariate cumulative distribution function (CDF), and probability distribution function (PDF) of the standard normal.

The parameters of the model are estimated using maximum likelihood. These are consistent under the stated assumptions, that is, if the error terms are homoskedastic and normally distributed.

The interpretation of the estimated coefficients is not straightforward. It is thus necessary to calculate the marginal effects. These are based on the expected value of the observed hours of market work, which is equal to:

$$E(y_i) = P(d_i^* > 0)E(y_i|d_i^* > 0) = \Phi(X_{2i}\beta_2) \left[ X_{1i}\beta_1 + \sigma \frac{\phi(X_{2i}\beta_2)}{\Phi(X_{2i}\beta_2)} \right] \quad (9)$$

The marginal effects are defined as the derivate of  $E(y)$  with respect to the variables in  $X_1$ :

$$\frac{\partial E[y_i]}{\partial X_{1j}} = \Phi(X_{2i}\beta_2)\beta_{1j} + \delta [\phi(X_{2i}\beta_2)\beta_{2j}(X_{1i}\beta_1 - \sigma X_{2i}\beta_2)] \quad (10)$$

Where  $\delta = 1$  if  $X_{1j}$  is included in  $X_2$ , and else  $\delta = 0$



The extent of these effects depends on the choice of an observation point in the sample. In this study, all effects are evaluated at the sample means of  $X_1$  and  $X_2$ , as it is usually done in the literature. It is noteworthy that these are solely representative for the mean point of the sample, and they cannot be generalized to the whole sample.

### ***A standard Tobit model for housework time***

The standard Tobit (Type I) model, devised by Tobin (1958), is a special case of the generalized Tobit model. It is obtained by dropping the index equation and modifying the threshold function to:

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0 \\ 0 & \text{else} \end{cases} \quad (11)$$

The likelihood function can be derived as:

$$L = \prod_{y=0} \{1 - \Phi(X_1\beta_1/\sigma)\} \prod_{y>0} \left\{ \frac{1}{\sigma} \phi((y - X_1\beta_1)/\sigma) \right\} \quad (12)$$

Finally, we can define the expected value of the observed hours of housework and the marginal effects as follows:

$$E(y_i) = P(y_i^* > 0)E(y_i|y_i^* > 0) = \Phi(X_{1i}\beta_1) \left[ X_{1i}\beta_1 + \sigma \frac{\phi(X_{1i}\beta_1)}{\Phi(X_{1i}\beta_1)} \right] \quad (13)$$

$$\frac{\partial E[y_i]}{\partial X_{1j}} = \Phi(X_{1i}\beta_1)\beta_{1j} \quad (14)$$

The standard Tobit model is more restrictive than the generalized Tobit model because it does not differ between the participation decision and the structural equation. However, specifying a reasonable model for the decision to participate in housework is not straightforward in practice.

### ***Decomposition of the total marginal effect***

McDonald and Moffitt (1980) show that the total marginal effect can be disaggregated into two parts:

$$\frac{\partial E(y)}{\partial x_1} = \frac{\partial p(d=1)}{\partial x_1} E(y|d=1) + \frac{\partial E(y|d=1)}{\partial x_1} p(d=1) \quad (15)$$

The first term on the right-hand side of equation (15) reflects the change in the probability of participating in market work given a change in the exogenous variable  $x_1$ , weighted by the expected value of market work time. The second term reflects the impact of the covariate  $x_1$  on market work time for those participating, weighted by the probability of participating in market work. With this decomposition we are then able to discriminate between the impact of covariates on participation in market work during the reference period, and the impact of covariates on market work time given participation.

This decomposition procedure is widely applied in the literature. However, it is not exempt of criticisms. Kang (2007) highlights two limitations of this decomposition. First, the decomposition ratio of the two effects depends on the choice of an observation point in the sample. In this study, all effects are evaluated at sample means. It is important to recall that a decomposition at the mean point of the sample cannot be generalized to the whole sample. Second, once the observation is fixed, the decomposition ratio is the same across all coefficients or across all predictors. Actually, this is true for the standard Tobit model but not for the generalized Tobit model. In the latter, the decomposition of the total marginal effect is, indeed, not assumed to be

constant for all exogenous variables. This limitation does therefore not apply in our case, since total marginal effects are decomposed only for market work time after estimation of a generalized Tobit model.

### **3.4 Results**

How do men and women allocate their time between market and household work? Do they perform the same amount of total work? What are the determinants of market and household work time across gender? Is there substitutability or complementarity between the two types of work? To what extent? Using the LFS 2005, this section attempts to answer these questions. All descriptive statistics and regression results are obtained using sampling weights so they are representative of the Ethiopian population as a whole, except for those located in the zones of Gambela, Afar and Somali regions not covered by the survey. Recall that in this study we focus on the working-age population, i.e. individuals aged 15 and above.

#### **3.4.1 Decomposition of total work time**

Table 3.1 shows the decomposition results of the average total work time for the overall sample population, as well as separately for men and women. Individuals aged 15 and older work 43 hours per week on average, which represents 26 per cent of the total time available (168 hours [hr] a week), and 39 per cent of the total time available for work (112 hr a week). They spend on average more time on market work (22 hr) than on household activities (21 hr). As compared to housework, the incidence of market work is higher while the average duration is lower, which means that there are more individuals working in the labour market (73 per cent) where they experience shorter average work shifts (30 hr), than in the household (66 per cent), where the average duration of work is higher (32 hr). About 42 per cent of individuals aged 15 and older are not engaged at all in domestic activities. According to this and to the fact that most of these

activities are time-intensive, it is not surprising that the average hours dedicated to domestic activities by those who perform it rises to 28 hours per week, which, for example, is nearly as high as the average duration of market work. Less than a third of the sample population collect firewood, and no more than 40 per cent fetch water for own consumption. The average duration of both activities is about 7 hours per week. These figures on water and firewood collection show only the peak of the iceberg and mask an even more cruel reality. Though the analysis of child labour is beyond the scope of this study, it is worth precising here that, in Ethiopia, water and firewood collection are the most common child activities. Many children are commonly involved in these activities at the expense of attending school or doing homework. Beyond poverty and low education, cultural norms also have an impact on child labour since these specific tasks are culturally designated as children's work. The performance of these household chores is more typically the domain of female children<sup>7</sup>.

As shown in Table 3.1, on average women work much more than men. Interestingly, gender inequality in total work time observed in Ethiopia, as in many developing countries, contrasts with the iso-work phenomenon observed in developed countries<sup>8</sup>. The average total work time per week rises to 50 hours for women, while it is only about 35 hours for men. Accordingly, women spend 30 per cent of their time working, which is 9 per cent higher than men. Expressed as a percentage of the total time available for work, the share of hours spent working rises to 45 per cent for women, while for men it only represents 31 per cent. On average, women allocate 16 hours and men 30 hours to market work. With regard to housework, the average time allocated is 35 hours for women and 6 hours for men. The incidence of market work is higher among men (83 per cent) than among women (65 per cent). In contrast, almost all women do housework, while nearly two-thirds of men are not involved in any of the household activities. Moreover, the

---

<sup>7</sup> See, for instance, Guarcello, Lyon and Rosati (2006), Woldehanna, Tefera, Jones and Bayrau (2005), and Bhalotra (2003).

<sup>8</sup> Burda, Hamermesh, and Weil (2007) found, contrary to the general belief, that in rich northern countries on four continents, men and women do the same amount of total work.

average duration of housework is two times and a half higher for women (38 hr), while the average duration of market work is 12 hours longer for men (36 hr). These results show a clear gender-based division of labour which is characterized by both women (men) participating more and spending longer hours in household (market) work. However, the incidence and the average duration of market work for women are important and much higher than those of housework for men. Thus, compared to men, who generally focus only on market work, women tend to accumulate both types of work, and thereby are double-burdened.

**Table 3.1: Decomposition of the average total work hours per week by gender**

	<b>Men</b>	<b>Women</b>	<b>Total</b>
<b>Total Work Time</b>	35.3	50.5	43.4
<b>Market Work</b>			
Incidence	83%	65%	73%
Duration	35.7	24.0	30.2
Incidence * Duration	29.7	15.5	22.2
<b>Household Work</b>			
Incidence	37%	91%	66%
Duration	15.2	38.3	32.2
Incidence * Duration	5.6	35.0	21.2
<b>Fetching Water</b>			
Incidence	14%	63%	40%
Duration	5.2	7.3	6.9
Incidence * Duration	0.7	4.6	2.8
<b>Collecting Fire Wood</b>			
Incidence	16%	42%	30%
Duration	6.5	7.2	7.0
Incidence * Duration	1.0	3.0	2.1
<b>Domestic Activities</b>			
Incidence	22%	90%	58%
Duration	17.5	30.3	28.0
Incidence * Duration	3.8	27.3	16.3

*Source:* LFS 2005. *Note:* Individuals aged 15 and older. Weighted data.

The gender average total work time differential is attributable mainly to gender inequalities in housework time and participation. If the incidence or the average duration of housework for women was the same as for men, women would work on average 21 hours less, and the work burden on men would exceed that on women by almost 6 hours. Conversely, the average total work time for men would increase by more than 8 hours, and the gender differential would be reduced to less than 7 hours, if the incidence or the average duration of housework for men was as high as for women. Performing these simulations again, but now switching incidences and average durations of market work, it turns out that, in the first case, women would work on average between 5 and 8 hours more, which would increase the gender differential to 20-23 hours, and in the second case, this differential would reach 22-25 hours, as men would work on average between 7 and 10 hours less.

Finally, it is noteworthy that in Ethiopia all household activities are predominantly considered “feminine”. The proportions of women fetching water and collecting firewood (63 and 42 per cent, respectively) are very much bigger than men’s (14 and 16 per cent, respectively). In addition, the average durations of these activities are higher for women (7.3 hr and 7.2 hr, respectively) than for men (5.2 hr and 6.5 hr, respectively). Gender inequalities are more apparent in domestic activities. In fact, 90 per cent of women perform domestic activities, while only 22 per cent of men do. Furthermore, the average duration of domestic activities for women (30 hr) is almost twice as long as for men (17 hr).

Table 3.2 displays the results of the decomposition of the average total work time by place of residence and gender. On average, individuals work much more in rural (44 hr) than in urban areas (38 hr). The average time per week dedicated to work by women is 43 hours in urban areas and 52 hours in rural areas. Similarly, men spend on average more time working in rural areas (36 hr) than in urban areas (33 hr). It is noteworthy that the gender gap in average total work time, which is already impressive in urban areas (10 hr), sharply widens in rural parts (16 hr).

The other striking picture that emerges from Table 3.2 is the gender-based division of labour, which is much more acute in rural areas. Indeed, gender gaps in the incidence, as well as in the average duration, of both market and household work are greater in rural areas. In urban parts, men's incidences of market and household work are, respectively, 16 per cent higher and 49 per cent lower than women's, while in rural parts, these are, respectively, 19 per cent higher and 56 per cent lower. Moreover, in urban areas, men's average durations of market and household work are, respectively, 6 hours longer and 15 hours shorter than women's, while in rural areas, these are, respectively, 12 hours longer and 25 hours shorter. According to these results, the sexual division of labour is greater in rural areas, where housework is even more feminine, and market work even more masculine, than in urban areas.

The incidence and the intensity of market work widely vary according to the place of residence. Rural men and women are much more engaged in this type of work, but for shorter hours, as compared to their urban counterparts. As regards housework, participation and average duration are higher for men and women in rural areas, but for men the area gap in incidence and intensity of housework is relatively small while for women it is huge. The double work burden on women, then, is more pronounced in rural areas, where significantly more men focus on market work but not on housework, while more women tend to accumulate both types of work.

Unsurprisingly, incidences and average durations of fetching water and collecting firewood are higher in rural parts, where people have limited access to basic infrastructure. Furthermore, the incidence of domestic activities is slightly higher in urban areas, while the average duration is much longer in rural ones. In both areas, women are more involved than men and for longer hours in every household activity. For instance, within the household, women are responsible for water collection (Un-habitat, 2004), and have to walk long distance to perform this task (World Bank, 2006). Finally, note that in rural areas all gender gaps in household activities are larger. Overall, poor households in rural areas depend on female household members for the provision of reproductive tasks since they lack the economic means to access market substitutes.

**Table 3.2: Decomposition of the average total work hours per week by place of residence and gender**

	Men	Urban Women	Total	Men	Rural Women	Total
<b>Total Work Time</b>	32.9	42.7	38.2	35.8	52.2	44.5
<b>Market Work</b>						
Incidence	62%	46%	53%	88%	69%	78%
Duration	45.5	39.3	42.6	34.3	21.8	28.4
Incidence * Duration	28.0	18.1	22.6	30.0	15.0	22.1
<b>Household Work</b>						
Incidence	35%	84%	62%	37%	93%	67%
Duration	13.9	29.2	25.2	15.4	40.1	33.6
Incidence * Duration	4.9	24.6	15.6	5.8	37.2	22.4
<b>Fetching Water</b>						
Incidence	11%	27%	20%	15%	71%	45%
Duration	4.6	5.7	5.5	5.3	7.4	7.1
Incidence * Duration	0.5	1.5	1.1	0.8	5.3	3.2
<b>Collecting Fire Wood</b>						
Incidence	3%	8%	6%	19%	50%	35%
Duration	6.5	6.7	6.7	6.5	7.2	7.0
Incidence * Duration	0.2	0.5	0.4	1.2	3.6	2.5
<b>Domestic Activities</b>						
Incidence	31%	84%	59%	20%	92%	58%
Duration	13.7	27.0	23.9	18.6	30.9	28.9
Incidence * Duration	4.2	22.5	14.2	3.8	28.4	16.8

Source: LFS 2005. Note: Individuals aged 15 and older. Weighted data.

To summarize, there is a strong gender-based division of labour in Ethiopia, which is much more acute in rural areas. Women work more and for longer hours than men in the household, while the reverse is true in the labour market. However, despite the fact that market work is predominantly masculine, the proportion of women working in the labour market and the time they spend in it are important and much higher than the incidence and the duration of housework for men. It appears then that women experience a double work burden, as they tend to accumulate both types of work, unlike men, who generally focus only on market work. Consequently, women spend more time at work than men, this phenomenon being observed to a greater extent in rural areas.



An increasing number of studies deal with time use issues in developing countries. However, empirical evidence remains scarce, especially for Sub-Saharan Africa, due to the lack of appropriate data. Comparisons with available studies on this topic are made difficult by the fact they use different kind of data, methods and approaches. Although not strictly comparable, our findings are in line with the literature. In his review of empirical evidence on time use in Africa, Charmes (2006) find that, in the four Sub-Saharan countries where time use surveys have been conducted at the national level since 1998, women spend more time than men at work. Women total workload exceeds men's by nearly a half in Benin, 29 per cent in South Africa, 18 per cent in Madagascar and 6 per cent in Mauritius. This is because their involvement in domestic and care activities is much bigger than men's: nearly 5 times more in Madagascar, 4 times more in Mauritius, and 3 times more in South Africa and Benin. In another study, Bardasi and Wodon (2010) show that in Guinea the mean working time in urban areas rises to 36 hours per week, 39 hours for women, and 34 for men. While men spend more time in the labour market, the amount of time women spend on domestic chores is much higher than for men. In rural areas, the average working time is 49 hours, again with a higher level for women than for men.

What is surprising in our results, however, is the fact that time use estimates are remarkably low. Although this is not an isolated case, since this *prima facie* problem is also observed in other time use studies, it raises important questions. Data quality can be rightly questioned, alleging that questionnaire-based estimates like the LFS 2005 are not the more suited to conduct this kind of analysis since they contain less information and they are more affected by recall or reporting biases, and therefore are less accurate, than diary-based estimates. Departing from the hypothesis that our estimates are downward biased, it could also be presumed that, owing to cultural and other factors, individuals tend to underreport their time spent in some activities as, for example, men doing household chores. But low estimates are not necessarily (only) the result of data quality issues or underreporting. According to Wodon and Beegle (2006), evidence for Malawi and other developing countries suggests the existence of labour shortages at the peak of the

cropping season, and substantial underemployment for most of the year, especially in rural areas. As regards Ethiopia, Dercon and Krishnan (2000) use data from rural areas to show high levels of seasonal and year-to-year variability in consumption and poverty, with households also responding to changes in labour demand and prices. The LFS 2005 data collection took place in a short period of time, during the month of March of that year. Since Ethiopia is characterized by diverse climates which translate into diverse vegetation zones, it is difficult to assert at the country scale whether this specific period was particularly marked by high underemployment. However, underemployment is a well-known widespread problem that affects countries like Ethiopia, and this is most probably what our findings are in part telling us.

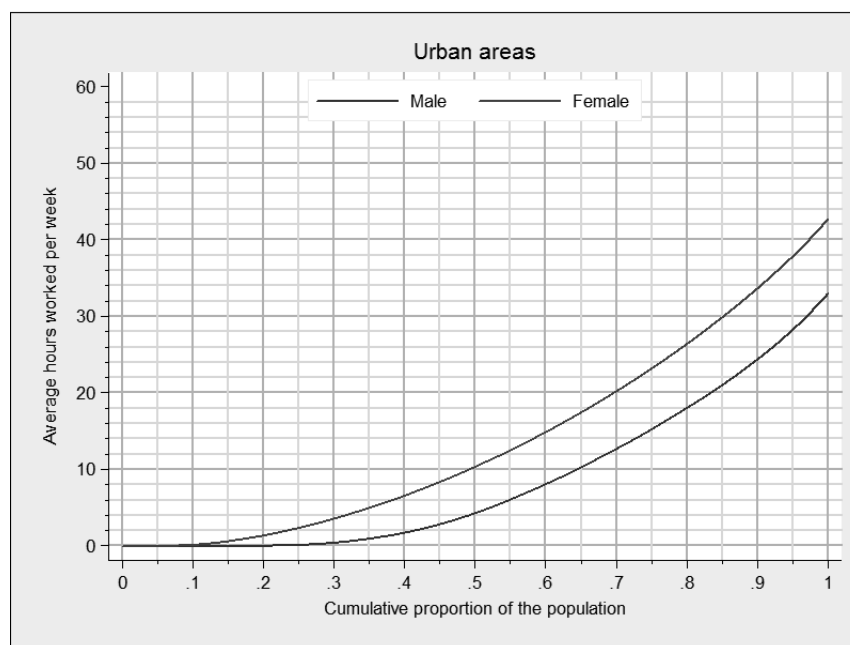
Beyond the sexual division of labour, what is indeed striking in Ethiopia is the work burden on women, which exceeds by far that on men. However, this finding is based on the averages of the total time devoted to work by men and women. The higher average for women could be the result of a small group of women who are more heavily burdened in total. To test this hypothesis, we constructed the Generalized Lorenz Curves for the total work time by gender. These are depicted in Graphs 3.1 and 3.2, respectively, for urban and rural areas. The Generalized Lorenz Curves – graphical representations of inequality that gather information about the shape and the level of men's and women's workload distributions – plot on the horizontal axis the cumulative proportion of the population, ranked according to the amount of time people work, and on the vertical axis the cumulative average work time (that is, the average work time is calculated by taking the cumulative work time of a given share of the population divided by the total population).

Graphs 3.1 and 3.2 show that, in both urban and rural areas, whatever the share of the population selected, the average cumulative work time of women is always higher than that of men. For instance, in urban areas, half the women who spend less time working devote on average 10 hours to work, while the equivalent half of the men devote on average 4 hours to work. The gender gap in the average cumulative work time increases continuously all along the distribution of the

population. In rural areas, half the women that spend less time working devote on average 16 hours to work, while the equivalent half of men devote on average 9 hours to work. This difference also increases as we move to higher shares of the population. The gender gap in the average cumulative work time is more or less equal across areas at the first deciles of the distribution of the population, but at the median and at higher deciles, this gap is increasingly higher in rural areas.

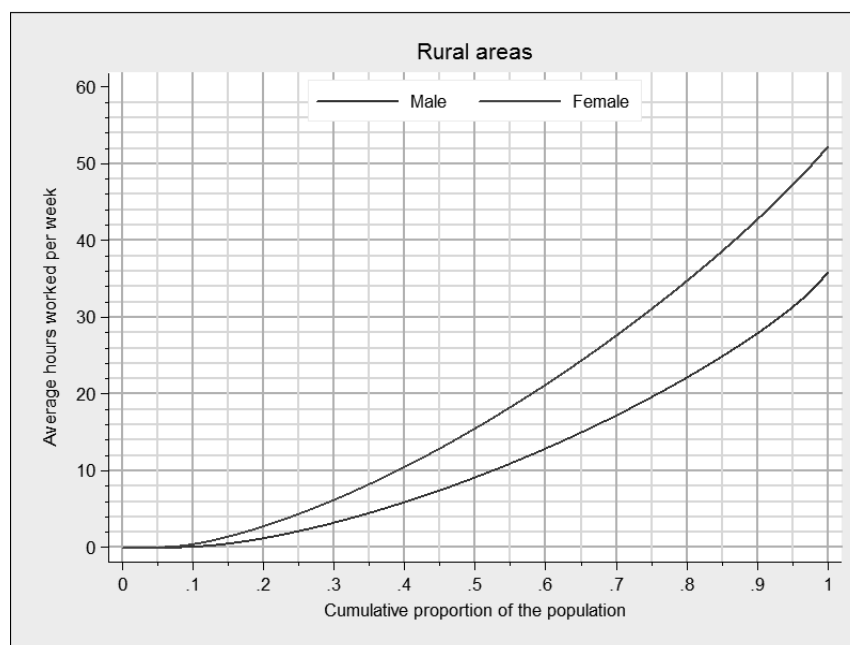
The Generalized Lorenz Curves of women always lie above those of men. The hypothesis previously formulated is rejected. Women spend more time working than men, on average and at all points of the distribution of the population. In addition, the Generalized Lorenz Curves show strong evidence of the higher gender inequality in total work time in rural areas.

Generalized Lorenz Curves by gender have also been constructed separately for market and household work time. These are depicted in Appendix A Graphs 3.3 and 3.4, respectively. They show that no matter the share of the population selected, women systematically work more than men in the household, while the reverse is observed in the labour market. Moreover, it appears that this unfavourable situation of women relative to men is more acute in rural areas, regardless of the point of the population distribution chosen.



**Figure 3.1: Generalized Lorenz curves for the total work time by gender in urban areas**

**Source:** *Ethiopia LFS, 2005. Note: Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*



**Figure 3.2: Generalized Lorenz curves for the total work time by gender in rural areas**

**Source:** *Ethiopia LFS, 2005. Note: Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*

### 3.4.2 Further disaggregations of time use estimates

It is worth going beyond the general and already well documented urban-rural divide, and try to get a better picture of the heterogeneity of the gender relation in Ethiopia. To this end, we disaggregate our results along a few dimensions believed to be important characteristics deemed to identify clusters of potentially more under- or over-loaded workers. We focus on individual and intra-household characteristics, which undoubtedly play an important and direct role in determining the way people allocate their time<sup>1</sup>.

**Age groups.** Disaggregating data by age cohort (15-24, 25-39 and 40+) shows that the group facing the highest total work burden is women aged 25-39 (57 hr). By contrast, men in the age group 15-24 are the most under-loaded, spending an average total of 32 hours a week only in all types of work. It is when they are middle-age adults, that men and women are the most involved in market work, both in terms of incidence and average duration. These findings are in line with the prediction of intertemporal time use models according to which life-cycle pattern is such that hours of market work will be greatest in an individual's middle years (Ghez and Becker 1975; Blinder and Weiss 1976; Ryder, Stafford and Stephan 1976). Women also experience the highest housework burden in middle-age adulthood, whereas men are the most involved in household chores when aged 15-24. Besides, we observe the sharpest gender gaps in total work time and in average durations of market and household work among people aged 25-39. Indeed, in this age group, women have a total workload which exceeds by almost 18 hours that on men, and they spend on average 34 more hours in housework, and 17 less hours in market work, than men. The biggest gender gaps in market and household work incidences are found among the oldest. Accordingly, while women disadvantages remain, as a matter of course, a deep concern among young people, our results could be indicative of some recent improvements towards gender

---

<sup>1</sup> Disaggregated time use estimates are not displayed in order to save space. They are available upon request.

equality or, at least, that it is at young ages that gender disparities in time allocation use to be the less widespread.

**Levels of schooling<sup>2</sup>.** The relation education-workload seems to follow a U-shaped profile: up to the general cycle, education translates into a continuous downsizing of the total amount of hours worked; then, for those who go beyond and reach higher levels of education, we observe an abrupt increase in total workload. The heaviest total workload among men is observed for those with beyond general education (39 hr), while among women those illiterates are found to be the most total work burdened (52 hr). A U-shaped profile is also observed between the level of education and the incidence of market work. Looking at the average durations of market work, it appears that, while numerous, illiterates performing market work devote on average much less time on it than people with general or beyond general education do. Data on housework show that education is negatively associated with involvement in household chores. While participation incidence of men in housework is more or less the same across levels of education, average duration continuously declines from 16 hours per week among illiterates to 13 hours per week among men with beyond general education. For women, the adverse relation between education and housework burden is much more acute, meaning that, besides contributing to increase their time spent in market work, education enables them to significantly relieve part of the household duties that usually fall on and overwhelm them. Finally, data leave no doubts on the beneficial impact of education in reducing gender inequalities in time allocation. Clearly, gender gaps in total work time as well as in market and household work time progressively vanish as people get educated and attain higher levels of education. Among the highly educated, women disadvantages

---

<sup>2</sup> The following levels of education are considered: illiterates, primary education, general education and beyond general education. Illiterates account for all working-age individuals who neither can read nor write. Primary education refers to primary education in the new system (basic education cycle, grades 1–4; general primary cycle, grades 5–8), non-formal education and literacy campaign. General education includes grades 9–12 in the new system (general secondary education, grades 9–10; preparatory secondary education, grades 11–12) and grades 9–12 in the old system. And, finally, beyond general education refers to new vocational education (grades 11–12), certificate, diploma (grades 11–13), degree completed or not, and above degree.

are still noticeable, but they are not as pervasive as they are among illiterates or the less-educated. From a policy point of view, these results stress and reinforce the importance to improve women's access to all levels of education, not only for the purpose of achieving the 2015 target of the UN Third Millennium Development Goal of gender equality and women empowerment, but also, beyond education, to improve women's relative situation in the economic sphere and in the domestic realm in order to dissipate traditional gender roles and lessen the gender division of labour.

**Male-headed vs. female-headed households.** In poor countries like Ethiopia, where the traditional and male-dominated society prevails, in which generally men endorse the role of breadwinner and women that of housekeeper or caregiver, it is not very common to see households headed by women. In the case of Ethiopia, 3 out of 4 households are male-headed. A closer look at these two types of household reveals substantial differences. Almost all male household heads are married (91 per cent) and live with a spouse (89 per cent). In turn, female household heads are only 27 per cent to be married and no more than 3 per cent live with a spouse. The fact that a significant proportion of female household heads are married but not living with their spouse is illustrative of a non-negligible phenomenon characterizing developing countries and in particular Ethiopia, which is the migration of male household heads from rural to urban areas, or from regions with low labour demand to others in which employment opportunities and earnings expectations are higher. Although they are quite rightly not considered as household heads given that they do not live permanently in the household, men – there are also many women in this situation – migrate to send money home to their families so, while being away, they continue to play a determinant role in their household. However, the striking point here is that, because they are for the vast majority not married and have no spouse, women heading households, unlike male household heads, can only rely on themselves and, to some extent, on their children to sustain their household, thus they are a vulnerable group at particular

risk of time poverty<sup>3</sup>. Households' composition is also quite distinct depending on the sex of the household head. 53 per cent of members of male-headed households are males, while men account for only 25 per cent of members of female-headed households. Time use estimates show that women work in total more than men, whether they belong to male-headed or female-headed households, and that both men and women have a lower total work burden when living in female-headed households. The sex of household head seems to affect the involvement of men and women in market and household work. Overall, our results indicate that individuals can expect to participate more (less) in market work and less (more) in housework if they have the same (opposite) sex as the household head. Besides, it is noteworthy that all gender gaps in time use estimates are significantly lower in female-headed households, suggesting that in this type of households attitudes are more gender-egalitarian.

**The presence of children.** We lastly analyse the way men and women allocate their time depending on whether they live with or without children. Children are defined as those aged 0-14. The detrimental effect of the presence of children on women's total workload and engagement in housework appears clearly. Living with children implies for women an increase of 7 hours of the total work time. This increase is fully attributable to the incidence and the average duration of

---

<sup>3</sup> Time poverty is a relatively new and interesting concept. Its measurement is under debate as there is no internationally agreed definition. The idea that poverty is a function of time as well as money was first articulated by Vickery in 1977. In her work, she defines a two-dimensional measure of household needs accounting for not just money, but also the time requirements to achieve a minimal level of consumption, corresponding to the poverty line. In the same spirit, Robert Goodin *et al.* (2008) define the "minimum necessary" time as the amount of time needed to earn an income corresponding to the poverty line and use this concept to distinguish between those who work long hours by necessity and those who work long hours by choice. More recently, Bardasi and Wodon (2006) apply the traditional concepts and techniques used for the analysis of income or consumption poverty to time poverty, and use the first three poverty measures of the so-called FGT class after Foster, Greer, and Thorbecke (1984) to define the headcount index of time poverty, the time poverty gap, and the squared time poverty gap. The time poverty line represents a certain threshold work time above which individuals do not have enough time for rest and leisure, and thereby are considered time poor. Because of the impossibility of determining the correct level of rest and leisure individuals need, the authors arbitrarily define two relative time poverty lines: a lower threshold equal to 1.5 times the median number of total individual working hours distribution, and a higher threshold equal to 2 times the median. Bardasi and Wodon (2010) propose a new approach that recognizes the multidimensional aspect of poverty. They refine the definition of time poverty by combining the concept of time poverty as in working long hours with consumption poverty. According to this new definition, an individual is time poor if he works long hours without choice because his household is poor or would be at risk of falling into poverty if the individual reduces his working hours below a certain time poverty line.



housework which rise, respectively, by 3 per cent and 6 hours with the presence of children. Babies and young children are completely dependent on adults to supply their essential needs. Rearing children means performing many daily activities to take care of them that are altogether really time-consuming. It seems that women are committed to children's care and welfare, but not men. On the contrary, men are 11 per cent less likely to do housework in households with children. Children take part in many household chores, even at young ages, some of which are primarily carried out by them. Thus, living with children is probably advantageous for men in the sense that it releases them from part of the housework they would have to do otherwise. Nonetheless, the need of additional income due to the presence of children leads men to participate more in the labour market. Indeed, having children barely lower the time devoted by men to market work, but it substantially increases their participation incidence. In total, the work burden on men is less than 2 hours greater in households with children. Finally, we observe that the presence of children deeply aggravate inequalities between how men and women make use of their time.

Summing up, the analysis of time use patterns across different subgroups reveals that illiterate individuals in middle-age adulthood living in male-headed households with children are the more heavily work burdened. Men have the higher workload when highly educated, while women exhibit the higher total work time when illiterate. It is also among illiterate middle-age adults living in male-headed households with children that gender disparities in time use are the more pervasive.

### **3.4.3 The determinants of market and household work time**

The next investigation looks at the determinants of market and household work time across gender and place of residence. Market and household work time are, respectively, estimated with the generalized Tobit and the standard Tobit methods, for the overall sample of individuals aged

15 and older, and separately for men and women in urban and in rural areas. For both dependent variables, we include, as explanatory variables, human capital characteristics (levels of education, training, potential work experience, which is defined by age minus years of schooling minus 6, and its square to take into account its possible decreasing returns), other individual characteristics (binary variables for marital and disability status), and other characteristics that reflect household composition (number of infants, children, senior people, male and female adults for various age cohorts). Because time is a limited resource and decisions on time spent in each type of work may influence each other, we include, as a regressor, housework hours in the market work time equations, and market work hours in the housework time equations. In this analysis, we are particularly interested in the impact of these variables in order to know to what extent decisions on market and household work time are interrelated. Regions dummies are also included in the regressions<sup>4</sup>.

### ***Preliminary considerations on models' specification and data quality issues***

Human capital and household characteristics such as education and potential work experience, the number of infants and children, and other household demographics, are relevant explanatory variables for market work time, since they are factors that are believed to affect individual's potential earnings or reservation wage<sup>5</sup>. Human capital characteristics are relevant determinants

---

<sup>4</sup> At first, predicted log hourly wages, either as a single regressor, in its quadratic form or by population-weighted quintiles of the wage distribution, were also added to the specification of the equations. We used the selectivity corrected (Heckman's two-step method) log hourly earnings in wage employment by gender of Chapter 2 to predict the log hourly wages for all surveyed working-age men and women. The results of these variables appeared to be not really intuitive and easy to interpret. It is worth precising here that, in the LFS 2005, earnings are only observed for waged employed workers in their main occupation. Therefore, the ability of predicted wages to explain hours worked is limited and distorted by the fact they do not account for multiple job holding and self-employment, where it is clear that earnings determination process is completely different than in wage employment. As a matter of course, any attempt to interpret these results is risky; therefore we finally decided to not include these variables in the final specification of the equations.

<sup>5</sup> It would have been interesting to account in the regression analysis also for the effect of household income (household labor income may give rise to endogeneity problems, thus it is probably better to rather consider the unearned non-labor income of the household). Indeed, household income is also expected to

for housework time as well, since they impact marginal productivity in non-market activities. Household composition is expected to determine in part individual's allocation of time, in particular the number of infants and children, whose presence may raise the demand for services in the home. In turn, the presence of other adults, especially adult women, may reduce individual's housework time through substitution or economies of scale effects.

There are concerns, however, over the inclusion of the number of infants and children as explanatory variables in the regressions, because fertility may be endogenous to choices regarding market and household work. If fertility and time allocation decisions are jointly determined, inferences of causality from the former to the latter will be incorrect (Glick, 1999). The usual way to deal with this kind of problem is to rely on exclusion restrictions to generate appropriate and suitable instruments for the endogenous variables. However, exclusion restrictions that usually attempt to control for the possible simultaneity of the fertility-work relation are often questionable. In addition, the LFS 2005 provides no variables which could reasonably be used as valid instruments for the number of infants and children. In spite of this, we decide to keep these variables in our regressions, because they are expected to be important determinants of individual's allocation of time.

The sample of individuals engaged in market work may be a non-random sample and have specific characteristics. To deal with this potential sample selection bias, we look, in the generalized Tobit regressions for market work time, for exclusion restrictions by seeking one or several variables that can generate nontrivial variation in the selection variable but do not affect the outcome variable directly. Because it is often hard to come up with a variable with these characteristics, a strong justification is necessary for imposing the exclusion restriction. In our case, we test several variables and combinations without success. Wald tests results of joint significance in the first stage and insignificance in the second stage in the regression analysis of

---

affect individual's reservation wage. Unfortunately, the LFS 2005 does not contain this kind of information.

market work do not produce convincing results. Then, identification is achieved through nonlinear functional form, and the same explanatory variables are used for modeling the participation decision and the structural equation.

The strength of the LFS 2005 is that it contains information on a large number of observations and that it is nationally representative. However, the weakness is that this is cross section data obtained from a short questionnaire. The LFS 2005 thus offers little opportunity to find valid exclusion restrictions and to comprehensively model individual's time allocation. Besides, lack of some crucial information in the LFS 2005 makes it difficult to investigate the determinants of market and household work time. Factors other than individual characteristics, household composition and location variables may also influence the number of hours spent in each type of work. For instance, information on access to basic infrastructure resources is missing, which is undoubtedly an important determinant of the time allocated by individuals to market and household work. Using a combination of nationally representative individual level time use data combined with household and community data from Lesotho, Dawson (2007) tries to figure out how infrastructure impacts on gender disaggregated time poverty. He finds that proximity to main water source, public transport and primary school all have the expected signs of influence on time poverty, with the closer the proximity to all of these being negatively associated with individuals being time poor. Females are 9.6 per cent (5.5 per cent) less likely to be time poor if their nearest water supply (public transport) is within 30 (60) minutes, relative to women with a water supply (public transport) that is more than 1 hour away. Infrastructure variables are of greater importance for women, than for men, given control over major assets – i.e. if there is a motorbike or a car in the household it is likely that men will control the usage. Women's dependence on close public transport is clearly very important, and accentuated because of their large number of domestic/home duties (Dawson 2007).

Appendix B Table 3.3 reports the total marginal effects from the standard Tobit model for housework time, separately by gender and place of residence. These are the marginal effects for

the unconditional expected value of housework time. The detailed marginal effects from the generalized Tobit model for market work time, disaggregated by gender, gender in urban areas, and gender in rural areas, are, respectively, reported in Appendix C Tables 3.4, 3.5 and 3.6. The total effect, that is the marginal effect for the unconditional expected value of market work time, is disaggregated into two parts: the conditional effect, which is the marginal effect for the expected value of market work time conditional on being observed, and the effect on participation, which is the marginal effect for the probability of market work time being observed. Before proceeding to analyse these effects, it is noteworthy that the commonly used McFadden pseudo R-squared, which theoretically measures the goodness of fit of the models, are extremely low in the Tobit regressions for housework time. This is not unusual when estimating Tobit models. On this point, it is worth recalling that Tobit regression does not have an equivalent to R-squared in OLS regression, which measures the proportion of variance of the response variable explained by the predictors. Like McFadden (1973), many people have tried to come up with one, leading to a wide variety of pseudo R-squared statistics available today that have no real meaning in Tobit models, as in many others. Nevertheless, a small pseudo R-squared should make us humble about the model's explanatory ability. Here, we use a probably better alternative which is to calculate the R-squared between the predicted and observed values<sup>6</sup>. However, in the Tobit regressions for housework time, this alternative R-squared measure also stands at very low levels, especially in equations for men. We tested several specifications accounting for individual and household characteristics in order to reach higher R-squared values, but without success. Observed characteristics failed in explaining most of the variance of housework time across individuals, men in particular, which could be explained, at least in part, by the omission of infrastructure variables. As a consequence, the following interpretation of the results should be taken with caution.

---

<sup>6</sup> For Stata users, it is also possible to make use of the Long and Freese utility command *fitstat*, which provides a number of pseudo R-squared in addition to other measures of fit.

### ***Ancillary parameters of the generalized Tobit models for market work time***

We first have a look at the results of the ancillary parameters of the generalized Tobit models for market work time, displayed in the bottom of Tables 3.4, 3.5 and 3.6. Rho is the correlation coefficient between the unobservables that determine selection into market work and the unobservables that determine hours of market work. Wald tests reject at 1 per cent the null hypothesis of  $\rho=0$  in all the generalized Tobit regressions (5 per cent in the case of urban women), which indicates that the residuals of the selection and outcome equations are well correlated, meaning that choices of participation and time devoted to market work are indeed dependent. In such cases, the use of generalized Tobit models is justified because standard regression techniques, which do not correct for potential selection bias, would have yield biased results<sup>7</sup>.

### ***Sex and area of residence***

Consistent with descriptive statistics, sex and area of residence are statistically significant and important determinants of individuals' time allocation. *Ceteris paribus*, being a woman increases by 28 hours the time spent in housework, and reduces by 12 hours the time devoted to market work. Decomposing the total marginal effect on market work time brings the following additional results: women are 21 per cent less likely than men to perform market work, and when they do, they spend on it 8 hours less than men. All else being equal, living in rural areas, where people have limited access to basic infrastructure and social services, translates into 2 and 7 additional hours of housework for men and women, respectively. As regards market work, effects on participation are positive, while effects on working hours given participation are negative, for

---

<sup>7</sup> The other ancillary parameters are sigma, which is the standard error of the residual in the market work time equations, and lambda, which is the selectivity effect calculated as the product of rho and sigma. Overall, they suggest that a slight negative selection is at work (except among rural women) since unobservable characteristics that are positively related to market work participation appear to be also slightly negatively related to market work hours.

both men and women. However, their magnitudes are not comparable across sexes, and the observed total effect of rural location appears to be positive for men and negative for women.

In the following, we only comment on the results of the male and female regressions that are performed separately for urban and rural areas. Labour market conditions are so different between cities and countryside – in rural areas almost all productive activities are in or related to agriculture, not to say traditional or subsistence agriculture –, that it is more appropriate to analyse them separately.

### ***Complementarity/substitutability of time allocation decisions***

In urban areas, an additional hour of market work translates into a decrease of 0.07 hour and 0.27 hour of housework time, respectively for men and women. In turn, an additional hour of housework time results in a reduction of 0.5 hour of market work for both men and women. While no causal inference can be established solely by estimating our econometric models, from our findings we can at least say that the impact of housework hours on market work time seems to be stronger than the reverse in urban areas. These results are suggestive that there is substitutability in time allocation decisions in urban areas but, while housework time appears to be quite insensitive and inelastic to market work hours, especially among men, the amount of hours devoted to productive activities seems to be, in turn, constrained and conditioned to some extent by the amount of hours allocated to household chores. As regards women, it is well recognized that their housework burden limits their time available to market work and allows them to engage only in productive activities compatible with their household duties.

Findings for urban men hold also for those living in rural areas, with the nuance that the adverse impact of housework hours on market work time is less strong in the countryside than in cities. As regards rural women, results are totally different. Indeed, the time they spend working at home does not affect and constrain, or it does but very marginally, their time spent working in the labour market, and vice versa. Women's time allocation decisions seem then to be disconnected

in rural areas, where gender disparities and traditional gender roles are the more pervasive and deeply rooted.

### ***Human capital and other individual characteristics***

The level of education has generally a statistically significant negative impact on men's and women's housework time. However, urban women can expect to reduce statistically significantly their housework burden only when they have a beyond general education. In turn, urban men exhibit a negative relation between schooling and housework time, which is not as important and statistically significant as for women, only when they have reached general and beyond general education levels. In rural areas, education has a greater impact on housework time. It reduces by 1, 5 and 12 hours the housework burden of rural women who have achieved, respectively, a primary education, a general education, and a beyond general education. As regards rural men, they decrease at most by 2 hours their housework time when they have acquired a general education. Primary education has no effect on men's and women's market work time in urban areas. By contrast, general education is associated with a decrease of 1 and 4 hours in, respectively, urban women's and men's market work hours. Interestingly, having achieved a beyond general education in urban areas translates into an important reduction in market work hours for men, while for women it implies higher participation incidence and duration of market work. The positive relations between education and market work time and participation incidence are very important for women in rural areas. Indeed, women are 10 and 18 per cent more likely to be engaged in productive activities, and they work 6 and 12 hours more in these activities, when they have, respectively, a general education and a beyond general education. With regard to rural men, only primary education exerts a really significant impact on market work time, which is negative.

Overall, the effect of education on labour supply appears to be strongly gendered. First, education has generally a negative influence on the time allocated to housework, especially among women.



Second, education is usually a negative predictor of the time devoted to market work by men, and a positive predictor of that devoted by women. From these general observations, what is remarkable is that being highly educated leads women to be both more engaged in productive activities and less involved in household chores, this phenomenon being observed to a much higher extent in rural areas.

The observed decrease of market work time with education level of urban men might reflect a negative income effect on labour supply. In turn, the positive relationship between education level and market work time and participation incidence of rural women might indicate that the substitution effect dominates the income effect. That is, as rural women become more educated, they substitute market work hours for leisure hours to take advantage of the higher expected labour income or, said differently, because of the higher opportunity cost of leisure.

Having benefited from training has a statistically significant impact on housework time for urban women only. These spend 1.5 hour less in household tasks than their non-trained counterparts. Training is, in addition, a strong positive determinant of market work time for both men and women. By training, we explicitly refer to having received a certificate or above in any technical or vocational education or training. Two explanations come into mind. On one hand, technical and vocational skills may be more valued in the labour market because they better fit the required qualifications for many jobs, and they can be immediately applied in the workplace. On the other hand, it may be that jobs, where such skills are needed, are more time-consuming and/or less rewarding, and thus push workers to work longer hours. Training has a greater impact on market work time in cities, suggesting that these skills are more in demand there than in the countryside. We also observe that training has a higher effect on women's market working hours than men's in rural areas, while the reverse is true in urban parts.

Interestingly, potential work experience has opposite effects on time devoted by men and women to household chores. Men's housework hours decrease linearly, and women's increase concavely, with potential work experience. It is worth mentioning that the use of potential experience, as a

proxy for workers' actual experience in the labour market, is probably a good approximation of the true experience for men, who usually have a higher labour force attachment, but may lead to overestimate the amount of experience for women, who participate discontinuously in the labour market, for instance, because of their involvement in childcare. Effects of potential work experience on men's and women's market work time have inverted U-shaped profiles, and are more pronounced in urban than in rural areas.

Marginal impacts of marital status shed light on labour division across spouses. They show that married women, especially when living in cities, tend to work fewer hours in the labour market and more in the household. As regards men, being married does not statistically significantly change their time spent in market work, but it reduces, to a higher extent in rural than in urban areas, their housework time.

Women who are stricken by disability are less involved in housework, as compared to those able, in urban, and even so more, in rural areas<sup>8</sup>. In addition, disability has a negative impact on market work participation, and working hours given participation, among rural women. Disability is a matter of greater concern in rural labour markets. Indeed, productive activities in rural zones, which are for the vast majority agricultural, generally require physical aptitudes and abilities that are out of reach of people with some disabilities.

### ***Composition of the household***

Of particular interest is the intra-household allocation of time which we analyse through time allocation effects of household composition. The impact of the number of infants on women's housework hours is significantly positive, and the effect of the number of children on men's housework time is significantly negative. Within households with infants, women are more

---

<sup>8</sup> In the LFS 2005, an individual is considered as disable if he has at least one of the following characteristics: i) seeing difficulties, ii) hearing difficulties, iii) speaking difficulties, iv) standing and moving difficulties, v) body movement difficulties, vi) mental retarded, vii) mental problem, and viii) other physical/mental-related difficulties.

heavily burdened because they have to take care of them. As infants grow up, they start taking part in many types of household activities, allowing men to be less involved in such activities. Market work effects of the number of infants and children are less statistically significant. Nevertheless, they show that, with the presence of children, men, especially those living in urban areas, raise their hours of market work, while women, but only those located in rural areas, reduce them. In addition, the presence of children has a 1 per cent statistically significant negative impact on market working hours of urban women. In Ethiopia, where the prevailing traditional society remains male-dominated, generally men endorse the role of breadwinner and women those of housekeeper and caregiver. Women are much more committed than men to children's care and welfare. Moreover, men benefit from the presence of children in the household in the sense that it releases them from part of the housework they would have to do otherwise. Nonetheless, the need of additional income due to the presence of children leads men to be more involved in market work. Many research works demonstrate that when women have better command over income resources, decisions on how these resources are spent tend to favour children more in terms of human capital investment (Hoddinott and Haddad 1995; Bourguignon and Chiappori 1992; Browning and Chiappori 1998; Bussolo, De Hoyos and Wodon 2009). In addition, higher involvement of women in decision making within the household has also been shown to lead to better outcomes in terms of nutrition (Piesse and Simister 2002).

An interesting picture emerges from time allocation effects of the number of male and female adults in various age cohorts. It first appears that the impact of the number of adult women on housework hours is strongly significant and negative, especially in rural areas and among women. The presence of middle-age adult women in the household reduces the most men's housework time, while it is the presence of young (prime-age adult) women in the household that reduces the most urban (rural) women's housework time. Because household activities are primarily carried out by women, it is not surprising then that the time spent on housework by men significantly decreases as the number of adult women in the household increases. In turn, the presence of other

adult women in the household relieves individual women of part of their housework burden. The presence of adult men increases significantly housework burden on women, especially in rural areas and when adult male household members are young or in prime-age adulthood. As regards market work, we observe that the presence of young women positively affects hours worked by men in urban areas. In addition, urban men are more likely to do longer hours of market work when living with prime-age adult men. Finally, women decrease their market working hours with the presence of adult men, in particular when living in rural areas with middle-age adult men. Overall, age and gender composition of the household are significant predictors of time allocation, especially in rural areas and among women. Besides, results are interesting in that they could be indicative of some form of hierarchy within households.

Finally, regression results show that time allocation effects of regional location are important. This is not surprising since Ethiopia is characterized by a diversity of climates and geographies, cultures, languages and ethnic groups that influence job opportunities and choice. Moreover, regions in Ethiopia are increasingly playing a substantial role in determining their own economic policies, and this is expected to increase the disparities across local labour markets.

### **3.5 Conclusion**

The main aim and the novelty of this study is to help fill the gap in current knowledge about the nature, the extent, and the determinants of a key dimension of work-related gender disparities in the Ethiopian context: time allocation between market and household work. In particular, the research question this Chapter attempted to address is whether and to what degree there is complementarity or substitutability between the two types of work. In other words, we were interested in understanding the connection between market and household work time allocation decisions.

Nationally representative time-diary and time use surveys, as well as household surveys including a time use module, have not been carried out in Ethiopia. We then drew upon the LFS 2005 which, despite its shortcomings, offer a unique opportunity to analyse gender-differentiated time-use patterns and to address the issue of substitutability or complementarity between market and household work time allocation decisions.

However, it leaves no doubts that the time-diary approach is of crucial importance for at least two reasons. First, it is essential to take into account the domestic realm in order to have a better understanding of the labour supply, and this can be achieved through the analysis, for instance, of time-diary estimates. Second, time-diary estimates are more accurate and detailed than stylized (questionnaire-based) estimates. It is thus worth mentioning that time allocation analysis would greatly benefit from using survey data collected using the time diary approach.

How do men and women allocate their time between market and household work? Do they perform the same amount of total work? What are the determinants of market and household work time across gender? Is there substitutability or complementarity between the two types of work? To what extent? These are the questions we attempted to answer throughout the Chapter. The main findings can be summarized as follows.

There is a strong gender-based division of labour in Ethiopia, which is much more acute in rural areas. Women work more and for longer hours than men in the household, while the reverse is true in the labour market. However, despite the fact that market work is predominantly masculine, the proportion of women working in the labour market and the time they spend in it are important and much higher than the incidence and the duration of housework for men. It appears, then, that women experience a double work burden, as they tend to accumulate both types of work, unlike men, who generally focus only on market work. Consequently, women spend more time at work than men, this phenomenon being observed to a greater extent in rural areas.

However, these findings are based on the averages of the total time devoted to work by men and women. The higher average for women could be the result of a small group of women who are

more heavily burdened in total. To test this hypothesis, we constructed the Generalized Lorenz Curves for the total work time by gender. It appears that women's curves always lie above those of men. This clearly indicates that women spend more time working than men, on average and at all points of the distribution. In addition, the Generalized Lorenz Curves show strong evidence of the higher gender inequality in total work time in rural areas.

The study estimated the determinants of market and household work time across gender and place of residence using, respectively, the generalized Tobit and the standard Tobit methods, accounting for individuals and household characteristics, and location variables. We also included, as a regressor, housework hours in the market work time equations, and market work hours in the housework time equations, in order to know to what extent time allocation decisions are interrelated. Measures of goodness of fit of the Tobit models were very low, especially in equations for men, meaning that these characteristics failed in explaining most of the variance of housework time across individuals, bringing to light the fact that some crucial information was omitted. For instance, access to basic infrastructure resources is unfortunately missing in the LFS 2005, while it is undoubtedly an important determinant of the time allocated by individuals to market and household work.

Commenting on the most robust results, it appears that, overall, the effect of education on labour supply is strongly gendered. First, education has generally a negative influence on the time allocated to housework, especially among women. Second, education is usually a negative predictor of the time devoted to market work by men, and a positive predictor of that devoted by women.

Our results further indicate that there is substitutability in time allocation decisions in urban areas. While housework time appears to be quite insensitive and inelastic to market work hours, especially among men, the amount of hours devoted to productive activities seems to be, in turn, constrained and conditioned to some extent by the amount of hours allocated to household chores. The same holds for rural men, with the nuance that the adverse impact of housework hours on

market work time is less strong in the countryside than in cities. By contrast, women's time allocation decisions seem to be disconnected in rural areas. Indeed, the time rural women spend working at home does not affect and constrain, or it does but very marginally, their time spent working in the labour market, and vice versa.

Age and gender composition of the household are significant predictors of time allocation. The presence of adult women in the household negatively affects men's housework time, while the presence of adult men is associated with lower hours of market work and longer hours of housework performed by women. In addition, the presence of other adult women in the household relieves women of part of their housework burden.

Within households with infants, women are more heavily burdened because they have to take care of them. As infants grow up, they start taking part in many types of household activities, allowing men to be less involved in such activities.

There is still a long and difficult way to go towards achieving the ultimate goal of gender equality in Ethiopia. Despite the positive legal framework and political context for gender equality, many barriers still hinder women's empowerment. Better information and enforcement of the law in support of women's economic and social well-being, especially in remote areas where usually tribal law prevails, is, of course, essential.

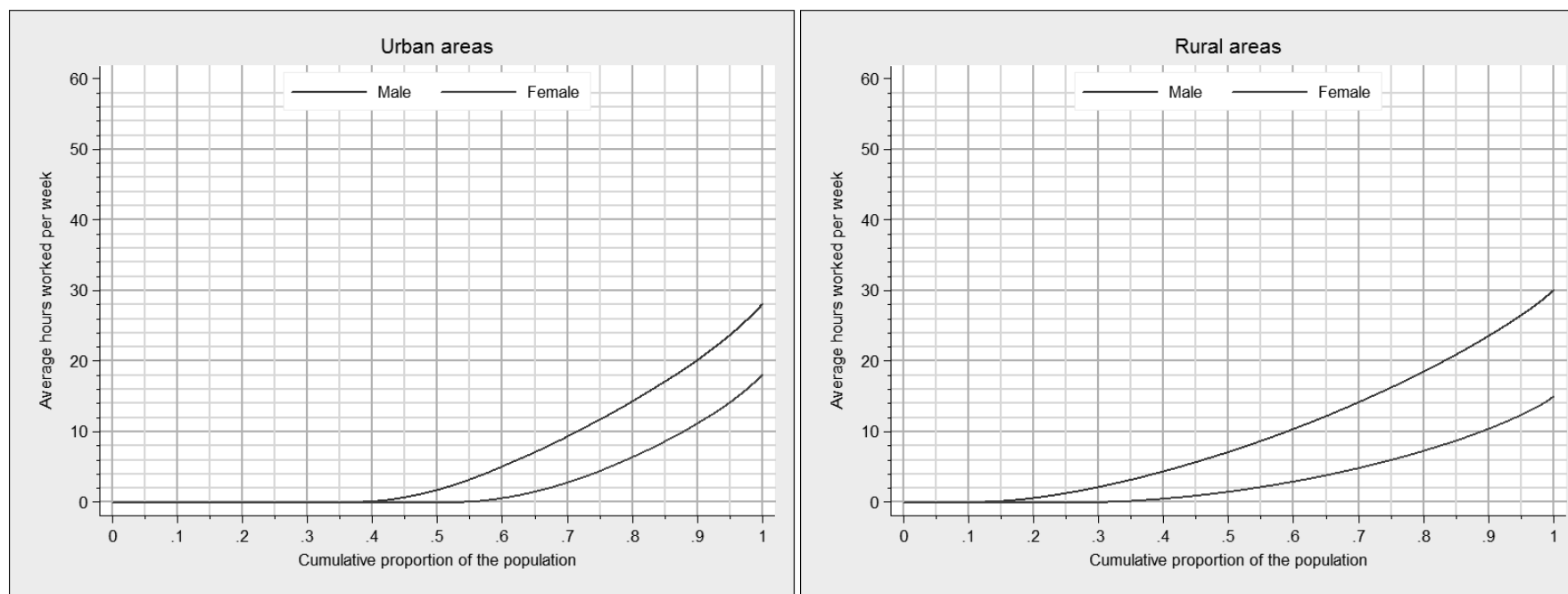
Social and cultural norms strongly influence the ability of Ethiopian women to realize their potential, especially in the social and economic spheres. Mentalities and attitudes in the Ethiopian population, who lives predominantly in rural and traditional communities, impede women to take full advantage of their abilities and keep women subordinated to men. Information campaigns and other awareness-raising efforts would be helpful in changing mentalities and attitudes and in lessening the perpetuation of traditional gender roles.

Ethiopia has seen an enormous and rapid increase in enrolment in primary education that has contributed to reduce the gender imbalance. However, gender disparities in education are still large, and little improvement has been achieved in gender equality in secondary or higher

education. Ethiopia would benefit from pursuing and intensifying its efforts to ensure better access to education at all levels for women because, as this study and others show, being highly educated leads women to participate more and for longer hours in productive and income-generating activities, while, at the same time, it relieves them from part of their housework burden.

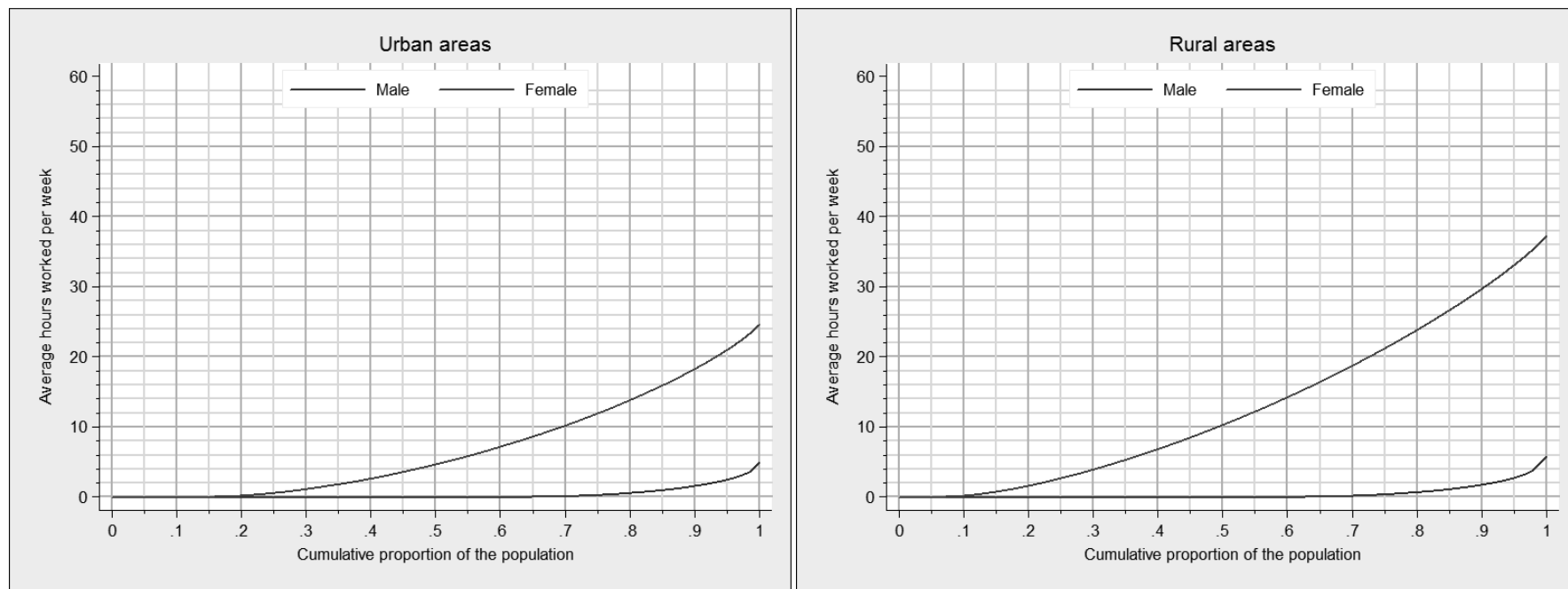


## Appendix A. Generalized Lorenz curves for the market and household work time



**Figure 3.3: Generalized Lorenz curves for the market work time by gender and place of residence**

**Source:** *Ethiopia LFS, 2005*. Note: *Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*



**Figure 3.4: Generalized Lorenz curves for the housework time by gender and place of residence**

*Source: Ethiopia LFS, 2005. Note: Individuals aged 15 and above. Generalized Lorenz curves account for sampling weights.*

## Appendix B. Housework time equations

**Table 3.3: Marginal effects from Tobit model for housework time (evaluated at sample mean) by gender and place of residence**

	All		Men		Women		Urban areas				Rural areas			
	Effect	Mean	Effect	Mean	Effect	Mean	Men	Mean	Women	Mean	Men	Mean	Women	Mean
Female	28.226*** (0.188)	0.532	—	—	—	—	—	—	—	—	—	—	—	—
Rural	3.963*** (0.182)	0.825	1.564*** (0.174)	0.830	6.748*** (0.267)	0.821	—	—	—	—	—	—	—	—
Market work hours	-0.134*** (0.005)	22.176	-0.086*** (0.005)	29.708	-0.088*** (0.007)	15.560	-0.072*** (0.004)	28.058	-0.266*** (0.005)	18.076	-0.090*** (0.007)	30.046	0.019* (0.011)	15.013
Primary education	-2.161*** (0.233)	0.277	-0.729*** (0.181)	0.397	-1.341*** (0.388)	0.172	-0.422 (0.310)	0.457	-0.432 (0.359)	0.362	-0.764*** (0.203)	0.384	-1.339*** (0.516)	0.131
General education	-2.817*** (0.342)	0.052	-1.273*** (0.267)	0.066	-2.132*** (0.520)	0.039	-0.724** (0.321)	0.287	-0.390 (0.437)	0.195	-1.873*** (0.531)	0.020	-5.130*** (1.724)	0.006
Beyond general education	-3.266*** (0.507)	0.019	-0.977** (0.360)	0.026	-6.449*** (0.790)	0.013	-0.733* (0.368)	0.122	-3.301*** (0.645)	0.062	-0.422 (0.831)	0.007	-12.069*** (2.432)	0.003
Potential experience	0.092*** (0.032)	25.985	-0.097*** (0.027)	25.902	0.569*** (0.043)	26.058	-0.079*** (0.021)	18.400	0.600*** (0.033)	20.928	-0.101*** (0.035)	27.436	0.571*** (0.059)	27.173
Potential experience squared	-0.004*** (0.000)	976.460	0.001 (0.000)	999.562	-0.011*** (0.001)	956.170	0.000 (0.000)	627.763	-0.011*** (0.001)	754.284	0.001 (0.000)	1075.609	-0.012*** (0.001)	1000.078
Training	-0.546 (0.410)	0.038	-0.042 (0.300)	0.056	-0.412 (0.609)	0.022	-0.204 (0.220)	0.209	1.503*** (0.471)	0.099	0.131 (0.571)	0.025	-1.584 (2.269)	0.005
Married	2.378*** (0.248)	0.616	-2.069*** (0.305)	0.616	5.694*** (0.322)	0.616	-1.166*** (0.247)	0.428	5.550*** (0.327)	0.402	-2.261*** (0.384)	0.654	5.277*** (0.407)	0.662
Disable	-5.278*** (0.622)	0.031	-1.083** (0.447)	0.034	-9.831*** (1.148)	0.028	-0.908 (0.540)	0.030	-8.413*** (0.980)	0.024	-1.112** (0.524)	0.035	-9.787*** (1.376)	0.029
Nb infants under 6	0.229** (0.114)	0.960	0.076 (0.103)	0.953	0.664*** (0.170)	0.966	0.048 (0.129)	0.482	1.353*** (0.181)	0.560	0.098 (0.117)	1.050	0.535*** (0.198)	1.054
Nb children between 6 and 14	-0.413*** (0.081)	1.277	-0.647*** (0.076)	1.298	-0.136 (0.117)	1.259	-0.579*** (0.083)	0.973	-0.011 (0.113)	1.045	-0.659*** (0.091)	1.365	-0.190 (0.144)	1.305
Nb men between 15 and 24	1.206*** (0.114)	0.581	-0.280*** (0.105)	0.757	1.259*** (0.183)	0.426	-0.264*** (0.081)	1.034	0.414*** (0.151)	0.504	-0.280* (0.146)	0.701	1.626*** (0.243)	0.409
Nb women between 15 and 24	-4.194*** (0.137)	0.577	-1.372*** (0.135)	0.472	-4.089*** (0.208)	0.668	-1.041*** (0.105)	0.630	-3.747*** (0.144)	0.965	-1.451*** (0.190)	0.440	-4.194*** (0.293)	0.604
Nb men between 25 and 44	0.403** (0.173)	0.486	-0.037 (0.155)	0.540	0.789*** (0.252)	0.439	-0.410*** (0.121)	0.649	0.447** (0.213)	0.464	0.173 (0.218)	0.517	1.145*** (0.352)	0.433
Nb women between 25 and 44	-3.490*** (0.201)	0.560	-1.479*** (0.184)	0.535	-4.531*** (0.296)	0.582	-0.883*** (0.144)	0.546	-3.659*** (0.209)	0.703	-1.675*** (0.259)	0.533	-4.734*** (0.427)	0.556
Nb men between 45 and 64	1.100*** (0.232)	0.280	0.222 (0.204)	0.319	0.635* (0.341)	0.245	-0.227 (0.209)	0.273	-0.148 (0.313)	0.219	0.382 (0.254)	0.328	1.072** (0.428)	0.251
Nb women between 45 and 64	-2.320*** (0.234)	0.275	-1.870*** (0.213)	0.257	-2.804*** (0.332)	0.292	-1.585*** (0.194)	0.240	-2.072*** (0.309)	0.300	-1.944*** (0.273)	0.260	-3.230*** (0.430)	0.290

Table 3.3: *continued*

	All		Men		Women		Urban areas				Rural areas			
	Effect	Mean	Effect	Mean	Effect	Mean	Men Effect	Men Mean	Women Effect	Women Mean	Men Effect	Men Mean	Women Effect	Women Mean
Nb seniors 65 and older	0.281 (0.243)	0.178	-0.705*** (0.233)	0.174	0.478 (0.347)	0.181	-0.565*** (0.217)	0.155	-1.343*** (0.342)	0.177	-0.694** (0.284)	0.178	1.020** (0.427)	0.182
Tigray	-1.908*** (0.285)	0.066	-2.120*** (0.224)	0.065	-0.607 (0.442)	0.066	-2.579*** (0.222)	0.067	3.291*** (0.498)	0.082	-2.070*** (0.333)	0.064	-3.658*** (0.883)	0.063
Affar	0.061 (0.366)	0.004	0.250 (0.323)	0.004	-0.790 (0.542)	0.003	0.601 (0.448)	0.010	-0.473 (0.688)	0.009	-0.213 (0.447)	0.003	-2.984*** (0.974)	0.002
Amhara	-1.317*** (0.251)	0.269	-0.887*** (0.224)	0.278	-1.238*** (0.367)	0.262	-1.531*** (0.173)	0.171	-1.060*** (0.344)	0.184	-0.915** (0.367)	0.300	-2.948*** (0.837)	0.279
Oromiya	0.864*** (0.246)	0.379	-0.490** (0.215)	0.376	1.822*** (0.355)	0.383	0.228 (0.228)	0.286	0.512 (0.355)	0.282	-0.750** (0.376)	0.394	0.394 (0.848)	0.405
Somalie	-3.564*** (0.307)	0.010	-3.428*** (0.199)	0.010	-1.399*** (0.508)	0.010	-3.222*** (0.182)	0.019	-4.887*** (0.503)	0.018	-3.414*** (0.295)	0.008	-1.483 (0.980)	0.008
Benshangul Gumuz	-0.769** (0.356)	0.012	-0.732** (0.302)	0.012	-0.950* (0.536)	0.011	-1.306*** (0.275)	0.009	-0.265 (0.621)	0.008	-0.734* (0.424)	0.012	-2.600*** (0.927)	0.012
SNNP	1.158*** (0.261)	0.203	0.816*** (0.242)	0.199	0.493 (0.374)	0.207	0.266 (0.213)	0.124	0.384 (0.389)	0.108	0.787** (0.401)	0.215	-1.380* (0.833)	0.228
Harari	-2.691*** (0.369)	0.003	-0.863** (0.350)	0.003	-4.719*** (0.495)	0.003	0.067 (0.455)	0.009	-3.178*** (0.532)	0.009	-2.096*** (0.445)	0.001	-8.504*** (0.986)	0.001
Dire Dawa	-3.903*** (0.373)	0.006	-2.786*** (0.267)	0.006	-4.114*** (0.602)	0.006	-1.816*** (0.285)	0.025	-3.260*** (0.624)	0.025	-4.734*** (0.242)	0.002	-6.939*** (1.048)	0.002
Expected value	17.81		6.010		34.50		4.767		23.87		6.267		36.86	
Number of observations	131112		60763		70349		33790		40685		26973		29664	
Censored observations	49182		40139		9043		22976		6797		17163		2246	
Log pseudolikelihood	-103455576.0		-31199460.3		-71345381.5		-4917904.8		-11760188.1		-26259432.5		-59424068.2	
Prob. > F	0.000		0.000		0.000		0.000		0.000		0.000		0.000	
R2	0.4068		0.0380		0.2260		0.0425		0.2290		0.0362		0.1282	

Source: LFS 2005. Note: Individuals aged 15 and older. Standard errors are in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. Marginal effects and sample means are computed after estimating a standard Tobit model for housework time accounting for sampling weights. The marginal effect is precisely defined as the marginal effect for the unconditional expected value of the dependent variable (housework time). In Tobit regressions, the McFadden pseudo R-squared is commonly used to measure the goodness of fit of the models. Here, we use a probably better alternative which is to calculate the R-squared between the predicted and observed values.

## Appendix C. Market work time equations

**Table 3.4: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender**

	All				Men				Women			
	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)
Female	0.532	-11.857*** (0.244)	-7.611*** (0.190)	-0.208*** (0.005)	—	—	—	—	—	—	—	—
Rural	0.825	-0.353*** (0.232)	-3.595*** (0.273)	0.076*** (0.005)	0.830	0.245*** (0.378)	-8.172*** (0.355)	0.178*** (0.007)	0.821	-1.161*** (0.275)	-3.508*** (0.283)	0.041*** (0.007)
Housework hours	21.250	-0.130*** (0.005)	-0.100*** (0.004)	-0.002*** (0.000)	5.621	-0.272*** (0.009)	-0.209*** (0.007)	-0.003*** (0.000)	34.978	-0.046*** (0.006)	-0.037*** (0.004)	-0.001*** (0.000)
Primary education	0.277	-1.448*** (0.208)	-1.016*** (0.169)	-0.024*** (0.005)	0.397	-2.669*** (0.316)	-1.896*** (0.274)	-0.029*** (0.006)	0.172	-0.368 (0.298)	-0.283 (0.230)	-0.008 (0.008)
General education	0.052	-0.447 (0.407)	1.790 (0.393)	-0.059 (0.009)	0.066	-1.058** (0.594)	1.131** (0.575)	-0.056** (0.010)	0.039	1.354*** (0.589)	3.317*** (0.578)	-0.028*** (0.014)
Beyond general education	0.019	-0.083 (0.559)	2.315 (0.588)	-0.059 (0.014)	0.026	-3.569*** (0.749)	-3.810*** (0.736)	-0.007*** (0.012)	0.013	4.030*** (0.910)	5.425*** (1.035)	0.019*** (0.023)
Potential experience	25.985	0.899*** (0.023)	0.519*** (0.020)	0.018*** (0.000)	25.902	0.762*** (0.039)	0.380*** (0.037)	0.012*** (0.001)	26.058	0.868*** (0.032)	0.493*** (0.026)	0.023*** (0.001)
Potential experience squared	976.460	-0.014*** (0.000)	-0.008*** (0.000)	-0.000*** (0.000)	999.562	-0.012*** (0.001)	-0.006*** (0.000)	-0.000*** (0.000)	956.170	-0.014*** (0.000)	-0.008*** (0.000)	-0.000*** (0.000)
Training	0.038	8.356*** (0.496)	5.643*** (0.453)	0.120*** (0.007)	0.056	6.704*** (0.594)	4.245*** (0.590)	0.076*** (0.005)	0.022	9.818*** (0.823)	6.798*** (0.787)	0.178*** (0.013)
Married	0.616	-0.107 (0.213)	-0.269 (0.170)	0.003 (0.004)	0.616	1.653 (0.428)	-0.478 (0.390)	0.058 (0.007)	0.616	-2.333*** (0.257)	-1.427*** (0.200)	-0.058*** (0.007)
Disable	0.031	-8.593*** (0.466)	-2.828*** (0.477)	-0.246*** (0.014)	0.034	-10.213 (0.827)	-0.539 (0.809)	-0.277 (0.019)	0.028	-6.798*** (0.522)	-2.838*** (0.515)	-0.231*** (0.019)
Nb infants under 6	0.960	0.007 (0.092)	-0.012 (0.072)	0.001 (0.002)	0.953	0.563 (0.155)	0.191 (0.134)	0.011 (0.003)	0.966	-0.394*** (0.110)	-0.266*** (0.082)	-0.009*** (0.003)
Nb children between 6 and 14	1.277	-0.185*** (0.070)	-0.187*** (0.055)	-0.001*** (0.001)	1.298	-0.289*** (0.118)	-0.320*** (0.104)	-0.000*** (0.002)	1.259	-0.148* (0.083)	-0.125* (0.063)	-0.003* (0.002)
Nb men between 15 and 24	0.581	-1.067*** (0.108)	-0.670*** (0.087)	-0.019*** (0.002)	0.757	-0.565 (0.169)	-0.226 (0.167)	-0.010 (0.003)	0.426	-0.622*** (0.137)	-0.403*** (0.105)	-0.015*** (0.004)
Nb women between 15 and 24	0.577	0.686*** (0.114)	0.509*** (0.094)	0.010*** (0.002)	0.472	0.267** (0.191)	0.449** (0.186)	-0.004** (0.003)	0.668	0.184 (0.142)	0.235 (0.113)	0.001 (0.004)
Nb men between 25 and 44	0.486	-1.618*** (0.161)	-0.619*** (0.138)	-0.040*** (0.003)	0.540	-0.200 (0.262)	0.307 (0.256)	-0.013 (0.004)	0.439	-1.588*** (0.196)	-0.584*** (0.159)	-0.050*** (0.005)
Nb women between 25 and 44	0.560	0.049 (0.173)	0.410 (0.144)	-0.009 (0.003)	0.535	-0.863 (0.277)	-0.165 (0.270)	-0.020 (0.004)	0.582	-0.320 (0.215)	0.193 (0.173)	-0.018 (0.006)
Nb men between 45 and 64	0.280	-1.946*** (0.204)	-0.959*** (0.163)	-0.042*** (0.004)	0.319	-1.117 (0.350)	-0.491 (0.336)	-0.019 (0.005)	0.245	-1.961*** (0.246)	-1.023*** (0.188)	-0.054*** (0.007)
Nb women between 45 and 64	0.275	-0.252 (0.212)	0.211 (0.173)	-0.014 (0.004)	0.257	-0.722* (0.334)	-0.572* (0.324)	-0.006* (0.005)	0.292	-0.803*** (0.256)	-0.134*** (0.203)	-0.030*** (0.007)

Table 3.4: *continued*

	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)
Nb seniors 65 and older	0.178	0.132 (0.225)	0.312 (0.181)	-0.004 (0.004)	0.174	0.725** (0.375)	0.907** (0.361)	-0.002** (0.006)	0.181	0.084 (0.259)	0.202 (0.204)	-0.002 (0.007)
Tigray	0.066	-4.170*** (0.364)	-7.294*** (0.410)	0.061*** (0.006)	0.065	-6.753*** (0.501)	-7.749*** (0.520)	0.001*** (0.007)	0.066	-1.010*** (0.472)	-4.263*** (0.568)	0.084*** (0.010)
Affar	0.004	0.675 (0.475)	-0.692 (0.485)	0.043 (0.008)	0.004	1.676 (0.614)	1.033 (0.605)	0.021 (0.008)	0.003	-0.010 (0.646)	-0.387 (0.708)	0.010 (0.016)
Amhara	0.269	-3.918*** (0.324)	-8.066*** (0.387)	0.087*** (0.006)	0.278	-4.421*** (0.431)	-6.367*** (0.440)	0.035*** (0.006)	0.262	-2.335*** (0.419)	-6.204*** (0.555)	0.077*** (0.009)
Oromiya	0.379	-5.027*** (0.312)	-8.536*** (0.375)	0.054*** (0.006)	0.376	-7.138*** (0.413)	-8.942*** (0.425)	0.020*** (0.005)	0.383	-2.293*** (0.407)	-5.660*** (0.538)	0.058*** (0.009)
Somalie	0.010	-4.316*** (0.392)	-6.019*** (0.434)	0.011*** (0.008)	0.010	-5.616*** (0.573)	-4.762*** (0.594)	-0.047*** (0.009)	0.010	-2.273*** (0.506)	-4.031*** (0.595)	0.015*** (0.013)
Benshangul Gumuz	0.012	-8.457*** (0.382)	-11.470*** (0.451)	0.014*** (0.008)	0.012	-12.254*** (0.542)	-13.565*** (0.567)	-0.018*** (0.010)	0.011	-4.119*** (0.487)	-7.201*** (0.606)	0.029*** (0.013)
SNNP	0.203	-3.490*** (0.330)	-7.458*** (0.395)	0.088*** (0.006)	0.199	-5.585*** (0.425)	-8.428*** (0.434)	0.058*** (0.005)	0.207	-0.554** (0.432)	-4.397** (0.552)	0.104** (0.009)
Harari	0.003	-5.504*** (0.453)	-7.166*** (0.513)	-0.003*** (0.009)	0.003	-8.228*** (0.643)	-7.936*** (0.660)	-0.047*** (0.010)	0.003	-1.879*** (0.601)	-4.246*** (0.712)	0.042*** (0.015)
Dire Dawa	0.006	-0.167 (0.667)	-1.848 (0.710)	0.046 (0.011)	0.006	-1.082 (0.795)	-0.221 (0.801)	-0.025 (0.012)	0.006	1.947** (0.908)	-0.065** (0.862)	0.082** (0.019)
Expected value		21.87	28.82	0.759		31.18	35.68	0.874		15.53	24.25	0.641
Number of observations			131112				60763				70349	
Censored observations			49739				16822				32917	
Wald chi2(29) & Wald chi2(28)			16783.4				7135.4				2911.4	
Log pseudolikelihood			-113626233.1				-58578734.1				-54632745.7	
Prob. > chi2			0.0000				0.0000				0.0000	
Rho			0.975 (0.003)				-0.0709 (0.014)				0.990 (0.001)	
Sigma			20.43 (0.120)				17.00 (0.079)				20.44 (0.148)	
Lambda			19.92 (0.163)				-1.206 (0.231)				20.24 (0.162)	
Wald test of indep. eqns. (rho=0):												
chi2(1)			1732.97				27.18				2390.72	
Prob. > chi2			0.0000				0.0000				0.0000	

Source: LFS 2005. Note: Individuals aged 15 and older. Standard errors are in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. Marginal effects and sample means are computed after estimating a generalized Tobit model for market work time accounting for sampling weights. Effect (1) is the total effect, i.e. the marginal effect for the unconditional expected value of market work time. Effect (2) is the conditional effect, i.e. the marginal effect for the expected value of market work time conditional on being observed. Effect (3) is the effect on participation, i.e. the marginal effect for the probability of market work time being observed.

**Table 3.5: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender in urban areas**

	Men				Women			
	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)
Housework hours	4.893	-0.454*** (0.023)	-0.346*** (0.018)	-0.005*** (0.000)	24.621	-0.461*** (0.011)	-0.443*** (0.013)	-0.007*** (0.000)
Primary education	0.457	-2.969 (0.840)	0.618 (0.688)	-0.077 (0.016)	0.362	-1.927 (0.470)	-0.646 (0.551)	-0.044 (0.011)
General education	0.287	-3.694*** (0.949)	2.079*** (0.794)	-0.113*** (0.018)	0.195	-0.932*** (0.608)	2.030*** (0.690)	-0.047*** (0.014)
Beyond general education	0.122	-5.692*** (1.105)	-3.987*** (0.913)	-0.076*** (0.023)	0.062	1.057*** (0.933)	-3.114*** (1.028)	0.070*** (0.021)
Potential experience	18.400	1.586*** (0.067)	0.520*** (0.060)	0.028*** (0.001)	20.928	1.496*** (0.061)	0.418*** (0.055)	0.035*** (0.001)
Potential experience squared	627.763	-0.025*** (0.001)	-0.008*** (0.001)	-0.000*** (0.000)	754.284	-0.025*** (0.001)	-0.008*** (0.001)	-0.001*** (0.000)
Training	0.209	10.257*** (0.657)	4.150*** (0.567)	0.164*** (0.011)	0.099	9.812** (0.796)	3.087** (0.837)	0.210** (0.015)
Married	0.428	7.132 (0.648)	0.641 (0.585)	0.153 (0.012)	0.402	-4.961*** (0.434)	-3.024*** (0.508)	-0.097*** (0.010)
Disable	0.030	-13.179 (1.413)	-0.900 (1.860)	-0.293 (0.027)	0.024	-9.170 (0.855)	-1.916 (1.697)	-0.232 (0.022)
Nb infants under 6	0.482	1.013* (0.334)	-0.434* (0.298)	0.029* (0.006)	0.560	-0.106 (0.249)	-0.084 (0.301)	-0.002 (0.005)
Nb children between 6 and 14	0.973	-0.312** (0.210)	-0.423** (0.205)	-0.001** (0.004)	1.045	-0.218*** (0.151)	-0.759*** (0.181)	0.003*** (0.003)
Nb men between 15 and 24	1.034	-1.359 (0.236)	-0.479 (0.254)	-0.024 (0.004)	0.504	-0.254 (0.196)	-0.100 (0.254)	-0.006 (0.005)
Nb women between 15 and 24	0.630	0.291*** (0.262)	0.823*** (0.269)	-0.005*** (0.005)	0.965	-1.622 (0.196)	0.065 (0.249)	-0.044 (0.004)
Nb men between 25 and 44	0.649	1.586*** (0.367)	1.658*** (0.373)	0.012*** (0.006)	0.464	-0.682*** (0.270)	1.777*** (0.349)	-0.039*** (0.006)
Nb women between 25 and 44	0.546	-1.705 (0.383)	-0.108 (0.376)	-0.037 (0.007)	0.703	-0.976 (0.275)	-0.079 (0.335)	-0.025 (0.006)
Nb men between 45 and 64	0.273	-3.181 (0.578)	0.145 (0.586)	-0.075 (0.010)	0.219	-1.228*** (0.427)	1.822*** (0.515)	-0.054*** (0.010)
Nb women between 45 and 64	0.240	-2.242 (0.547)	-0.350 (0.565)	-0.046 (0.010)	0.300	-2.957*** (0.416)	-1.879*** (0.511)	-0.056*** (0.010)

Table 3.5: *continued*

	Mean	Effect (1)	Effect (2)	Effect (3)		Mean	Effect (1)	Effect (2)	Effect (3)
Nb seniors 65 and older	0.155	-0.358 (0.633)	0.173 (0.639)	-0.011 (0.011)		0.177	-0.218 (0.453)	-0.200 (0.556)	-0.003 (0.010)
Tigray	0.067	-3.171*** (0.859)	-5.437*** (0.858)	0.009*** (0.017)		0.082	2.251*** (0.671)	-4.159*** (0.797)	0.120*** (0.015)
Affar	0.010	-0.651*** (1.141)	-3.772*** (0.918)	0.045*** (0.022)		0.009	1.210 (1.018)	1.305 (1.210)	0.016 (0.021)
Amhara	0.171	-3.696*** (0.588)	-7.564*** (0.586)	0.031*** (0.012)		0.184	0.960*** (0.469)	-7.234*** (0.576)	0.127*** (0.011)
Oromiya	0.286	-2.253*** (0.605)	-7.549*** (0.589)	0.065*** (0.011)		0.282	1.960*** (0.455)	-7.593*** (0.560)	0.156*** (0.011)
Somalie	0.019	-8.960** (0.835)	-2.443** (0.835)	-0.178** (0.017)		0.018	-3.758*** (0.650)	-3.431*** (0.900)	-0.065*** (0.016)
Benshangul Gumuz	0.009	-4.995*** (0.909)	-12.880*** (0.936)	0.107*** (0.019)		0.008	-1.089*** (0.731)	-12.200*** (0.907)	0.170*** (0.019)
SNNP	0.124	-1.779*** (0.612)	-8.131*** (0.629)	0.092*** (0.012)		0.108	1.020*** (0.521)	-8.514*** (0.650)	0.156*** (0.012)
Harari	0.009	-5.448** (1.076)	-2.818** (1.010)	-0.088** (0.021)		0.009	-0.900*** (0.813)	-4.370*** (1.040)	0.032*** (0.019)
Dire Dawa	0.025	-2.433* (1.247)	1.699* (1.090)	-0.078* (0.024)		0.025	1.687** (1.197)	3.282** (1.269)	0.005** (0.026)
Expected value		28.38	43.88	0.647			16.86	37.68	0.447
Number of observations			33790					40685	
Censored observations			13345					23463	
Wald chi2(27)			1465.2					3288.4	
Log pseudolikelihood			-8283110.1					-7907000.3	
Prob. > chi2			0.0000					0.0000	
Rho			-0.0851 (0.025)					-0.144 (0.068)	
Sigma			19.99 (0.140)					18.58 (0.178)	
Lambda			-1.701 (0.500)					-2.685 (1.272)	
Wald test of indep. eqns. (rho=0):									
chi2(1)			11.54					4.45	
Prob. > chi2			0.0007					0.0348	

Source: LFS 2005. Note: Individuals aged 15 and older. Standard errors are in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. Marginal effects and sample means are computed after estimating a generalized Tobit model for market work time accounting for sampling weights. Effect (1) is the total effect, i.e. the marginal effect for the unconditional expected value of market work time. Effect (2) is the conditional effect, i.e. the marginal effect for the expected value of market work time conditional on being observed. Effect (3) is the effect on participation, i.e. the marginal effect for the probability of market work time being observed.



**Table 3.6: Detailed marginal effects from generalized Tobit model for market work time (evaluated at sample mean) by gender in rural areas**

	Men				Women			
	Mean	Effect (1)	Effect (2)	Effect (3)	Mean	Effect (1)	Effect (2)	Effect (3)
Housework hours	5.770	-0.248*** (0.009)	-0.198*** (0.008)	-0.002*** (0.000)	37.230	0.013** (0.006)	0.008** (0.005)	0.000** (0.000)
Primary education	0.384	-2.821*** (0.324)	-2.138*** (0.294)	-0.027*** (0.005)	0.131	-0.101 (0.367)	-0.125 (0.269)	-0.001 (0.010)
General education	0.020	0.123* (1.110)	1.656* (1.013)	-0.039* (0.019)	0.006	6.383*** (2.074)	5.305*** (1.477)	0.103*** (0.038)
Beyond general education	0.007	3.103 (1.453)	0.228 (1.420)	0.085 (0.011)	0.003	11.882*** (2.804)	9.237*** (2.049)	0.181*** (0.043)
Potential experience	27.436	0.632*** (0.045)	0.362*** (0.043)	0.009*** (0.001)	27.173	0.702*** (0.039)	0.427*** (0.030)	0.019*** (0.001)
Potential experience squared	1075.609	-0.010*** (0.001)	-0.005*** (0.001)	-0.000*** (0.000)	1000.078	-0.011*** (0.001)	-0.007*** (0.000)	-0.000*** (0.000)
Training	0.025	4.663*** (1.062)	3.696*** (0.973)	0.036*** (0.014)	0.005	8.002*** (2.196)	5.050*** (1.610)	0.171*** (0.035)
Married	0.654	0.946 (0.496)	-0.605 (0.458)	0.044 (0.008)	0.662	-1.352*** (0.300)	-0.813*** (0.227)	-0.036*** (0.008)
Disable	0.035	-9.070 (0.888)	-0.464 (0.879)	-0.259 (0.020)	0.029	-6.104*** (0.592)	-2.809*** (0.503)	-0.220*** (0.022)
Nb infants under 6	1.050	0.620* (0.161)	0.296* (0.145)	0.010* (0.003)	1.054	-0.307** (0.121)	-0.202** (0.089)	-0.008** (0.003)
Nb children between 6 and 14	1.365	-0.307** (0.128)	-0.301** (0.116)	-0.001** (0.002)	1.305	-0.010 (0.095)	0.014 (0.069)	-0.001 (0.003)
Nb men between 15 and 24	0.701	-0.341 (0.210)	-0.185 (0.200)	-0.005 (0.003)	0.409	-0.790*** (0.167)	-0.513*** (0.122)	-0.020*** (0.005)
Nb women between 15 and 24	0.440	0.246 (0.243)	0.327 (0.230)	-0.001 (0.004)	0.604	0.489*** (0.182)	0.378*** (0.133)	0.011*** (0.005)
Nb men between 25 and 44	0.517	-0.835 (0.330)	-0.114 (0.317)	-0.022 (0.005)	0.433	-1.955*** (0.247)	-1.056*** (0.186)	-0.057*** (0.007)
Nb women between 25 and 44	0.533	-0.723 (0.356)	-0.209 (0.336)	-0.016 (0.006)	0.556	-0.338 (0.281)	-0.029 (0.205)	-0.015 (0.008)
Nb men between 45 and 64	0.328	-1.303* (0.397)	-0.791* (0.382)	-0.017* (0.006)	0.251	-2.154*** (0.287)	-1.243*** (0.210)	-0.060*** (0.008)
Nb women between 45 and 64	0.260	-0.687* (0.393)	-0.662* (0.375)	-0.003* (0.006)	0.290	-0.304 (0.307)	0.031 (0.230)	-0.015 (0.008)

Table 3.6: *continued*

	Mean	Effect (1)	Effect (2)	Effect (3)		Mean	Effect (1)	Effect (2)	Effect (3)
Nb seniors 65 and older	0.178	0.613** (0.421)	0.871** (0.407)	-0.005** (0.006)		0.182	-0.259 (0.296)	0.028 (0.222)	-0.013 (0.008)
Tigray	0.064	-7.820*** (0.792)	-9.086*** (0.816)	0.014*** (0.011)		0.063	0.280 (0.762)	-0.532 (0.599)	0.030 (0.019)
Affar	0.003	3.997*** (0.950)	3.742*** (0.923)	0.017*** (0.012)		0.002	0.809 (0.926)	1.972 (0.718)	-0.023 (0.025)
Amhara	0.300	-5.074*** (0.750)	-7.270*** (0.755)	0.047*** (0.009)		0.279	-1.454** (0.704)	-2.283** (0.550)	0.005** (0.019)
Oromiya	0.394	-8.123*** (0.746)	-10.086*** (0.742)	0.030*** (0.010)		0.405	-1.814*** (0.712)	-2.013*** (0.546)	-0.021*** (0.019)
Somalie	0.008	-5.667*** (0.888)	-6.340*** (0.905)	0.002*** (0.013)		0.008	-0.571 (0.832)	-0.308 (0.632)	-0.017 (0.023)
Benshangul Gumuz	0.012	-13.389*** (0.799)	-14.619*** (0.847)	-0.010*** (0.013)		0.012	-2.959*** (0.710)	-2.838*** (0.566)	-0.054*** (0.022)
SNNP	0.215	-6.719*** (0.730)	-9.451*** (0.744)	0.062*** (0.008)		0.228	0.522 (0.723)	-0.406 (0.557)	0.037 (0.018)
Harari	0.001	-12.705*** (0.877)	-13.547*** (0.921)	-0.024*** (0.016)		0.001	-0.359 (0.876)	-0.341 (0.669)	-0.006 (0.024)
Dire Dawa	0.002	-4.667*** (1.003)	-5.565*** (1.013)	0.012*** (0.013)		0.002	4.467*** (1.053)	2.786*** (0.801)	0.105*** (0.021)
Expected value		30.57	33.89	0.902			14.79	21.67	0.683
Number of observations			26973					29664	
Censored observations			3477					9454	
Wald chi2(27)			1887.2					1297.0	
Log pseudolikelihood			-50165173.8					-45817832.5	
Prob. > chi2			0.0000					0.0000	
Rho			-0.0924 (0.026)					0.997 (0.000)	
Sigma			16.50 (0.090)					18.24 (0.136)	
Lambda			-1.524 (0.422)					18.17 (0.137)	
Wald test of indep. eqns. (rho=0):									
chi2(1)			12.97					6424.63	
Prob. > chi2			0.0003					0.0000	

*Source:* LFS 2005. *Note:* Individuals aged 15 and older. Standard errors are in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. Marginal effects and sample means are computed after estimating a generalized Tobit model for market work time accounting for sampling weights. Effect (1) is the total effect, i.e. the marginal effect for the unconditional expected value of market work time. Effect (2) is the conditional effect, i.e. the marginal effect for the expected value of market work time conditional on being observed. Effect (3) is the effect on participation, i.e. the marginal effect for the probability of market work time being observed.

## **Chapter 4**

### **Assessing the Impact of Informality on Earnings in Tanzania: Is There a Penalty for Women?**

#### **4.1 Introduction**

According to recent economic figures, Tanzania seems to be on the right track. The average annual growth of the gross domestic product (GDP) stood at 6 per cent in the period 2000-2005 (Utz, 2008), which is remarkably high. This positive trend has been favoured by a series of macroeconomic and structural reforms initiated in the mid-1980s when Tanzania began its transition to a market economy.

While these measures have been fruitful, contributing in fine to the acceleration in economic growth, they have also had negative consequences such as the rapid growth of the informal sector.

The informal sector, which is an important entry point for the poor to engage in industrial and service sector activities, has undoubtedly contributed to poverty reduction, especially in rural areas where movements from traditional agriculture to other sources of income have been observed. Notwithstanding its beneficial overall effect on poverty alleviation and its contribution to economic growth, informal jobs are generally less secure and less rewarded than those found in the formal sector, and thus more likely refer to workers who are disadvantaged in the labour market and who potentially incur many risks given that work is usually their sole source of living.

In all developing regions of the world, informal employment today constitutes a very important share of total non-agricultural employment. In Sub-Saharan Africa it accounts for the vast majority of workers outside agriculture, and involves generally more women than

men, the former being in addition over-represented in the most precarious jobs within the informal sector (see Jütting and de Laiglesia, 2009).

Schneider (2004) analysed the size of the shadow economies of 145 countries all over the world, and found for Tanzania that the informal economy, including informal activities in agriculture, accounts for 60 per cent of the gross national income (GNI). The quasi-totality of economic activities is within extra-legal boundaries in the Tanzanian informal economy which has assets worth US\$29 billion, as revealed by the Instituto Libertad y Democracia (ILD, 2005). The National Bureau of Statistics of Tanzania found, in its analytical report of the Integrated Labour Force Survey 2000/01 (NBS, 2002), that the number of households with informal economy activities in urban areas has grown the last decade to reach 61 per cent of the total households in 2001, this negative trend being explained by the economic hardships, the inability to find formal jobs and the need for additional income that have forced families to engage in informal activities as a survival strategy.

In the 2006 World Economic Forum Global Gender Gap report (WEF, 2006), Tanzania was ranked number 1 out of 115 countries in terms of women's economic participation. As part of its commitment to achieving the third of the Millennium Development Goals (MDGs), which calls explicitly for the achievement of gender equality and the empowerment of women, Tanzania has addressed gender issues in a number of areas (see Ellis *et al.*, 2007). For instance, the Employment and labour Relations Act of 2004 prohibits discrimination in the workplace on the basis of gender, requires employers to promote equal opportunities, introduces maternity leave, and contains provisions protecting a mother's right to breastfeed and to be protected from engaging in hazardous employment.

Despite the positive legal framework and political context for gender equality, many barriers still hinder women's empowerment and contribution to the economy. Social and cultural norms strongly influence the ability of Tanzanian women to realize their potential, especially in the social and economic spheres. Indeed, mentalities and attitudes prevailing in the Tanzanian population, who lives predominantly in rural and traditional communities, impede women to take full advantage of their abilities and keep women subordinated to men. For

instance, the reluctance of husbands to allow their wives to engage in business activity and time constraints due to competing domestic responsibilities are common obstacles to women's economic empowerment.

We now well know that gender inequalities are detrimental to society at large, and that unlocking the full economic potential of women would importantly contribute to poverty reduction and growth stimulation (see Morrison *et al.*, 2007). Increased gender equality through better women's access to markets (labour, credit, land...), education and health, and through mother's greater control over decision-making in the household, would lead to increased women's labour force participation, productivity and earnings, and to improved children's well-being. These improvements would translate in fine into current poverty reduction and short-term growth thanks to higher incomes and consumption expenditures, and into future poverty reduction and long-term growth thanks to higher savings and better children's health and educational attainment, and productivity as adults.

The 2006 Integrated labour Force Survey is the most recent and well-suited nationally representative household survey to analyse informal employment in Tanzania according to the latest internationally-agreed statistical definition.

We use this dataset to address our research objective, which is to assess the impact of informality on earnings and determine whether, compared to men, women are penalized by working informally.

According to the theoretical literature on compensating earnings differentials<sup>1</sup>, firms have to offer a pecuniary compensation to encourage workers to bear unfavourable working conditions. If this theory turns out to be true in Tanzania, we should observe an informal employment earnings premium given that informal employed workers are generally exposed to worse conditions than those working formally.

---

<sup>1</sup> See, for instance, Brown (1980), Rosen (1986), and Murphy and Topel (1987).

Since the early 1970s, a number of different schools of thought on the informal economy have emerged, giving rise to discussion about its nature and drivers<sup>2</sup>. More recently, the discussion has focused on the causes of informal employment, taking a micro-perspective and looking at the motivation of individuals (see Maloney, 2004; Perry *et al.*, 2007). According to the exit (or voluntary) view, people or firms voluntarily exit from formality and choose to work informally as the result of weighting the costs and benefits brought by formality. Following this, we should expect our result to show that, in line with the theory of compensating earnings differentials, people get a pecuniary advantage from working informally.

In sharp contrast, the exclusion view claims that informal workers are denied access to formal jobs. The labour market is segmented due to the prohibitive cost of formality imposed on a subset of the population. According to this theory, we should instead find, at odds with the exit view, an informal employment earnings penalty.

A third view has emerged combining aspects of both exclusion and exit. According to this, in urban informal labour markets in developing countries there exist an “upper tier” and “lower tier” (Fields, 1990 and 2005). The upper tier has access requirements that make it unavailable to workers in the lower tier, and comprises those who voluntarily choose to be informal. The lower tier consists of individual who cannot afford to be unemployed but do not have access to more productive employment in either the formal or the informal sector. This view of the informal sector highlights the existence of segmentation within informal employment.

To adjudicate between these hypotheses, we conduct treatment effect analysis of informal employment on earnings, separately for men and women in paid- and self-employment, making three different assumptions for the treatment effect: (i) homogeneity, (ii) partial heterogeneity, and (iii) full heterogeneity of the population in the treatment response. We expect our results to advocate for the exclusion view by showing the following evidences. First, individuals have lower economic returns when working informally; second, negative

---

<sup>2</sup> See, for instance, Hart (1973) for the dualist school, Moser (1978), Castells and Portes (1989) for the structuralist school, de Soto (1989a, 1989b and 2000) for the legalist school, and Lewis (2004) for the parasitic school.

selection (or, in other words, negative sorting gain) to informal employment is at work, in the sense that it is individuals who are most likely to select into informal employment who lose the most from working informally.

We suspect that the hypothetical informal employment earnings penalty is significantly higher for women than for men. To explain such a result, we conjecture that the negative selection to informal employment, and therefore the exclusion hypothesis, is much more prevalent among women, i.e. women's decision to engage in informal employment is less exclusively and intentionally linked to economic gain and more the result of a constrained choice. This argument is quite intuitive to the extent that the decision to work informally is governed not only by rational choice but also by cultural and social norms, circumstances and constraints, which are believed to affect much more women than men. For instance, the burden of women's household responsibilities such as housework and childcare likely provide strong incentives for women to engage in informal employment, which generally offers lower earnings and protection but potentially more flexibility. Lack of adequate infrastructure is also believed to play a determinant role in explaining negative selection to informal employment, especially among women.

The structure of this Chapter is as follows. Data, concepts and descriptive statistics are documented in Section 2, and the methodology in Section 3. Section 4 presents the results and Section 5 concludes.

## **4.2 Data, concepts and descriptive statistics**

In this section, we start off by presenting the data used for our empirical analysis of treatment effects of informal employment on earnings in Tanzania. We further provide an explanation of the definitions and measures of key concepts such as employment, informal employment, and labour income. We finally provide a comprehensive picture of labour market structure and employment patterns across sexes, evidencing the multiple faces of informal employment and the over-representation of women in the more disadvantageous labour market segments.

#### **4.2.1 Tanzania Integrated Labour Force Survey 2006**

The 2006 Integrated Labour Force Survey (ILFS) is the fourth survey of its kind to be conducted by the National Bureau of Statistics (NBS) of Tanzania. This survey is a nationally representative household survey, collected during the four quarters of 2006 so as to capture seasonal variations, and is intended, among others, to meet the data needs for monitoring of the National Strategy for Growth and Reduction of Poverty (NSGRP/MKUKUTA). Its main objective is to provide up-to-date data on human economic activities, and particularly those related to the informal sector and its magnitude, unemployment, underemployment, child labour and time use.

Out of the 72 441 individuals who were interviewed in the ILFS<sup>3</sup>, 67 per cent were located in rural areas. Accounting for sampling weights, this figure rises to 75 per cent of overall population.

The ILFS consists of five questionnaires. In this Chapter we only rely on the first two. The first questionnaire (LFS1) includes information on human capital and other individual characteristics such as sex, age, marital status, migration, education and training, as well as information on household sources of income, assets, amenities and community services. The second questionnaire (LFS2) collects labour force details for individuals, i.e. information on usual and current activity, unemployment, main and other economic activities, informal sector, hours worked and income<sup>4</sup>. In sum, the 2006 ILFS offers relatively new, recent, detailed and reliable information on the situation of men and women in the Tanzanian labour market, an information useful in particular to deal with the purpose of this Chapter which is to analyse treatment effects of informal employment on earnings by employment status and gender.

---

<sup>3</sup> A three-stage sampling technique was used to select the sample survey. At the first stage, a simple random technique was adopted to determine representative samples of villages (140) and EAs (244) in rural and urban areas respectively. The second stage involved random selection of 80 households in each selected village and 30 households in each selected urban EA. This was then followed by the third stage of sampling which involved random selection of households to form representative samples of 20 and 30 households in each selected village and urban EA, respectively, which were interviewed in each quarter of the year 2006.

<sup>4</sup> The ILFS comprises in addition two questionnaires on child labour (CLS1 and CLS2) and a time use module (TUS) that collects information on the routine activities of respondents.



#### **4.2.2 Key concepts**

This section deals with definitions and measurement issues of the key concepts used in this study, which are employment, informal employment, and labour income. We implement the latest internationally agreed statistical definitions to measure these concepts in Tanzania using the ILFS 2006.

##### ***Employment***

All surveyed individuals aged 15 and above who declared doing any work of any type for pay, profit, barter or home use, during the last week prior to the date of interview, are considered to be part of the employed population. Those aged 15 and above who reported having a job or an own farm or an enterprise at which they did not work during the reference period but to which they will definitely return to, are also classified as employed. Tanzania is unusual – but in line with international recommendations – in considering collection of fuel and water as an economic activity<sup>5</sup>, with those who engage in this activity in the reference period being considered as employed.

##### ***Informal employment***

To measure informal employment in Tanzania, we make use of the international statistical definitions of employment in the informal sector and of informal employment, which were adopted by the 15<sup>th</sup> and 17<sup>th</sup> ICLS in January 1993 and December 2003, respectively (see Hussmanns, 2004).

Total employment is disaggregated according to two dimensions: type of production unit and type of job. Type of production unit is defined in terms of legal organization and other

---

<sup>5</sup> Human activities are often categorized on the basis of the System of National Accounts (SNA) which defines the rule that countries must use in calculating gross domestic product (GDP). On this basis, human activities can be divided into three categories: i) activities which are included in the production boundary of the SNA, ii) those which are recognized as work, but fall outside the SNA production boundary (extended SNA), and iii) non-productive activities. In the ILFS 2006, collection of fuel and water, which falls within the production boundary of the SNA, is included under this category even though such activities are not yet used in calculating Tanzania's national GDP (NBS, 2007).

enterprise-related characteristics, while type of job is defined in terms of status in employment and other job-related characteristics.

Using the information available in the ILFS 2006, we divide the employed population in the four following types of production unit:

- Formal sector enterprises comprise the following categories: central government, local government, parastatal, political party, partnership registered, nongovernmental organization, religion organization, cooperative registered and international/regional organization. It includes in addition individuals involved in other household economic activities (than fetching water or collecting firewood), private own account, private sector employed, and other private workers working in enterprises which have at least ten employees who are paid on a continuous basis, or working in enterprises which keep any written records or accounts (order book/record, sales book/records, accountancy book, final accounts for business year) showing all the balance sheets of assets and liabilities (investment/withdraw of capital and withdrawals of income by the business owner(s), earnings retained within the business as saving)<sup>6</sup>.
- Informal sector enterprises include cooperatives and partnerships unregistered as well as individuals involved in other household economic activities, private own account, private sector employed, and other private workers working in enterprises which have less than ten employees paid on a continuous basis, and which keep no written records or accounts, or keep them but without showing all the balance sheets of assets and liabilities.
- Households include all persons who were doing household-related work defined as economic activities such as collection of firewood and fetching water (other household chores are excluded).

---

<sup>6</sup> Workers who answered that they don't know whether their enterprise keeps any written records or accounts or whether these accounts show all the balance sheets of assets and liabilities, as well as workers for whom this information is simply missing, are considered, by default, to be part of formal sector enterprises.

- Traditional agriculture comprises all persons working on their own farm or shamba either in agriculture, livestock or fishing as either self-employed persons or unpaid family helpers. Traditional agriculture is considered as a separate category and *de facto* excluded from informal employment measurement<sup>7</sup>.

In accordance with the 15<sup>th</sup> ICLS (1993), we define employment in the informal sector as comprising all persons who are employed in an informal sector enterprise, irrespective of their status in employment.

Jobs are distinguished according to status in employment categories and according to their formal or informal nature. Based on employment status categories considered in the ILFS 2006, we divide the employed population in the five following types of job:

- Own account workers (outside traditional agriculture) are self-employed workers without employees. They are persons who perform work for profit or family gain in their own non-agricultural enterprise.
- Employers (outside traditional agriculture) are self-employed workers with employees. They are persons who perform work for profit or family gain in their own non-agricultural enterprise.

---

<sup>7</sup> There are several reasons why informal employment is usually measured excluding agriculture. First, the criteria for defining the informal sector (though it is less true for informal employment) are not adapted in the case of agriculture and usual data collection systems do not often distinguish formal and informal (or modern and traditional) agriculture. In fact, the recommendation of the 15<sup>th</sup> ICLS (1993) to exclude agricultural and related activities from the scope of informal sector surveys was made for practical data collection reasons. In our case, the survey definition of traditional agriculture, which accounts for the quasi-totality of agriculture workers, is quite weak in that it is based solely upon workers' employment status. Moreover, it was decided that the separate category was desirable as informal sector enterprises outside traditional agriculture would otherwise be lost in the dominant traditional agricultural group. Second, the shift from agricultural to non-agricultural activities is a sign of modernization. Therefore, it is preferable to distinguish between agriculture and non-agriculture (rather than rural/urban) to take account for major structural changes in developing countries (Jütting and de Laiglesia, 2009).

- Contributing family workers (outside traditional agriculture) are persons working completely without payment in cash or kind in family enterprises other than in the family farm or shamba.
- Employees are persons who perform work for a wage or salary in cash or kind.
- Traditional agriculture workers are persons working on their own farm or shamba either in agriculture, livestock or fishing as either self-employed persons or as unpaid family helpers. Again, traditional agriculture is considered as a separate category and *de facto* excluded from informal employment measurement.

The 17<sup>th</sup> ICLS (2003) defined informal employment as the number of informal jobs, whether carried out in formal sector enterprises, informal sector enterprises or households.

From an individual perspective, informal employment can be defined as comprising all persons who are employed in an informal job, irrespective of whether it is carried out in formal sector enterprises, informal sector enterprises or households.

Informal employment precisely comprises:

- Informal employment in the formal sector, which includes contributing family workers and informal employees in formal sector enterprises.
- Informal employment in the informal sector, which comprises own account workers, employers, contributing family workers and informal employees in informal sector enterprises.
- Informal employment in households, which includes own accounts workers, employers, contributing family workers and informal employees in households.

Informal employees are defined as those paid employees who have a casual oral work contract or who are not covered with any social security scheme.

### ***Labour income***

From the 2006 ILFS, we have the following information about labour income: (i) the gross cash income from workers' paid employment last month, (ii) the net income (difference between gross income/takings and expenses) from workers' business(es) outside agriculture in the last week or month, and (iii) the net income from workers' urban agricultural work in the last week or month<sup>8</sup>.

We first consider two categories of workers: paid- and self-employed workers (outside traditional agriculture). Then, we compute the hourly income for each of these two groups, and disaggregated it by gender and by formal versus informal employment. Gross cash hourly income from paid-employment is obtained by dividing the gross cash monthly income by the monthly hours of work. The latter is calculated by multiplying the number of hours worked each day last week prior the date of interview by 4.3. Net hourly income from self-employment is calculated in the same way, after computing the net monthly income (by multiplying the net weekly income by 4.3) for those who reported their net income in the last week instead of last month.

We focus exclusively on the hourly income from the main job due to the small number of multiple job holders in the sample survey that prevent us to obtain reliable income variables from the secondary job disaggregated for men and women in paid- and self-employment, and in formal and informal employment.

Assuming a lognormal distribution of hourly income from main job, we transform this variable by taking the natural logarithm of each of its values in order to analyse treatment effects of informal employment on log hourly income from main job by employment status (paid- and self-employment) and gender.

---

<sup>8</sup> We do not exploit this last income category because traditional agriculture workers, who comprise all persons working on their own farm or shamba as self-employed or unpaid family helpers, do not belong to neither formal nor informal employment.

### 4.2.3 Descriptive statistics on gender-differentiated employment patterns

Before documenting how men and women are distributed across labour market segments, we present in Table 4.1 some selected key labour market indicators.

In 2006, 85 per cent of the Tanzanian working-age population (individuals aged 15 and older) is employed, 12 per cent is inactive and less than 3 per cent is unemployed<sup>9</sup>. Almost half of workers hold multiple jobs. Statistics disaggregated by gender show that the employment ratio is lower, whereas the unemployment ratio and the inactivity rate are higher, for women than for men. In addition, female workers are 13 per cent more likely than male's to have one or more additional jobs. These findings give us a first insight of the gender disparities in Tanzania's labour market.

**Table 4.1: Selected key labour markets indicators**

	Men (%)	Women (%)	All (%)
Employment ratio	87.8	83.3	85.4
Multiple job holding rate	41.3	54.2	47.9
Unemployment ratio	1.7	3.6	2.7
Inactivity rate	10.5	13.1	11.9

*Source:* Tanzania ILFS 2006 *Note:* Individuals aged 15 and older. Sampling weights are accounted for.

Tables 4.2 and 4.3 present the distribution of workers by type of production unit and type of job (status in employment), respectively for men and women in main job.

---

<sup>9</sup> We define unemployed workers as all the working-age individuals (15+) with all the following characteristics:

- (i) they did not do any work of any type for pay, profit, barter or home use during the reference period (the calendar week prior to the date of interview),
- (ii) they did not have a job or own farm or enterprise to which they will definitely return to,
- (iii) they were available for work during the reference period,
- (iv) they had taken any steps within the last four weeks prior to the date of interview to look for work.

Inactive people comprise all those aged 15 and older who are neither employed nor unemployed (children are excluded from the analysis).

What is first appealing from these tables is the very high incidence of traditional agriculture and the very low incidence of formal employment among workers. 69, 24 and 7 per cent of male workers are respectively engaged in their main job in traditional agriculture, informal and formal employment. The corresponding figures for female workers are respectively 77, 20 and 3 per cent. In relative numbers, female workers are indeed less hit than male's by informal employment but they are clearly more involved in traditional agriculture, sufficiently to make them 4 per cent less likely to be part of formal employment as compared to their male counterpart.

Outside traditional agriculture, male workers are mostly participating in the formal sector (16 per cent), then in the informal sector (14 per cent), and very few work in households (1 per cent). For female workers the situation is quite different. Only 6 per cent are in the formal sector, while 11 per cent are engaged in the informal sector and 5 per cent work in households.

Tables 4.2 and 4.3 provide other interesting results. For instance, data reveal that the majority of workers' main jobs in the formal sector are informal (between 57 and 58 per cent), indicating that employees, who account for more than 80 per cent of workers in that sector, are for the most not covered by any social security scheme or have a casual oral work contract.

This brings us to the more general and unexpected result that a very high proportion of informal jobs are found outside the informal sector. More precisely, only 57 and 55 per cent of respectively male and female informal workers' main jobs are located in the informal sector. 38 per cent of male informal workers' main jobs are found in the formal sector and 5 per cent in households. The equivalent proportions for women are respectively 19 and 26 per cent.

Accordingly, overlooking what happens in the formal sector and in the households, the nature and the quality of the productive activities that are performed therein, would result in a very significant underestimation of the true value of the informal employment in Tanzania. In that country, there is a tiny minority of workers occupying formal jobs; these are lacking and

hardly accessible to common workers, most of the jobs occupied in the formal sector being unprotected jobs.

We are particularly interested in the composition of the informal sector and the informal employment. Having a closer look at the results displayed in Tables 4.2 and 4.3, we observe that the vast majority of men and women participating in the informal sector are own account workers, respectively 85 and 90 per cent. The second most frequent type of job in the informal sector is employer, which accounts for 12 per cent of men and 7 per cent of women working in that sector. Unpaid family helpers and employees represent a non-significant share of informal sector employment. With respect to unpaid family helpers, this is not surprising since the quasi-totality of them is found, as expected, in the households involved in economic activities such as collection of firewood and fetching water.

As regards informal employment, results show that for both men and women the proportion of own account workers among those who have an informal main job reaches 50 per cent. However, the remaining male and female informal workers are quite differently distributed among the other types of job. Indeed, 7 per cent of men occupying an informal main job are unpaid family helpers and 37 per cent are employees, while these numbers are respectively 28 and 18 per cent for their female counterpart.

For men and women alike, employers account for a small share of informal main job workers (7 and 4 per cent for men and women, respectively).



**Table 4.2: Distribution of workers by type of production unit and type of job (status in employment), men in main job**

		Formal employment				Informal employment					Traditional agriculture workers	Total
		Own account workers	Employers	Employees	Total	Own account workers	Employers	Unpaid family helpers	Employees	Total		
Formal sector	Row	6.0	4.3	31.9	42.2			2.3	55.5	57.7		100.0
	Column	100.0	100.0	100.0	100.0			22.3	97.7	37.8		15.9
	Cell	1.0	0.7	5.1	6.7			0.4	8.8	9.2		15.9
Informal sector	Row			0.0	0.0	85.2	12.4	1.2	1.2	100.0		100.0
	Column			0.0	0.0	99.8	100.0	10.1	1.9	57.4		13.9
	Cell			0.0	0.0	11.8	1.7	0.2	0.2	13.9		13.9
Households	Row			0.0	0.0	1.5	0.0	95.5	3.0	100.0		100.0
	Column			0.0	0.0	0.1	0.0	67.7	0.4	4.8		1.2
	Cell			0.0	0.0	0.0	0.0	1.1	0.0	1.2		1.2
Traditional agriculture	Row										100.0	100.0
	Column										100.0	69.1
	Cell										69.1	69.1
Total	Row	1.0	0.7	5.1	6.7	11.9	1.7	1.6	9.0	24.2	69.1	100.0
	Column	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Cell	1.0	0.7	5.1	6.7	11.9	1.7	1.6	9.0	24.2	69.1	100.0

Source: Tanzania ILFS 2006. Note: Individuals aged 15 and older. Sampling weights are accounted for. Dark grey cells refer to jobs that, by definition, do not exist in the employment sector in question.

**Table 4.3: Distribution of workers by type of production unit and type of job (status in employment), women in main job**

		Formal employment				Informal employment					Traditional agriculture workers	Total
		Own account workers	Employers	Employees	Total	Own account workers	Employers	Unpaid family helpers	Employees	Total		
Formal sector	Row	9.3	2.1	31.4	<b>42.8</b>			7.3	49.9	<b>57.2</b>		<b>100.0</b>
	Column	100.0	100.0	98.8	<b>99.1</b>			8.5	89.1	<b>18.5</b>		<b>6.4</b>
	Cell	0.6	0.1	2.0	<b>2.8</b>			0.5	3.2	<b>3.7</b>		<b>6.4</b>
Informal sector	Row			0.1	<b>0.1</b>	89.7	7.4	1.8	0.9	<b>99.9</b>		<b>100.0</b>
	Column			0.8	<b>0.6</b>	99.6	98.7	3.6	2.7	<b>55.3</b>		<b>11.0</b>
	Cell			0.0	<b>0.0</b>	9.9	0.8	0.2	0.1	<b>11.0</b>		<b>11.0</b>
Households	Row			0.2	<b>0.2</b>	0.7	0.2	93.3	5.7	<b>99.8</b>		<b>100.0</b>
	Column			0.4	<b>0.3</b>	0.3	1.3	87.9	8.2	<b>26.2</b>		<b>5.2</b>
	Cell			0.0	<b>0.0</b>	0.0	0.0	4.8	0.3	<b>5.2</b>		<b>5.2</b>
Traditional agriculture	Row										<b>100.0</b>	<b>100.0</b>
	Column										<b>100.0</b>	<b>77.4</b>
	Cell										<b>77.4</b>	<b>77.4</b>
Total	Row	<b>0.6</b>	<b>0.1</b>	<b>2.0</b>	<b>2.8</b>	<b>9.9</b>	<b>0.8</b>	<b>5.5</b>	<b>3.6</b>	<b>19.8</b>	<b>77.4</b>	<b>100.0</b>
	Column	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
	Cell	<b>0.6</b>	<b>0.1</b>	<b>2.0</b>	<b>2.8</b>	<b>9.9</b>	<b>0.8</b>	<b>5.5</b>	<b>3.6</b>	<b>19.8</b>	<b>77.4</b>	<b>100.0</b>

Source: Tanzania ILFS 2006. Note: Individuals aged 15 and older. Sampling weights are accounted for. Dark grey cells refer to jobs that, by definition, do not exist in the employment sector in question.

The picture that emerges from all these findings is that of an underdeveloped labour market, characterized by a great predominance of traditional agriculture, which affords very few formal job opportunities to secure a living and decent work conditions for the vast majority of Tanzanian workers, and especially for women.

The essential function of informal employment in poverty alleviation and the fact that informal activities are pursued in the absence of other options are well reflected in the data. Own account workers in the informal sector constitute the largest employment group among men and women in main job (respectively 12 and 10 per cent). Business owners were asked in the survey to report why they conduct their business. We restrict the sample to own account workers in the informal sector (by our definition, these are all considered to have informal jobs), and look at the main answer they give to this question, which is indicative of the main reasons that push individuals to engage in informal activities. We find that respectively 39 and 35 per cent of male and female own account workers in the informal sector declare that the main reason why they conduct their business is because they cannot find other work. Moreover, family's need of additional income constitute the most important reason to engage in this kind of activities for respectively 25 and 50 per cent of men and women in this employment group.

In light of the descriptive statistics presented so far, it becomes flagrant that, despite the positive legal framework and political context for gender equality, women in Tanzania are clearly worse-off than men in the labour market.

### **4.3 Methodology**

In this section, we describe the methodologies we implement to address our research objective, which is to assess the impact of informality on earnings in Tanzania and determine whether, compared to men, women are penalized by working informally. We conduct treatment effect analysis of informal employment on log hourly income from main job, separately for men and

women in paid- and self-employment, making three different assumptions for the treatment effect: (i) homogeneity, (ii) partial heterogeneity, and (iii) full heterogeneity of the population in the treatment response. To ease the reading, we skip in this section some detailed information on methodological and estimation issues, and report it to the end of the Chapter in Appendix A.

#### **4.3.1 Informal employment effects under the assumption of homogeneity**

We first depart from a (likely unrealistic) assumption of effect homogeneity, according to which the effect of informal employment is constant across individuals. We perform OLS regressions of log hourly income from main job, separately for men and women in paid- and self-employment. In all regressions, we restrict the sample to formal and informal workers outside agriculture, and we account for sampling weights. Formally, we use the following augmented version of the traditional human capital semi-logarithmic income equation developed by Mincer (1974):

$$y_i = \alpha + \delta D_i + \beta X_i + \varepsilon_i \quad (1)$$

Where  $y$  is the natural logarithm of hourly income from main job, which is observed only for paid- and self-employed workers outside traditional agriculture,  $D$  is a dummy variable for treatment status (which equals one if the individual has an informal main job, and zero if the individual has a formal main job), and  $X$  is a vector of income determinants that may also influence the probability of holding an informal main job, including human capital characteristics (dummies for three levels of educational attainment, and either age and its square, to take into account its possible decreasing returns, or dummies for two age groups – as proxies for potential

work experience <sup>-10</sup>, and a dummy for training), other individual characteristics (dummies for marital and migrant status), three quarterly dummy variables to account for seasonality (three last quarters of 2006), and urban and regional dummies to account for spatial heterogeneity. The parameters  $\beta$  are regression coefficients measuring the changes in log hourly income associated with changes in the income determinants  $X$ , and  $\delta$  is the treatment effect, i.e. the effect of holding an informal main job on log hourly income. Concretely, the exponential transformation of the regression coefficient  $\delta$  represents the multiplicative decrease in hourly income associated with having an informal main job, *ceteris paribus*, or holding all other factors constant<sup>11</sup>.  $\delta$  is assumed to be a constant parameter, invariant across all members of the population under study.  $\varepsilon$  is the

---

<sup>10</sup> Potential work experience is proxied by dummies for two age groups in male regressions whereas it is proxied by age and its square in female regressions. This choice is made for technical reasons, that is, to satisfy the balancing property (i.e. balancing of pre-treatment variables given the propensity score) when performing treatment effect analysis, as explained in Appendix A. Concern arises over the fact that work experience is deemed to be an essential factor determining labour income, and that, consequently, a comparison of the informal employment effect across sexes may be biased given that different proxies of work experience based on age are used for men and women. However, we do not account here for the true work experience (the ILFS 2006 does not provide information on that) but for the potential work experience, which is a quite weak measure of the former, especially among women, so we do not expect that having different proxies of potential work experience based on age will significantly affect the comparability of the informal employment effect between men and women.

<sup>11</sup> According to Halvorsen and Palmquist (1980), some articles misinterpret the coefficients of dummy variables in semi-logarithmic regression equations by assuming that the coefficient of a dummy variable, multiplied by 100, is equal to the percentage effect of that variable. This interpretation, which is correct for continuous variables, is not true for dummy variables and can result in substantial errors. The larger the coefficient of a dummy variable, the more important is the difference between the percentage effect and the coefficient. Given that the estimated coefficient  $\delta$  of the dummy variable for treatment status D is expected to be large, and in order to correct for standard deviation, we use the Kennedy (1981) correction method which can be expressed as follows: let  $\rho$  be the direct (marginal) returns of the dummy variable for treatment status D in the semi-logarithmic income equation (see Equation 1). We then have:

$$\rho = \left( e^{\delta - \frac{S_{\delta}^2}{2}} - 1 \right) * 100$$

Where  $\delta$  is the estimated coefficient of the dummy variable for treatment status D, and  $S_{\delta}$  its estimated standard error.

error term with an expected value of zero, and  $\alpha$  is the constant term to be estimated together with the other parameters  $\beta$  and  $\delta$ <sup>12</sup>.

Estimated informal employment effects will inform us on whether there is an informal employment income penalty, and whether women exhibit an income disadvantage relative to men from working informally. With this information, we will be able to come up with an initial conclusion on which of the exclusion or exit hypotheses prevails, and whether it is more acute among women.

However, the homogeneity assumption on which we base our results is very likely to be violated in practice. Assuming a constant effect of informal employment across individuals is not the best approach because it conceals systematic population heterogeneity. Accordingly, we turn to the next step of our investigation which is to relax the homogeneity assumption and conduct treatment effect analysis of informal employment under the assumption of partial heterogeneity.

#### **4.3.2 Informal employment effects under the assumption of partial heterogeneity**

We conduct heterogeneous treatment effect analysis using the Stratification-Multilevel method (SM-HTE), which assumes partial heterogeneity of the population in the treatment response<sup>13</sup>.

We first run propensity score Probit regression models predicting informal employment in main job, separately for men and women in paid- and self-employment. In all regressions, we restrict the sample to formal and informal workers outside agriculture, and we account for sampling weights.

Regressions include, in addition to the vector of income determinants  $X$  in Equation 1, some variables that are believed to affect informal employment participation decision because they reflect individual's constraints/advantages in terms of household responsibilities (number of

---

<sup>12</sup> In all regressions, we use the Huber/White sandwich estimator of the variance to correct for heteroskedasticity.

<sup>13</sup> This method has been applied in Xie et al. (2011), Brand and Davis (2011), Brand and Xie (2010), Brand (2010), Tsai and Xie (2008), and Xie and Wu (2005).

children aged 0-6 and aged 7-14 living in the household, and their squared value), household amenities and community services (whether the household owns a car; whether the household owns a cellular or a landline phone; whether electricity is the main source of energy of the household for either cooking, lighting or heating/cooling; whether the household has access to an improved drinking-water source; whether the household has an improved pit latrine; whether either bus, taxi or train are within a 30-minute (equivalent to 2 km) walk from the household).

Once the propensity score estimated, we construct balanced propensity score strata and estimate within each stratum the average treatment effect (level-1 model). The underlying assumption of the SM-HTE method is that we consider all treated and untreated within a stratum as homogeneous for estimating treatment effects. While the assumption of within-stratum homogeneity may not hold true in practice, it is less stringent relative to the strict homogeneity assumption described in Section 4.3.1<sup>14</sup>.

We then test for linear trend in treatment effects using variance-weighted least squares (level-2 model). Finally, we display the stratum-specific treatment effects and the estimated linear trend in a two-way graph. We do this separately for men and women in paid- and self-employment.

In a level-1 model, to estimate propensity score stratum-specific treatment effects within strata (level-1 slopes), we apply linear regressions within strata to further adjust for any remaining covariate imbalance within strata (control variables for within-strata models comprise all the regressors included in the vector  $X$  of Equation 1)<sup>15</sup>.

In a level-2 model, we evaluate a trend across strata by regressing the stratum-specific treatment effects (level-1 slopes) on strata rank using variance-weighted least squares – the variance is

---

<sup>14</sup> With the SM-HTE method, the treatment effect  $\delta$  (as well as the constant term  $\alpha$  and the other regression coefficients  $\beta$ ) of Equation 1 is assumed to be a constant parameter within each propensity score stratum, but not across propensity score strata.

<sup>15</sup> In all within-strata regressions, we account for sampling weights and we use the Huber/White sandwich estimator of the variance to correct for heteroskedasticity. Because we do not constrain the comparison of the treatment group and the control group across strata in any way, data analysis at this stage is non-parametric across strata.

based on the standard errors of the stratum-specific treatment effects<sup>16</sup>. The level-2 intercept is the predicted value of the effect of informal employment for the lowest propensity individuals, and the level-2 slope is the change in the effect of informal employment with each one-unit change to a higher propensity score stratum.

We thus model the heterogeneity pattern as a linear function across strata ranks in order to force the data to tell us whether the treatment effect is either a positive or a negative function of propensity.

In the two-way graph, ‘dots’ represent point estimates of level-1 slopes (i.e., stratum-specific OLS regression effects of informal employment on log hourly income from main job), and the linear plot is the level-2 variance-weighted least squares slope.

Estimates obtained with the SM-HTE method will inform us on whether there is an informal employment income penalty, whether women exhibit an income disadvantage relative to men from working informally, and how informal employment returns by gender evolve across propensity score strata. With this information at hand, we will have new evidence to adjudicate between the exclusion versus the exit hypotheses among men and women.

The SM-HTE method poses, however, two major drawbacks. First, the assumption of within-stratum homogeneity, as said before, may not hold true in practice. Second, modelling the heterogeneity pattern as a linear function across strata ranks is very restrictive. Linearity should only be taken as the first-order approximation of a trend. Further investigation is needed for detecting potential non-linearity.

---

<sup>16</sup> Variance-weighted least squares (VWLS) regression differs from ordinary least squares (OLS) regression in that it does not assume homogeneity of variance, but requires that the conditional variance of the dependent variable be estimated prior to the regression. The estimated variance need not to be constant across observations and is treated as if it were the true variance when computing standard errors of parameter estimates.



### 4.3.3 Informal employment effects under the assumption of full heterogeneity

In the last step of our investigation, we relax the strict homogeneity assumption and conduct heterogeneous treatment effect analysis using the Matching-Smoothing method (MS-HTE), which allows for full heterogeneity of the population in the treatment response<sup>17</sup>.

In contrast to the Stratification-Multilevel method (SM-HTE), the MS-HTE method, which is a non-parametric counterpart of the former, does not assume a global functional form on the heterogeneity in treatment effects and allows for heterogeneous treatment effects as a continuous function of propensity score rather than imposing homogeneity within strata.

Using the MS-HTE method, we first estimate propensity score as discussed in Section 4.3.2. We then match treated and untreated units with control units using different matching algorithms, and we plot the matched differences between treated and control units (treatment on the treated) and between control and untreated units (treatment on the untreated) against a continuous representation of the propensity score<sup>18</sup>. We finally fit nonparametric smoothed curves to obtain the pattern of treatment effect heterogeneity as a function of the propensity score<sup>19</sup>.

Only observations in the common support region are used for calculating treatment effects. Parameter estimates of the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the untreated (ATU), are bootstrapped by resampling the data with replacement 100 times<sup>20</sup>. By definition, ATE is the effect of randomly assigning a person to informal employment, ATT is the effect of treatment for informal workers, compared with what they would experience if they were working formally, and ATU is the effect

---

<sup>17</sup> This method has been proposed in Xie et al. (2011).

<sup>18</sup> Treated units are taken as the target group for matching in order to estimate treatment on the treated, and untreated units are taken as the target group for matching in order to estimate treatment on the untreated.

<sup>19</sup> We perform kernel-weighted local polynomial regressions accounting for sampling weights to construct the curves. We select the Epanechnikov kernel function in calculating the weighted local polynomial estimate, and we choose a one-degree polynomial for smoothing. A small amount of spherical random noise is added to each point before graphing in order to reduce over-plotting.

<sup>20</sup> The mean and standard error of each parameter estimate are computed from the 100 resampled estimates.

of treatment for formal workers, compared with what they would experience if they were working informally.

There is no clear consensus as to which matching estimator performs best in each application. We thus implement various propensity score matching methods for the sake of illustration and comparison, and as a way to assess the robustness of the estimates<sup>21</sup>.

The purpose of using the SM-HTE method is to confirm or, at least, reinforce the conclusions we draw from the previous method (MS-HTE), and to examine the potential non-linearity of the heterogeneity pattern of informal employment returns.

## 4.4 Results

In this section, we use the methods described previously to assess the impact of informality on earnings in Tanzania and determine whether, compared to men, women are penalized by working informally<sup>22</sup>.

### 4.4.1 Informal employment effects under the assumption of homogeneity

Appendix C Table 4.6 displays OLS regressions estimates of log hourly income from main job by employment status and gender. Results evidence a significant informal employment income penalty. Indeed, using the Kennedy correction (cf. footnote 11), we find that the decrease in hourly income associated with having an informal main job is, *ceteris paribus*, of 46.8 per cent for paid-employed men and 48 per cent for paid-employed women, and of 21.6 per cent for self-employed men and 7.8 per cent for self-employed women.

---

<sup>21</sup> The matching algorithms we use are the followings: single nearest neighbour matching with replacement but without caliper, radius matching within a radius size of 0.1, and Epanechnikov kernel matching with a bandwidth of 0.06.

<sup>22</sup> The list and definition of all the covariates we use in our regression models are presented in Appendix B Table 4.4.

The negative marginal returns to informal employment for paid-employed workers are large, statistically significant at 1 per cent level of significance, and only slightly higher for women as compared to men.

In turn, the negative marginal returns to informal employment for self-employed workers are much smaller (especially for women), and are statistically significant, but only at the 95 per cent confidence level, for men, and not statistically significant for women. In other words, holding all other factors constant, on one hand, self-employed women working informally have an expected income which do not significantly differ from that of those working formally, while, on the other hand, for self-employed men the expected income significantly differs with respect to formal-informal status, but the estimated negative effect of informal employment is not statistically significant at the highest confidence level and its absolute value is substantially lower than that for paid-employed men<sup>23</sup>.

Overall, observed results advocate for the exclusion hypothesis for both men and women in paid-employment but not in self-employment. Moreover, they do not evidence that, among paid-employed workers, the exclusion hypothesis is clearly more prevalent among women.

---

<sup>23</sup> Parameter estimates of our OLS regressions of log hourly income from self-employment need to be considered cautiously because pseudo R-squared are very low, which means that observed characteristics failed in explaining most of the variance of log hourly income across individuals, suggesting that our model is not properly specified and that some relevant variables are omitted. In our model, we include as regressors a set of human capital and other individual characteristics, as well as some controls to account for seasonality and spatial heterogeneity. However, self-employment income is typically generated at the household level, making it difficult to analyse it in relation to individual characteristics. In addition, self-employment income is usually derived not only from human capital but also from physical capital. It is important to account for enterprise attributes, because they arguably constitute crucial determinants of income dispersion in self-employment. For that purpose, the following enterprise characteristics could be, for instance, considered: employment size, value of physical capital, and access to credit. Given that the number of employees is already used to discriminate between formal and informal sector enterprises, we prefer not to include this variable in our regressions, because otherwise it would capture part of the effect of informal employment on self-employment income. Note, moreover, that, in the case of Tanzania, employment size of the enterprise is not so relevant because the vast majority of self-employed are own-account workers, i.e. self-employed workers without employees. As regards the two other suggested enterprise characteristics, the ILFS 2006 does not provide information on the value of physical capital while it contains some information on access to credit, but this is available only for enterprises that operate informally. Another point, which is worth mentioning, is that the vast majority of self-employed workers are informal. The significance and value of marginal returns to informal employment for self-employed workers is probably affected by the small number of cases of men and women working formally in our regression samples.

The fact that income prospects of paid-employed workers differ substantially with respect to formal/informal status is due, in part, to differences in background characteristics between formal and informal workers<sup>24</sup>. Appendix B Table 4.5 displays the mean values of the covariates we account for in the propensity score Probit regression models predicting informal employment. These help us to elucidate the characteristics of a typical individual within each worker category and figure out if, in line with the exclusion hypothesis, decision to engage in informal employment is less exclusively and intentionally linked to economic gain and more the result of a constraint choice. In addition to human capital and other individual characteristics, controls for seasonality and spatial heterogeneity, we also consider factors that reflect individual's constraints/advantages in terms of household responsibilities, household amenities and community services, and, therefore, may influence informal employment participation decision. Informal paid-employed workers, in particular women, are younger, as compared to their formal counterparts, which is illustrative of the increasing informalisation of employment and of the fact that, since age can be seen as a proxy for potential work experience<sup>25</sup>, informal activities, especially those performed primarily by women, are associated with low efficiency and productivity and, therefore, are less skill- and experience-demanding. In other words, older workers enjoy better employment prospects than younger workers because of their competitive advantage, as a result of their higher work experience, and the deterioration of labour market conditions for new entrants. Informal paid-employed workers, especially women, have also lower

---

<sup>24</sup> Note that, given we control for some observed individual's background characteristics in income regressions, marginal return to informal employment can be regarded as a "direct effect" in the sense that it does not capture, in principle, the indirect impact of these characteristics on income through their effect on informal employment participation. Note, moreover, that individual characteristics and a simplistic formal/informal employment divide cannot single-handedly determine income dispersion. There is increasing recognition that wage gaps are, in a large extent, due to a sorting process by which individuals get allocated to different types of jobs and different types of firms depending on their individual characteristics. For instance, Fafchamps *et al.* (2008) investigate if there is job sorting in African labor markets and find that much of the wage gap associated with education is driven by selection across occupations and firms.

<sup>25</sup> Because women often participate discontinuously in the labor market due to their competing domestic responsibilities, age is probably a proxy of potential work experience that is not as suitable for women as it is for men who usually have a higher labor force attachment.

education and training qualifications, i.e. poorer human capital endowments that result in lower employability. As an aside, it is worth noting that, among formal paid-employed workers, women are relatively more numerous than men to be highly educated, suggesting that access to formal jobs is more competitive for them.

Marriage and procreation are usually intimately linked, which explains why paid-employed women, in addition of having a lower occurrence of marriage, have also fewer infants than their male counterparts. The number of infants (aged 0-6), whose presence may raise the demand for services in the home, is slightly higher for informal paid-employed workers, who usually benefit from more flexibility to deal, for instance, with their household responsibilities, than for those working formally. Babies and infants are completely dependent on adults to supply their essential needs. Rearing infants means performing many daily activities to take care of them that are altogether really time- and energy-consuming. Within household with infants, women are more heavily burdened than men because they are more committed to children's care and welfare.

By contrast, informal paid-employed workers, and in particular men, have fewer children (aged 7-14) at home than those holding a formal job. In fact, many children are commonly involved in fetching water, firewood collection and other household-related activities, at the expense of attending school or doing homework. Beyond poverty and low education, cultural norms also have an impact on child labour since these specific tasks are culturally designated as children's work. The performance of these household chores is more typically the domain of female children. Thus, living with children is probably advantageous for paid-employed workers, especially for men, who can engage more easily in less flexible but more rewarding jobs in formal employment, because the presence of children releases them from part of the housework they would have to do otherwise.

Informal paid-employed workers do neither possess household amenities nor enjoy community services as formal paid-employed workers do. Indeed, they are, for instance, less likely to own a car and a cellular or a landline phone, to use electricity as the main source of energy for either

cooking, lighting, or heating/cooling, and to have access to local public transport services (either bus, taxi or train)<sup>26</sup>. Lack of adequate infrastructure imposes greater work burdens and lengthens the time it takes people, in particular women, to perform activities related to household survival, reducing the time for participating in other economic or income-generating activities. Lower availability of facilities and deficient infrastructure increase the opportunity cost of time and incur in many additional costs constituting important barriers that impede individuals' engagement in more rewarding jobs in formal employment.

All in all, the following tentative conclusion on paid-employment can be drawn from the findings presented in this section. The decision to work informally is probably not the mere result of a rational choice in which people weigh the expected economic returns against the costs based solely on economic factors. This decision is very likely to be governed also by cultural and social norms, circumstances and constraints. As seen in this section, informal workers, and in particular women, are affected by a range of constraints, such as, for instance, the burden of household responsibilities and the lack of adequate infrastructure, that constitute important obstacles which may explain why they participate in informal employment. Accordingly, working informally is likely to be less exclusively and intentionally linked to economic gain and more the result of a constrained choice. In other words, workers engaged in informal jobs find themselves in this situation probably not because of a potential competitive advantage but because, according to the exclusion hypothesis, they are denied access to formal jobs due to the disproportionate constraints

---

<sup>26</sup> With regard to Tanzania's recent acceleration in economic growth, analysis of factor inputs suggests that this acceleration is not so much grounded in a rapid expansion of human and physical capital but is primarily due to an increase in cultivated land in the agriculture sector and increased factor productivity for the other sectors. In particular, the analysis of public investment, which has recovered from an average of about 3 per cent of GDP during the late 1990s to about 8 per cent of GDP in recent years, suggests that only about one-third of it was used on public infrastructure such as roads or electricity, while the remainder was devoted to the rehabilitation and expansion of administrative and social infrastructure. For instance, in Tanzania's export processing zones, infrastructure weaknesses, especially reliable access to electricity and water, are perceived as a main constraint for firms located in those zones. Empirical analysis indicates that the main constraints in infrastructure are related to electricity, water, road transportation, and factors such as telecommunications, Internet connectivity, and waste disposal facilities. The costs for telecommunications services and electricity are particularly high in Tanzania (see Utz, 2008).

they face – they cannot afford to be unemployed but they do not have access to more productive employment in the formal sector –, and have no choice but to work in informal employment, which is mostly a residual where activities are being pursued in the absence of other options.

#### **4.4.2 Informal employment effects under the assumption of partial heterogeneity**

The marginal returns to informal employment, presented and discussed in the previous section, are computed as sample average effects. By doing this, we implicitly assume homogeneity of effect on all individuals. But, the homogeneity assumption is very likely to be violated in practice. Assuming a constant effect of informal employment across individuals is, indeed, not the best approach because it conceals systematic population heterogeneity. Accordingly, we turn to the next step of our investigation which is to relax this strict homogeneity assumption and conduct heterogeneous treatment effect analysis of informal employment using the Stratification-Multilevel method (SM-HTE), which assumes partial heterogeneity of the population in the treatment response.

Appendix D Table 4.7 displays the estimation results of the Probit models for the probability of being in informal relative to formal employment, separately for men and women in paid- and self-employment. As covariates, in addition to human capital and other individual characteristics, controls for seasonality and spatial heterogeneity, we also include factors that reflect individual's constraints/advantages in terms of household responsibilities, household amenities and community services, and that, therefore, may influence informal employment participation decision.

As regards self-employment, most of explanatory variables are not statistically significant and pseudo R-squared are quite low, especially in the Probit regression for men, which should make us humble about models' explanatory ability. Consequently, we adopt a conservative approach

and decide to focus on findings for paid-employed workers while leaving momentarily aside results for self-employed workers.

As expected, human capital characteristics (potential work experience, education and training) are statistically significant and important negative predictors of participation in informal employment among paid-employed workers, and especially among paid-employed women, whereas factors that reflect individual's constraints/advantages, in terms of household responsibilities, household amenities and community services, are in most cases not or poorly statistically significant. This, by no means, implies that these factors don't play any role in the decision-making process of working informally. It is rather a reflection of some kind of models' misspecification. On this point, it is worth precisising that some of the variables included as regressors are likely to be endogenous. For instance, fertility may be endogenous to choices regarding labour allocation between formal and informal jobs. If fertility and labour allocation decisions are jointly determined, inferences of causality from the former to the latter will be incorrect. The same holds for household amenities and community services<sup>27</sup>.

Appendix D Tables 4.8 and 4.9 probably give us a better idea of the factors that influence the likelihood to work informally in paid-employment. They provide the mean covariate values by propensity score stratum and formal-informal status, respectively for male and female paid-employed workers.

From these tables, it appears that paid-employed workers with the lowest propensity to have an informal job (first propensity score stratum), have the highest human capital endowments

---

<sup>27</sup> As an example, the causality between infrastructure and informal employment participation is evidently two-way. While infrastructure influences firms' productivity and formalization, and, hence, creation of formal jobs, higher productivity formal firms would also chose to locate in infrastructure abundant locations. Infrastructure variables are likely to be endogenous as a firm's location decision depends on perception of infrastructure availability. A region with growing firms will invest more in infrastructure than otherwise. Firms decide where to locate based on access to infrastructure that in turns affects their productivity. In other words, on one hand, better access and availability of infrastructure raise firms' productivity, their propensity to formalize and people's likelihood to get a formal job, whereas, on the other hand, the presence of higher productivity formal firms that generate formal employment increases public and private investors' propensity to improve infrastructure services.



(potential work experience, education and training) and are the most advantaged in terms of household amenities and community services. As we move to upper propensity score strata, human capital endowments and availability of household amenities and community services diminish until the last propensity score stratum (paid-employed workers with the highest propensity to have an informal job), which is characterized by the predominance of men and women with substantially lower potential work experience, education, training, and access to infrastructure services such as telecommunications, electricity and public transport<sup>28</sup>. For both men and women, the occurrence of marriage is much lower in the last than in the first propensity score stratum, and the number of infants tends to increase, while the number of children tends to decrease, with the propensity to have an informal job.

Overall, the above findings reinforce our idea that the decision to work informally in paid-employment is likely to be less exclusively and intentionally linked to economic gain and more the result of a constrained choice. Indeed, the higher is the propensity to work informally, the higher are the constraints in terms of household responsibilities and lack of adequate infrastructure, suggesting that the exclusion hypothesis is at work.

Appendix D Table 4.13 displays estimation results of the informal employment effects under the assumption of partial heterogeneity<sup>29</sup>. Results indicate a clear increase in the income-decreasing

---

<sup>28</sup> It is worth noting that, among paid-employed workers in the highest propensity score strata, women seem not to be disfavoured as compared to men but, on the contrary, they are more educated and have better access to infrastructure services. These findings have to be relativized because men's final stratum includes those with a propensity score of at least 90 per cent, while, in contrast, women's final stratum accounts for those with a propensity score of at least 80 per cent. However, it is important to make clear that, given the underdeveloped nature of the Tanzanian labour market, which affords very few good job opportunities to secure a living and decent work conditions for the vast majority of workers, and especially for women who face many barriers that hinder their empowerment and contribution to the economy, being paid-employed, even if it is in informal employment, is, per se, a good outcome for women. What these results are also probably telling us is that, in paid employment, access not only to formal jobs, but also to informal jobs, is more competitive for women than for men.

<sup>29</sup> Appendix D Table 4.13 displays estimates of the propensity score stratum-specific treatment effects within strata (level-1 slopes), and an estimate of the level-2 variance-weighted least squares slope obtained by regressing the level-1 slopes on strata rank, separately for men and women in paid- and self-employment. Appendix D Figure 4.1 offers a graphical representation of the heterogeneous effects of informal employment on log hourly income from main job by employment status and gender using the SM-

effect of informal employment across propensity score strata for paid-employed women. All the coefficient estimates are statistically significant at the highest confidence level and reach sizeable negative values. A unit change in propensity score stratum rank is associated with a 12 per cent increase in the informal employment effect, such that the predicted negative effect of informal employment on hourly income from main job among paid-employed women ranges, in absolute value and using the Kennedy correction, from 35 per cent in the first propensity score stratum to 57 per cent in the last propensity score stratum. Accordingly, women most constrained by household responsibilities and lack of adequate infrastructure, i.e. women with the highest propensity to have an informal main job (last propensity score stratum), benefit less from working informally, on the magnitude of an estimated 22 per cent, than would women who are least burdened by household responsibilities and most advantaged in terms of household amenities and community services, i.e. women with the lowest propensity to have an informal main job (first propensity score stratum). All in all, informal employment participation of paid-employed women is associated, as expected, with a significant decreasing economic return across propensity score strata, which offer support for the exclusion hypothesis.

As regards paid-employed men, all the informal employment effects are statistically significant at the highest confidence level and reach sizeable negative values as well, but, contrary to paid-employed women, who exhibit a fairly linear trend in informal employment effects, paid-employed men display an oscillating return to informal employment across propensity score strata. Their predicted negative effect of informal employment on hourly income from main job ranges, in absolute value and using the Kennedy correction, from 49 per cent in the first propensity score stratum to 54 per cent in the last propensity score stratum. The pattern of return to informal employment for paid-employed men seems to follow, more or less, an inverted-U shaped profile. Nonetheless, those with the highest as compared to those with the lowest

---

HTE method, where ‘dots’ represent point estimates of level-1 slopes and the linear plot is the level-2 variance-weighted least squares slope.

propensity to have an informal main job (last vs. first propensity score stratum), appear to benefit less from working informally, on the magnitude of an estimated 5 per cent. The estimated linear downward trend in informal employment effects across propensity score strata is much less steeper for paid-employed men than for paid-employed women. Indeed, the change in the effect of informal employment, with each one-unit change to a higher propensity score stratum, does not exceed 3 per cent for paid-employed men. Moreover, the estimated linear downward trend for paid-employed men is statistically significant, but only at the 90 per cent confidence level. Accordingly, our results evidence, but with a relatively low confidence interval, a small decreasing return of informal employment across propensity score strata for paid-employed men. Therefore, though these results nevertheless suggest the prevalence of the exclusion hypothesis among paid-employed men, they are clearly not as supportive of this hypothesis as compared to the results we obtain among paid-employed women.

In sum, we observe that, relaxing the unrealistic assumption of strict homogeneity and allowing for partial heterogeneity, paid-employed workers exhibit a significant informal employment income penalty, which is disproportionately higher for women.

While the exclusion hypothesis among paid-employed workers is pretty much likely to happen and dominate the alternative exit hypothesis, this phenomenon seems to be much more prevalent among females. Women are markedly disadvantaged compared with men, in that, with respect to the latter, their decision to work informally is very likely to be more governed by a range of factors that constitute important obstacles to access better jobs, such as the burden of household responsibilities and the lack of adequate infrastructure. This may explain why the significant informal employment income penalty in paid-employment is higher for women than for men.

With regard to self-employed workers, the quasi-totality of the coefficient estimates is not statistically significant. As a consequence, we abstain from drawing any conclusion for this workers' category.

#### 4.4.3 Informal employment effects under the assumption of full heterogeneity

We turn to the last step of our investigation where we conduct heterogeneous treatment effect analysis using the Matching-Smoothing method (MS-HTE) which allows for full heterogeneity of the population in the treatment response.

Appendix E Table 4.14 displays the average treatment effects (ATE), the average treatment effects on the treated (ATT), and the average treatment effects on the untreated (ATU), obtained using different matching algorithms, separately for men and women in paid- and self-employment.

It first appears that almost none of the treatment effects for self-employed workers are statistically significant. Because no reliable conclusions on the self-employed can be inferred, from now on, we limit ourselves to interpreting the results obtained for the paid-employed.

Treatment effects for paid-employed workers are all statistically significant at the highest confidence level. Recall that, in section 4.4.1, OLS estimates reveal that, under the assumption of homogeneity, women in paid-employment do not exhibit an obvious income disadvantage relative to men from working informally. By contrast, we now observe that all our ATE and ATT estimates<sup>30</sup> are markedly higher, in absolute value, for women than for men.

Treatment effects estimates are more reliable than OLS regression estimates, because the causal interpretation of the latter relies on a number of strong assumptions that are unlikely to hold true, in particular, as already discussed, the stringent assumption of homogeneity<sup>31</sup>. Accordingly, we conclude, on the basis of our estimates of treatment effects, that, in paid employment, women face a significantly higher informal employment income penalty than men.

Appendix E Table 4.14 also shows that, no matter the matching algorithm used, for paid-employed men and, to a much higher extent, for paid-employed women,  $ATT < ATE < ATU$ .

---

<sup>30</sup> Especially when using radius and kernel matching.

<sup>31</sup> In fact, in the presence of heterogeneity and self-selection, the use of conventional methods, not only OLS but also IV (instrumental variable), may yield a biased estimate of the treatment effect and, therefore, fail in accurately reflecting the causal effect of the treatment.

Accordingly, paid-employed workers and, especially, paid-employed women who work informally are those who get less out of it, whether compared to paid-employed workers who work formally ( $ATT < ATU$ ) or to randomly assigned paid-employed workers to informal employment ( $ATT < ATE$ ). These findings imply a negative sorting gain<sup>32</sup> for informal paid-employed workers, which is much more acute among women. Moreover, they confirm or, at least, reinforce our conclusion, made in the previous section using the Stratification-Multilevel (SM-HTE) method, according to which the exclusion hypothesis among paid-employed workers is at work, and is much more prevalent among women, which may explain why, in paid-employment, women face a significantly higher informal employment income penalty than men<sup>33</sup>. Appendix E Figures 4.2 and 4.3 provide, respectively for paid-employed men and women, a graphical representation of informal employment effects under the assumption of full heterogeneity<sup>34</sup>. The different matching algorithms yield similar results. We still observe higher income-decreasing effect of informal employment as the propensity of paid-employed workers, especially women, to work informally increases. However, the curves evidence that for men and, much more particularly, for women, the heterogeneity pattern of informal employment returns is not strictly linear.

To better illustrate this, we draw again the curves for paid-employed men and women (see Appendix E Figures 4.6 and 4.7, respectively), but now we suppress superimposing a scatter-plot

---

<sup>32</sup> i.e., negative treatment-effect heterogeneity bias or type II selection bias.

<sup>33</sup> Two additional points deserve to be noted. First, the fact we observe  $ATE \neq ATT \neq ATU$  also evidences that the homogeneity assumption does not fit our data, and neither does the conventional approach assuming a homogeneous treatment effect (OLS or IV). Second, treatment effects, in most cases, do not significantly differ depending on propensity score matching algorithm, which attest to the robustness of our estimates.

<sup>34</sup> Appendix E Figures 4.2 and 4.3, which offer, respectively for paid-employed men and women, a graphical representation of heterogeneous effects of informal employment on log hourly income from main job using the MS-HTE method, plot the matched differences between treated and control units (right-hand side graphs) and between control and untreated units (left-hand side graphs) against a continuous representation of the propensity score, and fit nonparametric smoothed curves to obtain the pattern of treatment effect heterogeneity as a function of the propensity score.

over the smooth<sup>35</sup> and we shorten the range of y-axis values considerably. Of course, doing this amounts to checking out our results with a magnifying glass, which brings into the light noticeable differences in the heterogeneity pattern depending on the matching algorithm being used.

Observed variations in the heterogeneity pattern when using radius versus kernel matching<sup>36</sup> appear to be relatively small and not so pronounced as to seriously detract from the robustness of our results and subsequent conclusions that can be drawn. Overall, we find that the heterogeneity pattern is, indeed, clearly not linear among paid-employed men and, especially, among paid-employed women. Accordingly, as expected, neither the assumption of homogeneity nor the assumption of partial heterogeneity (and, in particular, the linearity assumption of the heterogeneity pattern) of the population in the treatment response seem to hold in our data for Tanzania. Indeed, paid-employed workers clearly differ not only in their background characteristics, but also in how they respond to a particular treatment, which is to work informally. Assuming full heterogeneity of the population is therefore our preferred approach as it more adequately portrays and captures the situation observed in practice.

To sum up, results obtained with the assumption of full heterogeneity confirm our conclusions drawn under the assumption of partial heterogeneity. Paid-employed workers exhibit a significant informal employment income penalty, which is clearly higher for women than for men. The exclusion hypothesis is at work among paid-employed workers, and is much more prevalent among women, which may explain why, in paid-employment, women face an income disadvantage relative to men from working informally. The full heterogeneity assumption seems

---

<sup>35</sup> This option is very useful in our case because the number of plotted points is so large as to clutter the graphs.

<sup>36</sup> Though there is no clear consensus as to which matching estimator performs best in each application, we assume that radius and kernel matching methods are more reliable than single nearest neighbour matching (see Appendix A).

also to be strongly supported by our data and is therefore our preferred approach, while we do not find any robust ground for the homogeneity and partial heterogeneity hypotheses.

#### **4.5 Conclusion**

Using the Tanzanian 2006 Integrated Labour Force Survey and the latest internationally-agreed statistical definition for the measurement of informal employment, this Chapter has sought to assess the impact of informality on earnings and determine whether, compared to men, women are penalized by working informally. To this end, we conducted treatment effect analysis of informal employment on log hourly income from main job, separately for men and women in paid- and self-employment, making three different assumptions for the treatment effect: (i) homogeneity, (ii) partial heterogeneity, and (iii) full heterogeneity of the population in the treatment response<sup>37</sup>.

Our findings allow us to draw reliable conclusions only for paid-employed workers, because most of the results on the self-employed are not statistically significant.

In paid-employment, our analysis of informal employment effects under the assumption of strict homogeneity showed that workers display a significant informal employment income penalty and that, among them, women do not exhibit an obvious income disadvantage relative to men from working informally. These results advocate for the exclusion hypothesis for both paid-employed men and women but they do not evidence that, in paid-employment, the exclusion hypothesis is clearly more prevalent among women.

---

<sup>37</sup> A methodological weakness that deserves to be mentioned is that, to overcome potential biases resulting from non-randomness in treatment, we controlled for some observed pre-treatment covariates and invoked a conditional independence assumption called ignorability, which is unverifiable and cannot be tested with data. However, our approach does not have the major drawbacks of the IV approach that has been proposed recently to estimate heterogeneous treatment effects, and, in addition, is easier to implement.

By contrast, relaxing the strict homogeneity assumption and allowing for partial heterogeneity, we found that, in paid-employment, women face a significantly higher informal employment income penalty and are more affected by the exclusion hypothesis than men.

Allowing for full heterogeneity yields similar findings.

In addition, our data indicate that paid-employed workers clearly differ in how they respond to a particular treatment, which is to work informally, thus supporting the full heterogeneity assumption over the strict homogeneity and partial heterogeneity hypotheses.

Overall, our investigation led us to believe that the decision to work informally is probably not the mere result of a rational choice in which people weigh the expected economic returns against the costs based solely on economic factors. This decision is very likely to be governed also by cultural and social norms, circumstances and constraints.

The fact that the exclusion hypothesis is more prevalent among women may explain why they face a significantly higher informal employment income penalty than men.



## Appendix A. Methodological and estimation issues

### Heterogeneity bias

The main problem of causal inference is that, for a given individual, we observe either the potential outcome (log hourly income from main job) if treated (work informally) or the potential outcome (log hourly income from main job) if untreated (work formally), but not both. One solution to identify individual-level treatment effects would be to rely on a strong homogeneity assumption according to which there is only a difference between the treated state and the untreated state, but there is no variability across individuals within the same state, so that all treated individuals are identical and all untreated individuals are identical. However, we know that such a restrictive assumption is very unlikely to hold true in practice. A better solution would be, for example, to compare the average difference between a randomly selected set of individuals that were treated to another randomly selected set of individuals that were untreated. The comparison of these two randomly selected groups yields the Average Treatment Effect (ATE):

$$ATE = E(Y^1 - Y^0) = E(Y^1|D=1) - E(Y^0|D=0) = \delta \quad (2)$$

Where  $Y^1$  is the outcome if treated, and  $Y^0$  the outcome if untreated. Analogously, we define the Treatment Effect of the Treated (TT – the average difference by treatment status among those individuals who are actually treated) and the Treatment Effect of the Untreated (TUT – the average difference by treatment status among those individuals who are not treated):

$$\begin{cases} TT = E(Y^1 - Y^0|D=1) = E(Y^1|D=1) - E(Y^0|D=1) = \delta^1 \\ TUT = E(Y^1 - Y^0|D=0) = E(Y^1|D=0) - E(Y^0|D=0) = \delta^0 \end{cases} \quad (3)$$

Where D is the treatment status, which equals one if treated and zero if untreated.

If treated and untreated are randomly selected from the whole population, then we have:

$$\begin{cases} E(Y^1) = E(Y^1|D=1) = E(Y^1|D=0) \\ E(Y^0) = E(Y^0|D=0) = E(Y^0|D=1) \end{cases} \Rightarrow ATE = TT = TUT \Leftrightarrow \delta = \delta^1 = \delta^0 \quad (4)$$

In reality, due to population heterogeneity, there is no guarantee that the group that actually receives the treatment is comparable, in observed and particularly in unobserved characteristics, to the group that does not receive the treatment, i.e. that conditions expressed in Equation 4 are satisfied. We are in fact concerned with two types of bias:

$$\begin{cases} ATE = E(Y^1 - Y^0) \\ ATE = E(Y_{D=1}^1)p + E(Y_{D=0}^1)q - E(Y_{D=1}^0)p - E(Y_{D=0}^0)q \\ ATE = E(Y_{D=1}^1) - E(Y_{D=1}^0) + E(Y_{D=0}^1)q - E(Y_{D=1}^0) + E(Y_{D=1}^0)q - E(Y_{D=0}^0)q \\ ATE = E(Y_{D=1}^1) - E(Y_{D=0}^0) - [E(Y_{D=1}^0) - E(Y_{D=0}^0)] - [E(Y_{D=1}^1) - E(Y_{D=1}^0)] - [E(Y_{D=0}^1) - E(Y_{D=0}^0)]q \\ ATE = \delta - [E(Y_{D=1}^0) - E(Y_{D=0}^0)] - (\delta^1 - \delta^0)q \end{cases} \quad (5)$$

Where  $E(Y_{D=b}^a) = E(Y^a|D=b)$ , with  $a, b = 0, 1$ .  $p$  is the proportion treated ( $D = 1$ ), and  $q$  the proportion untreated ( $D = 0$ ).

$[E(Y_{D=1}^0) - E(Y_{D=0}^0)]$  is the pre-treatment heterogeneity bias, or type I selection bias, and corresponds to the average difference in outcomes between treated and untreated if neither group receives the treatment.

$(\delta^1 - \delta^0)_q$  or  $(TT - TUT)_q$  is the treatment-effect heterogeneity bias, or type II selection bias, and corresponds to the difference in the average treatment effect between treated and untreated<sup>38</sup>.

When there is type II selection bias,  $ATE \neq TT \neq TUT \Leftrightarrow \delta \neq \delta^1 \neq \delta^0$ . Unlike type I selection bias, treatment-effect heterogeneity bias cannot be controlled for by covariates or fixed-effects. Both sources of bias average to zero under randomized assignment. Random assignment ensures that individuals receive either the treatment or control condition by chance only, which means that, in theory, outcomes of treated individuals and outcomes of untreated individuals are independent from treatment status:

$$(Y^1, Y^0) \perp D \quad \text{and} \quad ATE = TT = TUT = E(Y^1_{D=1}) - E(Y^0_{D=0}) \quad (6)$$

This independence condition is very unlikely to hold true because individuals self-select into treatment or control groups for a number of reasons, some of which may be unknowable from data.

There are two ways to proceed when assignment to treatment is not random: find one or more instrumental variables (IV) that affect assignment to treatment exogenously but affect the outcome only indirectly through treatment (exclusion restriction), or control for differences between the treatment group and the control group with observed covariates. Some recent papers have proposed a method of estimating heterogeneous treatment effects using the IV approach (see Heckman *et al.* 2006; Heckman and Vylatcil 2005, 2001)<sup>39</sup>. As raised by Xie *et al.* (2011), the IV

---

<sup>38</sup> From Equation 1, type I selection bias is observed if  $\text{corr}(\epsilon, D) \neq 0$ , and type II selection bias is observed if  $\text{corr}(\delta, D) \neq 0$ .

<sup>39</sup> Zhou and Xie (2011) find that, as long as the ignorability of type I selection bias is satisfied, propensity score-based (PS-based) methods (including matching, stratification and weighting) can still identify TT, even in the presence of a heterogeneous treatment effect bias. Furthermore, when type I selection bias cannot be ignored, the bias for TT is in the same direction as the type I selection bias. By comparison,

approach has three major weaknesses: i) it is very hard to find a meaningful IV that satisfies the exclusion restriction assumption, ii) a weak IV may give rise to imprecise IV estimates and estimates that are biased in finite samples, and iii) if treatment effects are heterogeneous, the estimates using the IV formula should be only interpreted as local average treatment effects (LATE), average effects that pertain only to the units whose treatment assignment statuses are affected by the instrument. In this study, we choose to control for differences between the treatment group and the control group with observed covariates.

To overcome potential biases resulting from non-randomness in treatment assignment, it is necessary to control for observed pre-treatment covariates. Because it is impossible to control for all relevant variables that may affect the outcome, we invoke a conditional independence assumption called ignorability, unconfoundedness or selection on observables:

$$(Y^1, Y^0) \perp D | X \quad (7)$$

This assumption states that, conditional on observed confounding covariates denoted by vector  $X$ , the potential outcomes of treated and untreated are independent from treatment status. Confounding covariates refer to relevant pre-treatment variables that confound the observed relationship between treatment and outcome by meeting the condition of affecting both the treatment assignment and the outcome. Since some of the potential outcomes are unobserved, the ignorability assumption is unverifiable and cannot be tested with data. Results for causal inference under this assumption thus should always be interpreted cautiously.

---

marginal treatment effect-based (MTE-based) methods (IV approach) are robust to different types of violation of the ignorability assumption. However, they require strong instrumental variables that satisfy the exclusion restriction assumption to achieve statistical efficiency. If exclusion restriction is violated, MTE-based methods can be subject to severe over- or under-estimation of treatment effects.

Conditioning on observed confounding covariates can be difficult in applied research due to the curse of dimensionality<sup>40</sup>. Rosenbaum and Rubin (1984, 1983) propose an ingenious solution to this problem. They show that, when the ignorability assumption holds true, it is sufficient to condition on the propensity score (the conditional probability of treatment) as a function of the observed confounding covariates:

$$\left( Y^1, Y^0 \right) \perp\!\!\!\perp D \mid p(D = 1 \mid X) \quad (8)$$

Where  $p(D = 1 \mid X)$  is the propensity score, the probability of treatment given all the relevant information in covariates  $X$ . For simplicity, we also denote the propensity score by  $p(X)$ . Equation 8 tells us that only through the propensity score may covariates confound the observed relationship between treatment and outcome. In other words, according to Equation 8, under the ignorability assumption there is no selection bias conditional on the propensity score (i.e., any imbalance between the treated group and the untreated group can be adequately captured and characterized by the propensity score):

$$\begin{cases} E(Y^1 \mid p(X)) = E(Y_1^1 \mid p(X)) = E(Y_0^1 \mid p(X)) \\ E(Y^0 \mid p(X)) = E(Y_0^0 \mid p(X)) = E(Y_1^0 \mid p(X)) \end{cases} \Rightarrow ATE = TT = TUT \Leftrightarrow \delta = \delta^1 = \delta^0 \quad (9)$$

---

<sup>40</sup> Under assumption of ignorability, matching on observed measured confounders can be used to overcome the confounding bias problem. However, the problem with matching is that sample sizes are often not big enough to achieve matching on all observed covariates. As a result, researchers are often constrained to match on a relatively small set of variables. This problem of matching applies if one has several discrete covariates but becomes particularly severe if the covariates are of a continuous nature and is known as the curse of dimensionality.

### Estimating propensity score and constructing balanced propensity score strata

Formally, let  $d_i^*$  be a latent variable which is the underlying propensity of having an informal main job, and  $d_i$  the observed outcome:

$$d_i^* = \eta + \gamma W_i + u_i \quad \text{with} \quad \begin{cases} d_i = 1 & \text{if } d_i^* > 0 \\ d_i = 0 & \text{otherwise} \end{cases} \quad (10)$$

Since  $d_i^*$  is unobserved, we use an indicator variable  $d_i$  which is equal to one if the individual  $i$  has an informal main job, and zero if the individual  $i$  has a formal main job. In Equation 10,  $W$  is a vector of explanatory variables,  $u$  is the error term with an expected value of zero, and  $\eta$  is the constant term to be estimated together with the regression coefficients  $\gamma$ . In all regressions, we use the Huber/White sandwich estimator of the variance to correct for heteroskedasticity.

The following Probit model is applied for estimating the propensity score:

$$\Pr(d_i^* > 0 | W_i) = \Pr(d_i = 1 | W_i) = \Phi(\gamma W_i) \quad (11)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution. The vector of income determinants  $X$  in Equation 1 is a subset of the vector of explanatory variables  $W$  in Equations 10 and 11. Vector  $W$  consists of the same regressors than vector  $X$ , plus some variables that are believed to affect informal employment participation decision.

Once the propensity score is estimated, we construct balanced propensity score strata. In addition to the ignorability assumption – defined from Equation 8 as  $(Y^1, Y^0) \perp D | p(W)$  –, which is

empirically unverifiable, we need to satisfy another property, which is the balancing hypothesis (balancing of pre-treatment variables given the propensity score):

$$D \perp W | p(W) \quad (12)$$

If this hypothesis is satisfied, observations with the same propensity score must have the same distribution of observable and unobservable characteristics independently of treatment status. Said differently, for a given propensity score, exposure to treatment is random and therefore treated and untreated units should be on average observationally identical.

We restrict the analysis of the balancing property to all treated plus those untreated in the region of common support, and we set the significance level of the balancing property tests at 0.01. The number of blocks of equal range used at the beginning of the tests of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated in each stratum.

We make sure that balancing is satisfied (i.e., within each propensity score stratum, the average propensity score and the means of each covariate do not significantly differ between treated and untreated) in all cases (men and women in paid- and self-employment). On this point, it is worth mentioning that, in order to achieve the balancing property, we had to use a less parsimonious specification of the propensity score Probit regression models, include higher order terms (age squared, squared number of children aged 0-6 and aged 7-14 living in the household), and replace some variables (potential work experience is proxied by dummies for two age groups in male regressions instead of by age and its square as in female regressions). Accordingly, the choice of covariates in vector  $W$  and in subvector  $X$  is not solely determined by their relevance to

characterize informal employment participation decision, but also, to some extent, by the need to satisfy the balancing property in order to obtain balanced propensity score strata.

### **Matching algorithms**

There is no clear consensus as to which matching estimator performs best in each application. We thus implement various propensity score matching methods for the sake of illustration and comparison, and as a way to assess the robustness of the estimates. The matching algorithms we use are the followings: single nearest neighbour matching with replacement but without caliper, radius matching within a radius size of 0.1, and Epanechnikov kernel matching with a bandwidth of 0.06.

With single nearest neighbour matching, treated (untreated) units are matched to control units that are closest to the treated (untreated) units in their estimated propensity scores. ‘With replacement’ means that a control unit can be a best match for more than one treated (untreated) unit. Once each treated (untreated) unit is matched with a control unit, we compute the difference between the outcome of the treated (matched control) units and the outcome of the matched control (untreated) units. We then obtain the ATT (ATU) of interest by averaging these differences. While the single nearest neighbour method allows all treated (untreated) units to find a match, it is clear that some of these matches are fairly poor because for some treated (untreated) units the nearest neighbour may have a very different propensity score and nevertheless it would contribute to the estimation of the treatment effect independently of this difference.

The radius matching and kernel matching methods offer a solution to this problem. With radius matching each treated (untreated) unit is matched only with the control units whose propensity score falls in a predefined neighbourhood of the propensity score of the treated (untreated) unit. If the dimension of the neighbourhood (i.e., the radius) is set to be very small it is possible that some treated (untreated) units are not matched because the neighbourhood does not contain control units. On the other hand, the smaller the size of the neighbourhood the better is the quality



of the matches. With kernel matching all treated (untreated) are matched with a weighted average of all controls with weights that are inversely proportional to the distance between the propensity score of treated (untreated) and controls (i.e., bandwidth).

## Appendix B. Descriptive statistics

**Table 4.4: List and definition of covariates**

Covariate	Definition
<b>agegroup1 (ref.)</b>	Youths : Individuals aged 15-24
<b>agegroup2</b>	Prime-age adults: Individuals aged 25-34
<b>agegroup3</b>	Middle-age adults and seniors: Individuals aged 35+
<b>age</b>	Age (complete years)
<b>age2</b>	Age squared
<b>edu1 (ref.)</b>	Never attended school
<b>edu2</b>	<u>Primary education incomplete</u> : Attending or completed preschool (2 years of pre-primary education) or standard 1-6 (6 years of primary education), or attending standard 7-8 (7 <sup>th</sup> and last year of primary education)
<b>edu3</b>	<u>Primary education completed</u> : Completed standard 7-8 (7 years of primary education)
<b>edu4</b>	<u>Secondary education and above</u> : Attending or completed form 1-4 (4 years of junior secondary - ordinary level) or form 5-6 (2 years of senior secondary - advanced level) or tertiary non university or tertiary university
<b>training</b>	Training of at least one month duration (on the job, informal apprenticeship, vocational certificate, college certificate, diploma or advanced diploma, university degree or other)
<b>married</b>	Married
<b>migrant</b>	She/he is living for less than five years in her/his present town/district
<b>nb_child06</b>	Number of children aged 0-6 living in the household
<b>nb_child06sq</b>	Squared number of children aged 0-6 living in the household
<b>nb_child714</b>	Number of children aged 7-14 living in the household
<b>nb_child714sq</b>	Squared number of children aged 7-14 living in the household
<b>urban</b>	Living in an urban area
<b>geog_zone1 (ref.)</b>	Dar es Salaam
<b>geog_zone2</b>	Central-Eastern (Dodoma, Singida, Morogoro, Pwani, Lindi)
<b>geog_zone3</b>	Lake (Shinyanga, Kagera, Mwanza, Mara)
<b>geog_zone4</b>	Northern (Arusha, Kilimanjaro, Tanga, Manyara)
<b>geog_zone5</b>	Southern-Southern Highlands-Western (Mtwara, Ruvuma, Iringa, Mbeya, Tabora, Rukwa, Kigoma)
<b>q1 (ref.)</b>	Survey interview took place during the first quarter of 2006
<b>q2</b>	Survey interview took place during the second quarter of 2006
<b>q3</b>	Survey interview took place during the third quarter of 2006
<b>q4</b>	Survey interview took place during the fourth quarter of 2006
<b>car</b>	The household owns a car
<b>telecom</b>	The household owns a cellular or a landline phone
<b>electricity</b>	Electricity is the main source of energy of the household for either cooking, lighting or heating/cooling
<b>public_transport</b>	Either bus, taxi or train are within a 30-minute (equivalent to 2 km) walk from the household
<b>water</b>	Household access to an improved drinking-water source: protected private/public well or spring, or bottled water
<b>toilet</b>	Household type of toilet: improved pit latrine

**Table 4.5: Mean covariates values by employment status and gender**

	Paid Employment				Self Employment			
	Men (N = 2515)		Women (N = 1098)		Men (N = 2403)		Women (N = 1814)	
	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
<b>agegroup1 (ref.)</b>	0.04	0.25	–	–	0.14	0.16	–	–
<b>agegroup2</b>	0.26	0.38	–	–	0.34	0.40	–	–
<b>agegroup3</b>	0.70	0.37	–	–	0.52	0.44	–	–
<b>age</b>	–	–	38.38	27.90	–	–	34.62	34.45
<b>age2</b>	–	–	1569.79	892.51	–	–	1305.57	1329.37
<b>edu1 (ref.)</b>	0.01	0.08	0.01	0.09	0.05	0.08	0.05	0.15
<b>edu2</b>	0.04	0.10	0.02	0.08	0.07	0.11	0.08	0.10
<b>edu3</b>	0.39	0.66	0.33	0.67	0.68	0.69	0.69	0.66
<b>edu4</b>	0.55	0.15	0.65	0.16	0.21	0.12	0.18	0.08
<b>training</b>	0.74	0.36	0.80	0.22	0.38	0.29	0.27	0.15
<b>married</b>	0.83	0.57	0.55	0.22	0.67	0.68	0.61	0.56
<b>migrant</b>	0.17	0.18	0.18	0.32	0.11	0.13	0.11	0.13
<b>nb_child06</b>	0.80	0.86	0.60	0.63	0.77	0.93	0.90	0.95
<b>nb_child06sq</b>	1.46	1.73	0.97	1.06	1.36	1.88	2.13	1.88
<b>nb_child714</b>	0.92	0.74	0.82	0.80	0.76	0.82	0.97	1.00
<b>nb_child714sq</b>	2.07	1.68	1.75	1.66	1.69	1.83	2.56	2.33
<b>urban</b>	0.79	0.70	0.83	0.81	0.67	0.72	0.70	0.79
<b>geog_zone1 (ref.)</b>	0.35	0.32	0.30	0.34	0.30	0.28	0.17	0.27
<b>geog_zone2</b>	0.14	0.13	0.13	0.09	0.10	0.15	0.10	0.13
<b>geog_zone3</b>	0.16	0.18	0.15	0.17	0.22	0.21	0.23	0.20
<b>geog_zone4</b>	0.18	0.17	0.22	0.21	0.26	0.15	0.45	0.17
<b>geog_zone5</b>	0.16	0.19	0.20	0.19	0.12	0.22	0.06	0.23
<b>q1 (ref.)</b>	0.26	0.31	0.26	0.33	0.30	0.32	0.34	0.32
<b>q2</b>	0.28	0.27	0.30	0.25	0.32	0.23	0.36	0.24
<b>q3</b>	0.25	0.25	0.23	0.22	0.21	0.27	0.12	0.25
<b>q4</b>	0.21	0.17	0.21	0.20	0.17	0.19	0.18	0.19

**Table 4.5: Continued**

	Paid Employment				Self Employment			
	Men (N = 2515)		Women (N = 1098)		Men (N = 2403)		Women (N = 1814)	
	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
<b>car</b>	0.10	0.04	0.16	0.11	0.13	0.03	0.09	0.03
<b>telecom</b>	0.77	0.47	0.85	0.60	0.56	0.43	0.66	0.40
<b>electricity</b>	0.57	0.31	0.62	0.51	0.38	0.27	0.49	0.30
<b>public_transport</b>	0.89	0.83	0.91	0.88	0.80	0.87	0.92	0.90
<b>water</b>	0.16	0.19	0.16	0.18	0.18	0.23	0.20	0.22
<b>toilet</b>	0.13	0.09	0.16	0.15	0.12	0.07	0.14	0.10
<b>Sample Size</b>	<b>976</b>	<b>1539</b>	<b>420</b>	<b>678</b>	<b>229</b>	<b>2174</b>	<b>103</b>	<b>1711</b>

*Source:* Tanzania ILFS 2006. *Note:* Individuals aged 15 and older. Sampling weights are not accounted for.

## Appendix C. Informal employment effects under the assumption of homogeneity

Table 4.6: OLS regressions of log hourly income from main job by employment status and gender

	Paid Employment		Self Employment	
	Men	Women	Men	Women
<b>informal employment</b>	-0.6295*** (-15.47)	-0.6515*** (-10.44)	-0.2386** (-2.33)	-0.0690 (-0.44)
<b>agegroup2</b>	0.2583*** (4.74)	— —	0.0682 (0.76)	— —
<b>agegroup3</b>	0.3935*** (6.47)	— —	0.0824 (0.87)	— —
<b>age</b>	— —	0.0655*** (4.81)	— —	0.0518*** (4.09)
<b>age2</b>	— —	-0.0005*** (-2.74)	— —	-0.0007*** (-4.39)
<b>edu2</b>	0.1320 (1.18)	0.4631*** (2.78)	-0.1862 (-1.29)	0.0926 (0.75)
<b>edu3</b>	0.3070*** (3.45)	0.4212*** (3.60)	0.0361 (0.28)	-0.0007 (-0.01)
<b>edu4</b>	0.8198*** (8.46)	0.9856*** (7.57)	0.0919 (0.59)	0.3919*** (2.78)
<b>training</b>	0.3312*** (8.47)	0.5645*** (8.20)	0.0111 (0.17)	0.0685 (0.78)
<b>married</b>	0.1524*** (3.42)	0.1835*** (3.09)	0.1305* (1.95)	-0.1185* (-1.80)
<b>migrant</b>	-0.1163*** (-2.70)	-0.1128** (-2.12)	-0.0073 (-0.09)	0.0576 (0.66)
<b>urban</b>	0.1297*** (3.02)	0.0334 (0.49)	-0.1152 (-1.62)	-0.1015 (-1.29)
<b>geog_zone2</b>	-0.1539** (-2.55)	-0.1187 (-1.37)	-0.2249** (-2.21)	-0.2574** (-2.40)
<b>geog_zone3</b>	-0.0975* (-1.80)	-0.2664*** (-3.26)	-0.2240** (-2.54)	-0.1812* (-1.86)
<b>geog_zone4</b>	-0.0641 (-1.26)	-0.2362*** (-3.85)	-0.1685* (-1.81)	-0.3026*** (-3.02)
<b>geog_zone5</b>	-0.1598*** (-2.86)	-0.1416** (-2.01)	-0.1591* (-1.81)	-0.1564* (-1.65)
<b>q2</b>	0.1505*** (3.12)	0.0094 (0.14)	0.0501 (0.63)	0.0350 (0.43)
<b>q3</b>	0.0276 (0.57)	0.0178 (0.25)	0.1010 (1.30)	0.2594*** (3.03)
<b>q4</b>	0.1001** (2.06)	0.0314 (0.46)	0.2576*** (3.20)	0.2509*** (2.69)
<b>constant</b>	4.9870*** (42.20)	3.4412*** (12.93)	6.5988*** (37.04)	5.4353*** (17.56)
<b>Sample Size</b>	<b>2515</b>	<b>1098</b>	<b>2403</b>	<b>1814</b>
<b>Pseudo R-squared</b>	<b>0.4107</b>	<b>0.6489</b>	<b>0.0180</b>	<b>0.0442</b>
<b>Mean of depvar</b>	<b>5.625</b>	<b>5.333</b>	<b>6.432</b>	<b>6.126</b>

Source: Tanzania ILFS 2006. Note: Individuals aged 15 and older. Regressions account for sampling weights. Heteroskedasticity-robust T statistics in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

## Appendix D. Informal employment effects under the assumption of partial heterogeneity

**Table 4.7: Propensity score Probit regression models predicting informal employment by employment status and gender**

	Paid Employment		Self Employment	
	Men	Women	Men	Women
<b>agegroup2</b>	-0.4461*** (-3.79)	—	0.0133 (0.10)	—
<b>agegroup3</b>	-0.9549*** (-7.89)	—	-0.1259 (-0.96)	—
<b>age</b>	—	-0.2036*** (-6.72)	—	-0.0260 (-1.08)
<b>age2</b>	—	0.0021*** (5.30)	—	0.0003 (0.98)
<b>edu2</b>	-0.2849 (-1.40)	-0.4588 (-1.14)	-0.0368 (-0.18)	-0.4255 (-1.54)
<b>edu3</b>	-0.3988** (-2.27)	-0.7395** (-2.29)	-0.1992 (-1.17)	-0.4034* (-1.81)
<b>edu4</b>	-1.1602*** (-6.32)	-1.4060*** (-4.09)	-0.3646* (-1.88)	-0.6114** (-2.26)
<b>training</b>	-0.4427*** (-6.73)	-0.8903*** (-7.41)	-0.1756** (-2.08)	-0.2445* (-1.85)
<b>married</b>	-0.2715*** (-3.30)	-0.0735 (-0.65)	0.1402 (1.39)	-0.0351 (-0.31)
<b>migrant</b>	-0.0312 (-0.38)	0.1962 (1.60)	0.1099 (0.89)	0.1910 (1.13)
<b>nb_child06</b>	0.0154 (0.19)	0.0932 (0.59)	-0.1172 (-1.13)	0.2675*** (2.26)
<b>nb_child06sq</b>	0.0077 (0.27)	-0.0562 (-0.86)	0.0669* (1.87)	-0.0896*** (-2.64)
<b>nb_child714</b>	-0.0682 (-1.07)	0.2451* (1.86)	0.0502 (0.58)	0.1707* (1.92)
<b>nb_child714sq</b>	0.0179 (1.06)	-0.0807* (-1.90)	0.0002 (0.01)	-0.0397* (-1.86)
<b>urban</b>	0.0136 (0.15)	-0.2722* (-1.71)	0.2423** (2.29)	0.2831* (1.96)
<b>geog_zone2</b>	-0.4350*** (-4.32)	-0.2802 (-1.47)	0.3768*** (2.69)	-0.1799 (-0.85)
<b>geog_zone3</b>	-0.2330** (-2.38)	-0.1641 (-0.90)	-0.0197 (-0.17)	-0.3322* (-1.75)
<b>geog_zone4</b>	-0.2408** (-2.57)	-0.4337*** (-2.82)	-0.1987* (-1.77)	-0.8128*** (-4.86)
<b>geog_zone5</b>	-0.3219*** (-3.29)	-0.5680*** (-3.62)	0.4241*** (3.20)	0.2344 (1.06)

**Table 4.7: Continued**

	Paid Employment		Self Employment	
	Men	Women	Men	Women
<b>q2</b>	-0.0932 (-1.15)	-0.0931 (-0.70)	-0.1528 (-1.56)	-0.1432 (-1.08)
<b>q3</b>	-0.1047 (-1.26)	0.1509 (1.06)	0.1117 (1.08)	0.4344** (2.54)
<b>q4</b>	-0.1608* (-1.84)	-0.0407 (-0.27)	0.0285 (0.25)	0.1214 (0.77)
<b>car</b>	-0.0257 (-0.21)	-0.2917* (-1.71)	-0.6397*** (-4.09)	-0.2914 (-1.21)
<b>telecom</b>	-0.3680*** (-4.89)	-0.5779*** (-4.19)	-0.1365 (-1.50)	-0.4030*** (-3.19)
<b>electricity</b>	-0.3202*** (-4.30)	0.2454* (1.85)	-0.0891 (-0.88)	-0.1427 (-1.14)
<b>public_transport</b>	-0.1906* (-1.91)	0.0930 (0.49)	0.4218*** (3.56)	-0.0446 (-0.21)
<b>water</b>	-0.0462 (-0.58)	-0.0215 (-0.15)	0.0561 (0.58)	-0.2010 (-1.47)
<b>toilet</b>	0.2387** (2.44)	0.0279 (0.19)	-0.1601 (-1.20)	-0.1168 (-0.66)
<b>constant</b>	2.7881*** (12.72)	6.5441*** (10.57)	0.9998*** (4.46)	2.7628*** (5.36)
<b>Sample Size</b>	<b>2515</b>	<b>1098</b>	<b>2403</b>	<b>1814</b>
<b>Log pseudolikelihood</b>	<b>-1218.32</b>	<b>-411.18</b>	<b>-709.05</b>	<b>-332.03</b>
<b>Wald Chi2 (26)</b>	<b>653.68</b>	<b>398.56</b>	<b>108.72</b>	<b>90.14</b>
<b>Prob &gt; Chi2</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
<b>Pseudo R-squared</b>	<b>0.2691</b>	<b>0.4340</b>	<b>0.0838</b>	<b>0.1638</b>

Source: Tanzania ILFS 2006. Note: Individuals aged 15 and older. Regressions account for sampling weights. Heteroskedasticity-robust Z statistics in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%.

**Table 4.8: Mean covariates values by propensity score strata and employment status, men in paid employment**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5		Stratum 6		Stratum 7		Stratum 8		Stratum 9	
	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE
<b>agegroup1 (ref.)</b>	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.07	0.05	0.03	0.05	0.12	0.09	0.10	0.21	0.47	0.65
<b>agegroup2</b>	0.02	0.05	0.19	0.15	0.36	0.29	0.34	0.26	0.24	0.25	0.27	0.38	0.35	0.47	0.62	0.51	0.46	0.32
<b>agegroup3</b>	0.98	0.95	0.81	0.85	0.64	0.71	0.65	0.71	0.69	0.70	0.70	0.57	0.53	0.44	0.28	0.27	0.06	0.03
<b>edu1 (ref.)</b>	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.04	0.00	0.02	0.03	0.07	0.10	0.16	0.21
<b>edu2</b>	0.00	0.00	0.01	0.00	0.00	0.03	0.05	0.04	0.06	0.06	0.11	0.10	0.10	0.14	0.11	0.11	0.20	0.17
<b>edu3</b>	0.00	0.00	0.13	0.08	0.36	0.37	0.58	0.60	0.72	0.66	0.77	0.75	0.78	0.75	0.78	0.76	0.64	0.62
<b>edu4</b>	1.00	1.00	0.87	0.92	0.64	0.60	0.37	0.35	0.22	0.27	0.08	0.14	0.11	0.09	0.04	0.03	0.00	0.00
<b>training</b>	0.99	1.00	0.93	0.84	0.86	0.82	0.84	0.77	0.69	0.63	0.50	0.55	0.21	0.35	0.17	0.21	0.10	0.07
<b>married</b>	0.97	1.00	0.87	0.89	0.78	0.88	0.78	0.84	0.82	0.78	0.84	0.74	0.74	0.73	0.65	0.54	0.30	0.22
<b>migrant</b>	0.15	0.17	0.18	0.20	0.23	0.15	0.18	0.16	0.16	0.11	0.09	0.13	0.09	0.17	0.24	0.16	0.15	0.22
<b>nb_child06</b>	0.68	0.70	0.96	0.63	0.82	0.78	0.85	0.89	0.76	0.89	1.03	0.83	0.94	0.92	0.88	0.87	0.55	0.94
<b>nb_child06sq</b>	1.16	1.08	1.67	0.95	1.69	1.20	1.55	1.67	1.37	1.67	2.18	1.84	1.64	1.79	1.73	1.71	1.12	2.21
<b>nb_child714</b>	1.05	1.27	1.13	0.95	0.98	0.84	0.86	0.84	0.99	0.94	0.90	0.90	0.79	0.58	0.49	0.58	0.39	0.77
<b>nb_child714sq</b>	2.27	3.38	2.78	1.73	2.24	1.79	2.22	1.86	2.02	1.98	1.92	1.98	1.65	1.15	0.96	1.16	0.47	2.11
<b>urban</b>	0.80	0.72	0.75	0.74	0.76	0.76	0.79	0.74	0.64	0.70	0.66	0.69	0.60	0.61	0.61	0.60	0.58	0.52
<b>geog_zone1 (ref.)</b>	0.43	0.33	0.28	0.43	0.39	0.53	0.40	0.34	0.38	0.42	0.32	0.39	0.30	0.28	0.40	0.29	0.34	0.27
<b>geog_zone2</b>	0.11	0.15	0.16	0.07	0.12	0.13	0.15	0.13	0.20	0.11	0.18	0.14	0.15	0.16	0.06	0.15	0.04	0.12
<b>geog_zone3</b>	0.15	0.17	0.22	0.21	0.12	0.05	0.21	0.17	0.18	0.07	0.13	0.16	0.17	0.15	0.12	0.19	0.07	0.25
<b>geog_zone4</b>	0.12	0.13	0.18	0.14	0.17	0.15	0.16	0.22	0.14	0.18	0.20	0.16	0.22	0.18	0.20	0.18	0.16	0.15
<b>geog_zone5</b>	0.19	0.22	0.15	0.16	0.19	0.14	0.08	0.13	0.10	0.22	0.17	0.14	0.16	0.22	0.22	0.19	0.39	0.21
<b>q1 (ref.)</b>	0.23	0.25	0.18	0.26	0.18	0.20	0.26	0.31	0.33	0.25	0.29	0.32	0.30	0.27	0.33	0.31	0.30	0.33
<b>q2</b>	0.25	0.33	0.31	0.30	0.33	0.30	0.31	0.25	0.22	0.28	0.30	0.31	0.22	0.22	0.27	0.24	0.26	0.28
<b>q3</b>	0.25	0.27	0.25	0.22	0.26	0.26	0.18	0.28	0.28	0.22	0.21	0.20	0.23	0.29	0.28	0.27	0.34	0.23
<b>q4</b>	0.26	0.16	0.25	0.22	0.24	0.24	0.24	0.16	0.17	0.25	0.20	0.17	0.25	0.21	0.13	0.17	0.10	0.16



**Table 4.8: Continued**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5		Stratum 6		Stratum 7		Stratum 8		Stratum 9	
	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE
<b>car</b>	0.28	0.12	0.09	0.05	0.04	0.08	0.03	0.07	0.05	0.05	0.00	0.02	0.03	0.03	0.00	0.04	0.03	0.04
<b>telecom</b>	0.99	1.00	0.87	0.84	0.87	0.81	0.82	0.85	0.65	0.63	0.53	0.60	0.50	0.45	0.29	0.36	0.28	0.21
<b>electricity</b>	0.81	0.67	0.56	0.53	0.61	0.66	0.56	0.55	0.42	0.42	0.35	0.39	0.26	0.25	0.22	0.21	0.04	0.13
<b>public_transport</b>	0.94	0.90	0.85	0.88	0.81	0.89	0.89	0.92	0.87	0.86	0.83	0.85	0.84	0.81	0.79	0.81	0.82	0.69
<b>water</b>	0.14	0.06	0.13	0.13	0.14	0.19	0.17	0.14	0.26	0.23	0.21	0.18	0.15	0.23	0.14	0.21	0.27	0.18
<b>toilet</b>	0.08	0.10	0.22	0.25	0.16	0.15	0.16	0.14	0.10	0.09	0.12	0.12	0.15	0.08	0.01	0.06	0.05	0.06
<b>Sample Size</b>	<b>250</b>	<b>32</b>	<b>138</b>	<b>50</b>	<b>124</b>	<b>62</b>	<b>120</b>	<b>83</b>	<b>102</b>	<b>135</b>	<b>80</b>	<b>167</b>	<b>59</b>	<b>256</b>	<b>58</b>	<b>341</b>	<b>36</b>	<b>413</b>

*Source:* Tanzania ILFS 2006. *Note:* Paid employed men aged 15 and older. Sampling weights are accounted for. The balancing property is satisfied. The balancing property analysis has been restricted to all treated plus those controls in the region of common support, and the significance level of the balancing property tests has been set at 0.01. The number of blocks of equal score range used at the beginning of the test of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated. FE means formal employment and IE refers to informal employment.

**Table 4.9: Mean covariates values by propensity score strata and employment status, women in paid employment**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5	
	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE
<b>age</b>	42.57	45.14	38.50	35.83	35.18	33.33	32.33	35.15	29.07	24.69
<b>age2</b>	1872.86	2086.36	1576.08	1359.84	1348.70	1171.90	1101.16	1347.64	931.00	701.58
<b>edu1 (ref.)</b>	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.10	0.14
<b>edu2</b>	0.00	0.00	0.01	0.02	0.03	0.04	0.02	0.08	0.14	0.10
<b>edu3</b>	0.13	0.09	0.41	0.25	0.43	0.44	0.69	0.74	0.60	0.73
<b>edu4</b>	0.87	0.91	0.58	0.73	0.54	0.52	0.29	0.16	0.16	0.04
<b>training</b>	0.99	1.00	0.90	0.90	0.70	0.70	0.31	0.28	0.08	0.06
<b>married</b>	0.65	0.83	0.60	0.50	0.37	0.43	0.39	0.23	0.23	0.16
<b>migrant</b>	0.15	0.12	0.20	0.21	0.28	0.19	0.10	0.20	0.23	0.37
<b>nb_child06</b>	0.56	0.53	0.61	0.81	0.52	0.60	0.54	0.38	0.62	0.66
<b>nb_child06sq</b>	0.97	0.89	0.94	1.26	0.81	0.82	0.72	0.63	1.03	1.13
<b>nb_child714</b>	0.97	1.07	0.68	0.90	0.77	0.74	0.74	0.53	0.95	0.85
<b>nb_child714sq</b>	2.22	2.60	1.32	1.67	1.64	1.42	1.24	0.83	2.10	1.79
<b>urban</b>	0.75	0.72	0.82	0.76	0.67	0.82	0.73	0.66	0.76	0.75
<b>geog_zone1 (ref.)</b>	0.35	0.19	0.32	0.41	0.26	0.39	0.31	0.21	0.36	0.39
<b>geog_zone2</b>	0.10	0.27	0.15	0.05	0.16	0.15	0.13	0.11	0.12	0.08
<b>geog_zone3</b>	0.16	0.25	0.13	0.28	0.23	0.17	0.14	0.06	0.18	0.17
<b>geog_zone4</b>	0.21	0.19	0.21	0.17	0.18	0.16	0.25	0.34	0.21	0.18
<b>geog_zone5</b>	0.18	0.12	0.19	0.10	0.17	0.14	0.17	0.28	0.13	0.19
<b>q1 (ref.)</b>	0.20	0.18	0.32	0.25	0.22	0.30	0.34	0.40	0.33	0.32
<b>q2</b>	0.34	0.23	0.21	0.23	0.25	0.24	0.32	0.18	0.11	0.26
<b>q3</b>	0.26	0.26	0.24	0.23	0.32	0.28	0.09	0.16	0.20	0.22
<b>q4</b>	0.20	0.34	0.22	0.30	0.22	0.17	0.25	0.26	0.35	0.20

**Table 4.9: Continued**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5	
	FE	IE	FE	IE	FE	IE	FE	IE	FE	IE
<b>car</b>	0.31	0.17	0.04	0.12	0.06	0.05	0.11	0.06	0.07	0.13
<b>telecom</b>	0.97	1.00	0.87	0.95	0.74	0.71	0.46	0.50	0.59	0.54
<b>electricity</b>	0.63	0.74	0.65	0.64	0.61	0.54	0.39	0.37	0.40	0.48
<b>public_transport</b>	0.89	0.90	0.95	0.90	0.79	0.87	0.84	0.84	0.85	0.85
<b>water</b>	0.16	0.13	0.19	0.14	0.16	0.28	0.22	0.13	0.17	0.19
<b>toilet</b>	0.14	0.16	0.16	0.23	0.22	0.17	0.19	0.11	0.09	0.15
<b>Sample Size</b>	<b>164</b>	<b>24</b>	<b>121</b>	<b>33</b>	<b>57</b>	<b>63</b>	<b>37</b>	<b>89</b>	<b>28</b>	<b>469</b>

*Source:* Tanzania ILFS 2006. *Note:* Paid employed women aged 15 and older. Sampling weights are accounted for. The balancing property is satisfied. The balancing property analysis has been restricted to all treated plus those controls in the region of common support, and the significance level of the balancing property tests has been set at 0.01. The number of blocks of equal score range used at the beginning of the test of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated. FE means formal employment and IE refers to informal employment.

**Table 4.10: Mean covariates values by propensity score strata and employment status, men in self-employment**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4	
	FE	IE	FE	IE	FE	IE	FE	IE
<b>agegroup1 (ref.)</b>	0.10	0.14	0.10	0.15	0.16	0.13	0.19	0.18
<b>agegroup2</b>	0.16	0.27	0.30	0.30	0.35	0.39	0.47	0.42
<b>agegroup3</b>	0.74	0.58	0.60	0.55	0.49	0.48	0.35	0.41
<b>edu1 (ref.)</b>	0.02	0.02	0.04	0.05	0.02	0.03	0.09	0.11
<b>edu2</b>	0.06	0.01	0.13	0.03	0.00	0.07	0.12	0.16
<b>edu3</b>	0.60	0.63	0.67	0.62	0.69	0.72	0.73	0.68
<b>edu4</b>	0.32	0.34	0.16	0.29	0.29	0.18	0.06	0.06
<b>training</b>	0.46	0.58	0.57	0.52	0.36	0.39	0.27	0.20
<b>married</b>	0.75	0.66	0.71	0.62	0.59	0.61	0.64	0.71
<b>migrant</b>	0.07	0.06	0.05	0.09	0.12	0.12	0.14	0.13
<b>nb_child06</b>	0.62	0.82	0.72	0.61	0.58	0.65	1.03	1.09
<b>nb_child06sq</b>	0.92	1.46	0.93	0.91	0.83	1.08	2.23	2.40
<b>nb_child714</b>	0.80	0.76	0.76	0.59	0.89	0.71	0.70	0.90
<b>nb_child714sq</b>	1.93	1.46	1.62	1.14	2.01	1.58	1.38	2.08
<b>urban</b>	0.43	0.47	0.34	0.57	0.74	0.60	0.72	0.68
<b>geog_zone1 (ref.)</b>	0.22	0.24	0.19	0.29	0.39	0.35	0.38	0.27
<b>geog_zone2</b>	0.06	0.07	0.18	0.12	0.07	0.11	0.15	0.18
<b>geog_zone3</b>	0.23	0.25	0.28	0.24	0.21	0.22	0.21	0.19
<b>geog_zone4</b>	0.46	0.42	0.28	0.31	0.28	0.23	0.03	0.05
<b>geog_zone5</b>	0.02	0.02	0.07	0.05	0.04	0.09	0.23	0.30
<b>q1 (ref.)</b>	0.29	0.24	0.29	0.26	0.28	0.31	0.30	0.32
<b>q2</b>	0.44	0.37	0.41	0.40	0.23	0.27	0.17	0.17
<b>q3</b>	0.14	0.18	0.13	0.18	0.22	0.18	0.30	0.32
<b>q4</b>	0.13	0.20	0.17	0.17	0.26	0.24	0.22	0.19

**Table 4.10: Continued**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4	
	FE	IE	FE	IE	FE	IE	FE	IE
<b>car</b>	0.45	0.36	0.04	0.03	0.00	0.01	0.00	0.00
<b>telecom</b>	0.76	0.72	0.64	0.66	0.62	0.54	0.30	0.32
<b>electricity</b>	0.57	0.50	0.42	0.41	0.36	0.32	0.17	0.18
<b>public_transport</b>	0.62	0.59	0.67	0.66	0.84	0.77	0.87	0.91
<b>water</b>	0.07	0.20	0.16	0.13	0.18	0.21	0.29	0.25
<b>toilet</b>	0.18	0.22	0.15	0.13	0.03	0.07	0.08	0.04
<b>Sample Size</b>	<b>57</b>	<b>129</b>	<b>36</b>	<b>148</b>	<b>54</b>	<b>464</b>	<b>82</b>	<b>1433</b>

*Source:* Tanzania ILFS 2006. *Note:* Self-employed men aged 15 and older. Sampling weights are accounted for. The balancing property is satisfied. The balancing property analysis has been restricted to all treated plus those controls in the region of common support, and the significance level of the balancing property tests has been set at 0.01. The number of blocks of equal score range used at the beginning of the test of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated. FE means formal employment and IE refers to informal employment.

**Table 4.11: Mean covariates values by propensity score strata and employment status, women in self-employment**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4	
	FE	IE	FE	IE	FE	IE	FE	IE
<b>age</b>	36.14	34.62	33.38	35.19	32.20	33.88	35.10	34.74
<b>age2</b>	1369.47	1290.11	1222.46	1355.20	1067.09	1253.95	1405.19	1376.66
<b>edu1 (ref.)</b>	0.00	0.00	0.03	0.04	0.00	0.04	0.15	0.22
<b>edu2</b>	0.04	0.04	0.00	0.07	0.07	0.10	0.19	0.11
<b>edu3</b>	0.73	0.74	0.69	0.70	0.78	0.71	0.66	0.62
<b>edu4</b>	0.23	0.22	0.28	0.18	0.15	0.14	0.00	0.04
<b>training</b>	0.38	0.39	0.30	0.27	0.14	0.22	0.18	0.08
<b>married</b>	0.66	0.66	0.62	0.55	0.68	0.62	0.51	0.53
<b>migrant</b>	0.00	0.08	0.13	0.09	0.26	0.15	0.10	0.13
<b>nb_child06</b>	1.35	0.86	0.62	0.89	0.97	0.92	0.74	0.98
<b>nb_child06sq</b>	5.07	2.44	0.67	2.21	1.60	1.89	1.09	1.88
<b>nb_child714</b>	1.23	0.94	0.96	0.98	0.84	1.01	0.87	1.02
<b>nb_child714sq</b>	3.37	3.13	3.24	2.79	1.22	2.19	1.87	2.27
<b>urban</b>	0.38	0.49	0.63	0.59	0.82	0.71	0.76	0.75
<b>geog_zone1 (ref.)</b>	0.00	0.05	0.14	0.14	0.29	0.27	0.34	0.32
<b>geog_zone2</b>	0.11	0.00	0.05	0.12	0.15	0.14	0.15	0.15
<b>geog_zone3</b>	0.16	0.17	0.27	0.28	0.28	0.29	0.24	0.17
<b>geog_zone4</b>	0.73	0.77	0.53	0.45	0.29	0.26	0.03	0.05
<b>geog_zone5</b>	0.00	0.01	0.00	0.02	0.00	0.04	0.23	0.31
<b>q1 (ref.)</b>	0.38	0.32	0.29	0.37	0.33	0.29	0.32	0.31
<b>q2</b>	0.38	0.30	0.41	0.37	0.28	0.34	0.23	0.18
<b>q3</b>	0.04	0.05	0.06	0.09	0.16	0.15	0.24	0.32
<b>q4</b>	0.20	0.33	0.24	0.16	0.24	0.22	0.21	0.20

**Table 4.11: Continued**

	Stratum 1		Stratum 2		Stratum 3		Stratum 4	
	FE	IE	FE	IE	FE	IE	FE	IE
<b>car</b>	0.13	0.19	0.11	0.10	0.05	0.02	0.04	0.01
<b>telecom</b>	0.92	0.89	0.65	0.66	0.79	0.55	0.28	0.28
<b>electricity</b>	0.54	0.54	0.59	0.41	0.39	0.43	0.19	0.20
<b>public_transport</b>	0.92	0.98	0.79	0.89	1.00	0.93	0.95	0.85
<b>water</b>	0.18	0.25	0.22	0.23	0.13	0.27	0.35	0.21
<b>toilet</b>	0.05	0.20	0.15	0.16	0.16	0.12	0.17	0.06
<b>Sample Size</b>	<b>30</b>	<b>54</b>	<b>29</b>	<b>207</b>	<b>15</b>	<b>324</b>	<b>28</b>	<b>1126</b>

*Source:* Tanzania ILFS 2006. *Note:* Self-employed women aged 15 and older. Sampling weights are accounted for. The balancing property is satisfied. The balancing property analysis has been restricted to all treated plus those controls in the region of common support, and the significance level of the balancing property tests has been set at 0.01. The number of blocks of equal score range used at the beginning of the test of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated. FE means formal employment and IE refers to informal employment.

**Table 4.12: Frequency counts per propensity score stratum**

Paid Employment						Self Employment					
Men			Women			Men			Women		
P-Score	FE	IE	P-Score	FE	IE	P-Score	FE	IE	P-Score	FE	IE
[.07, .20]	250	32	[.05, .20]	164	24	[.45, .80]	57	129	[.55, .80]	30	54
[.20, .30]	138	50	[.20, .40]	121	33	[.80, .85]	36	148	[.80, .90]	29	207
[.30, .40]	124	62	[.40, .60]	57	63	[.85, .90]	54	464	[.90, .95]	15	324
[.40, .50]	120	83	[.60, .80]	37	89	[.90, 1.00]	82	1433	[.95, 1.00]	28	1126
[.50, .60]	102	135	[.80, 1.00]	28	469						
[.60, .70]	80	167									
[.70, .80]	59	256									
[.80, .90]	58	341									
[.90, 1.00]	36	413									

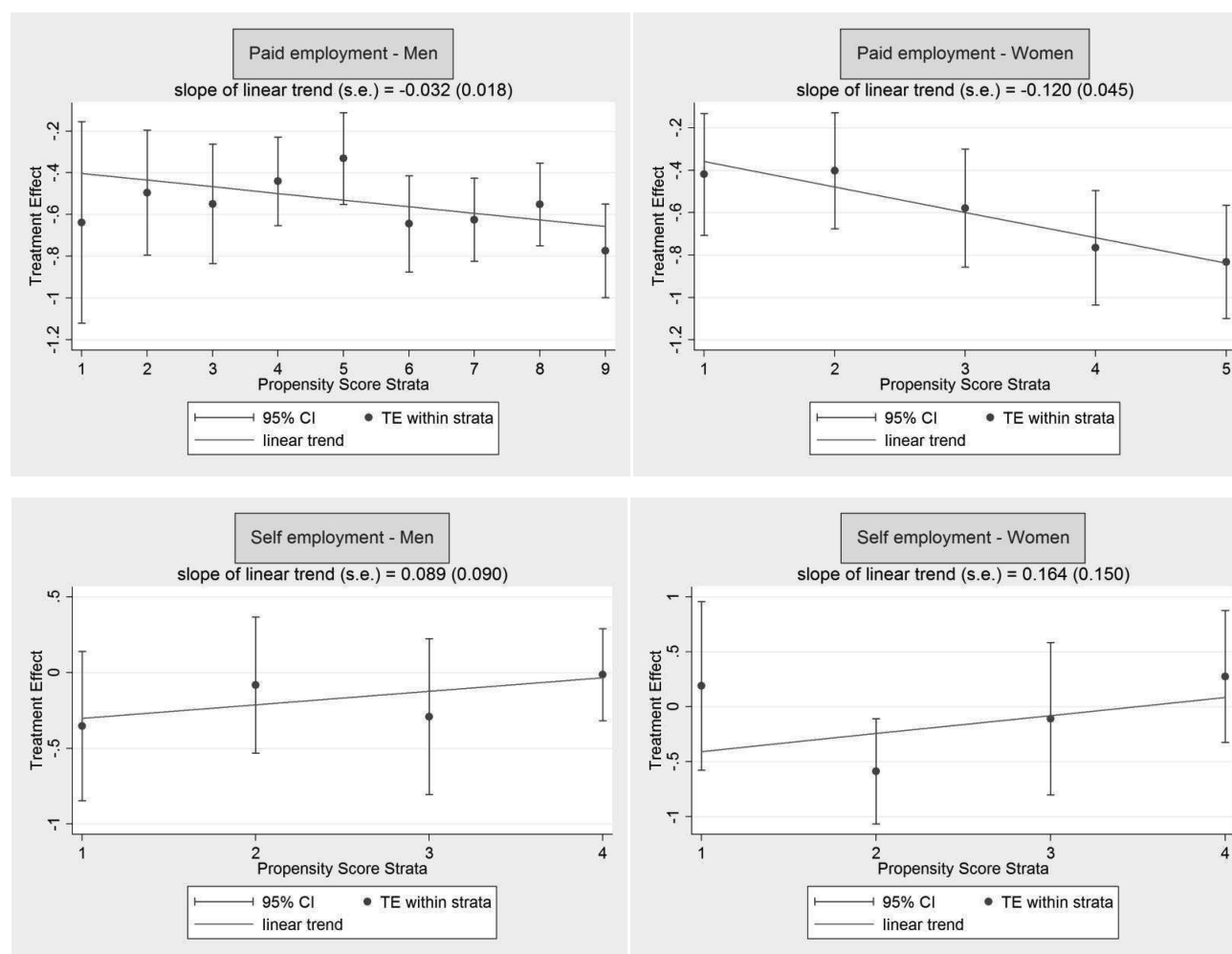
*Source:* Tanzania ILFS 2006. *Note:* Individuals aged 15 and older. Sampling weights are accounted for. The balancing property is satisfied. The balancing property analysis has been restricted to all treated plus those controls in the region of common support, and the significance level of the balancing property tests has been set at 0.01. The number of blocks of equal score range used at the beginning of the test of the balancing hypothesis has been set to five. Small strata at the low and high end of the propensity score have been merged with subsequent or precedent strata, respectively, so that the number of observations is at least 20 for both the treated and the untreated. FE means formal employment and IE refers to informal employment.



**Table 4.13: Heterogeneous effects of informal employment on log hourly income from main job by employment status and gender (SM-HTE)**

	Paid Employment		Self Employment	
	Men	Women	Men	Women
	<b>Level-1 Slopes</b>		<b>Level-1 Slopes</b>	
	OLS Regression		OLS Regression	
<b>P-Score Stratum 1</b>	-0.6378*** (-2.59)	-0.4189*** (-2.86)	-0.3522 (-1.40)	0.1909 (0.49)
<b>P-Score Stratum 2</b>	-0.4947*** (-3.25)	-0.4025*** (-2.88)	-0.0827 (-0.36)	-0.5876** (-2.40)
<b>P-Score Stratum 3</b>	-0.5485*** (-3.75)	-0.5781*** (-4.06)	-0.2910 (-1.11)	-0.1087 (-0.31)
<b>P-Score Stratum 4</b>	-0.4404*** (-4.07)	-0.7648*** (-5.54)	-0.0127 (-0.08)	0.2768 (0.90)
<b>P-Score Stratum 5</b>	-0.3313*** (-2.97)	-0.8318*** (-6.10)	—	—
<b>P-Score Stratum 6</b>	-0.6441*** (-5.47)	—	—	—
<b>P-Score Stratum 7</b>	-0.6244*** (-6.14)	—	—	—
<b>P-Score Stratum 8</b>	-0.5511*** (-5.47)	—	—	—
<b>P-Score Stratum 9</b>	-0.7739*** (-6.77)	—	—	—
	<b>Level-2 Slope</b>		<b>Level-2 Slope</b>	
	Variance-Weighted Least Squares		Variance-Weighted Least Squares	
<b>slope</b>	-0.0317* (-1.76)	-0.1200*** (-2.69)	0.0894 (0.99)	0.1636 (1.09)
<b>constant</b>	-0.3715*** (-3.33)	-0.2385 (-1.59)	-0.3902 (-1.38)	-0.5710 (-1.39)
<b>Sample Size</b>	<b>2515</b>	<b>1098</b>	<b>2403</b>	<b>1814</b>

*Source:* Tanzania ILFS 2006. *Note:* Individuals aged 15 and older. Regressions account for sampling weights. Heteroskedasticity-robust T statistics in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. To obtain level-1 slopes, we estimate propensity score stratum-specific treatment effects within strata. To do this, we apply a linear regression model within strata to further adjust for any remaining covariate imbalance within strata. Because we do not constrain the comparison of the treatment group and the control group across strata in any way, data analysis at this stage is non-parametric across strata. Control variables for within-strata models for men are: agegroup2, agegroup3, edu2-edu4, training, married, migrant, urban, geog\_zone2-geog\_zone5, and q2-q4. And control variables for within-strata models for women are: age, age2, edu2-edu4, training, married, migrant, urban, geog\_zone2-geog\_zone5, and q2-q4. We estimate Level-2 slope to evaluate a trend of the stratum-specific treatment effects across the strata. Level-2 slope is obtained by regressing the stratum-specific treatment effects on strata rank using variance-weighted least squares (the variance is based on the standard errors of the stratum-specific treatment effects). We thus model the heterogeneity pattern as a linear function across strata ranks in order to force the data to tell us whether the treatment effect is either a positive or a negative function of propensity.



**Figure 4.1: Heterogeneous effects of informal employment on log hourly income from main job by employment status and gender (SM-HTE)**

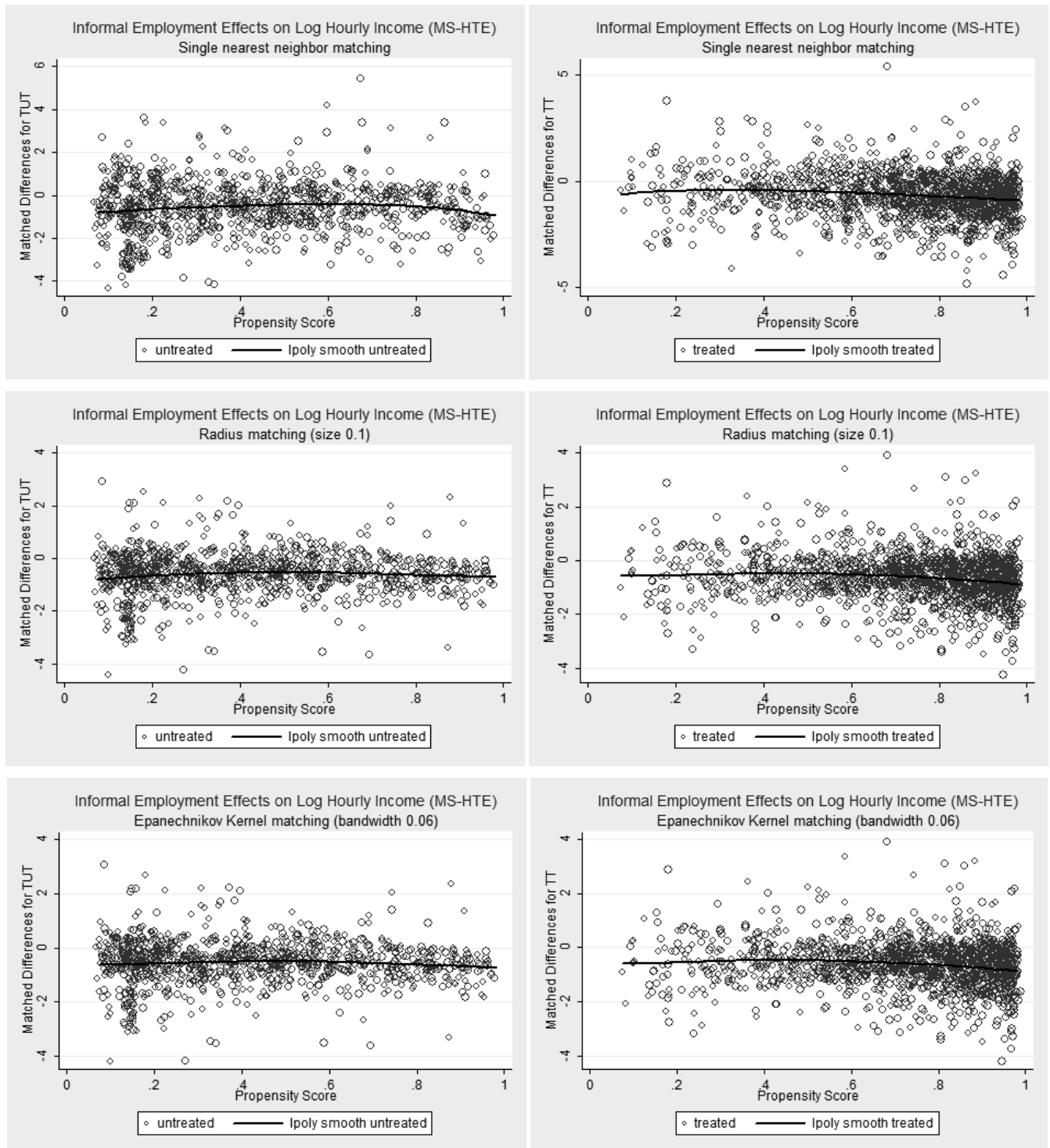
*Source:* Tanzania ILFS 2006. *Note:* Individuals aged 15 and older. “Dots” represent point estimates of level-1 slopes, i.e. stratum-specific OLS regression effects of informal employment on log hourly income from main job. The linear plot in the figure is the level-2 variance-weighted least squares slope (see footnote of Table 4.13).

## Appendix E. Informal employment effects under the assumption of full heterogeneity

**Table 4.14: Matching estimates effects of informal employment on log hourly income from main job by employment status and gender**

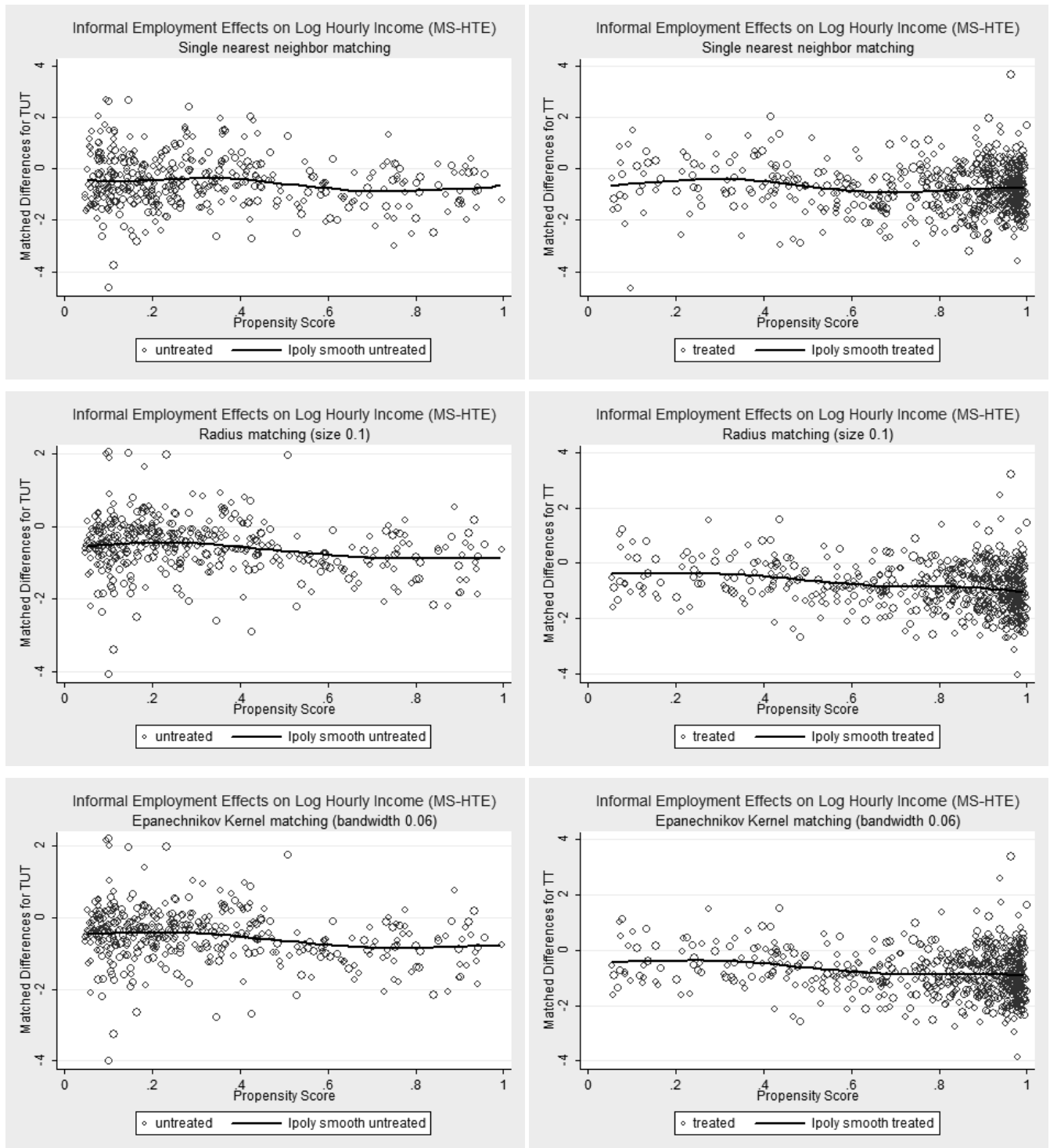
	Paid Employment						Self Employment					
	Men			Women			Men			Women		
	ATE	ATT	ATU	ATE	ATT	ATU	ATE	ATT	ATU	ATE	ATT	ATU
<b>Single nearest neighbour matching</b>	-0.6128***	-0.6571***	-0.5449***	-0.6624***	-0.7760***	-0.4953***	-0.0916	-0.0825	-0.1775	0.0555	0.0601	-0.0158
Matching with replacement (without caliper)	(0.0484)	(0.0517)	(0.0818)	(0.0749)	(0.0827)	(0.1183)	(0.1091)	(0.1148)	(0.1789)	(0.2361)	(0.2487)	(0.2003)
<b>Radius matching</b>	-0.6174***	-0.6351***	-0.5902***	-0.7288***	-0.8644***	-0.5292***	-0.1994*	-0.1967*	-0.2253**	-0.0143	-0.0137	-0.0234
Radius size 0.1	(0.0447)	(0.0477)	(0.0611)	(0.0838)	(0.0957)	(0.0873)	(0.1019)	(0.1040)	(0.0955)	(0.1769)	(0.1812)	(0.1513)
<b>Kernel matching</b>	-0.5968***	-0.6243***	-0.5545***	-0.6924***	-0.8207***	-0.5034***	-0.1667	-0.1642	-0.1908	0.0513	0.0533	0.0202
Epanechnikov kernel - bandwidth 0.06	(0.0386)	(0.0395)	(0.0618)	(0.0792)	(0.0918)	(0.0838)	(0.1306)	(0.1330)	(0.1201)	(0.1978)	(0.2050)	(0.1340)
<b>Sample Size</b>	<b>2515</b>			<b>1098</b>			<b>2403</b>			<b>1814</b>		

*Source:* Tanzania ILFS 2006. *Note:* Individuals aged 15 and older. Bootstrapped standard errors (100 replications) in parenthesis. \* significant at 10%, \*\* significant at 5%, \*\*\* significant at 1%. ATE is the average treatment effect, ATT the average treatment effect on the treated, and ATU the average treatment effect on the untreated. Only observations in the common support region are used for calculating treatment effects. For information about propensity score estimation, see Table 4.7 and footnote of Table 4.12. There is no clear consensus as to which matching estimator performs best in each application. We thus implement various propensity score matching methods (single nearest neighbour, radius, and kernel matching) for the sake of illustration and comparison.



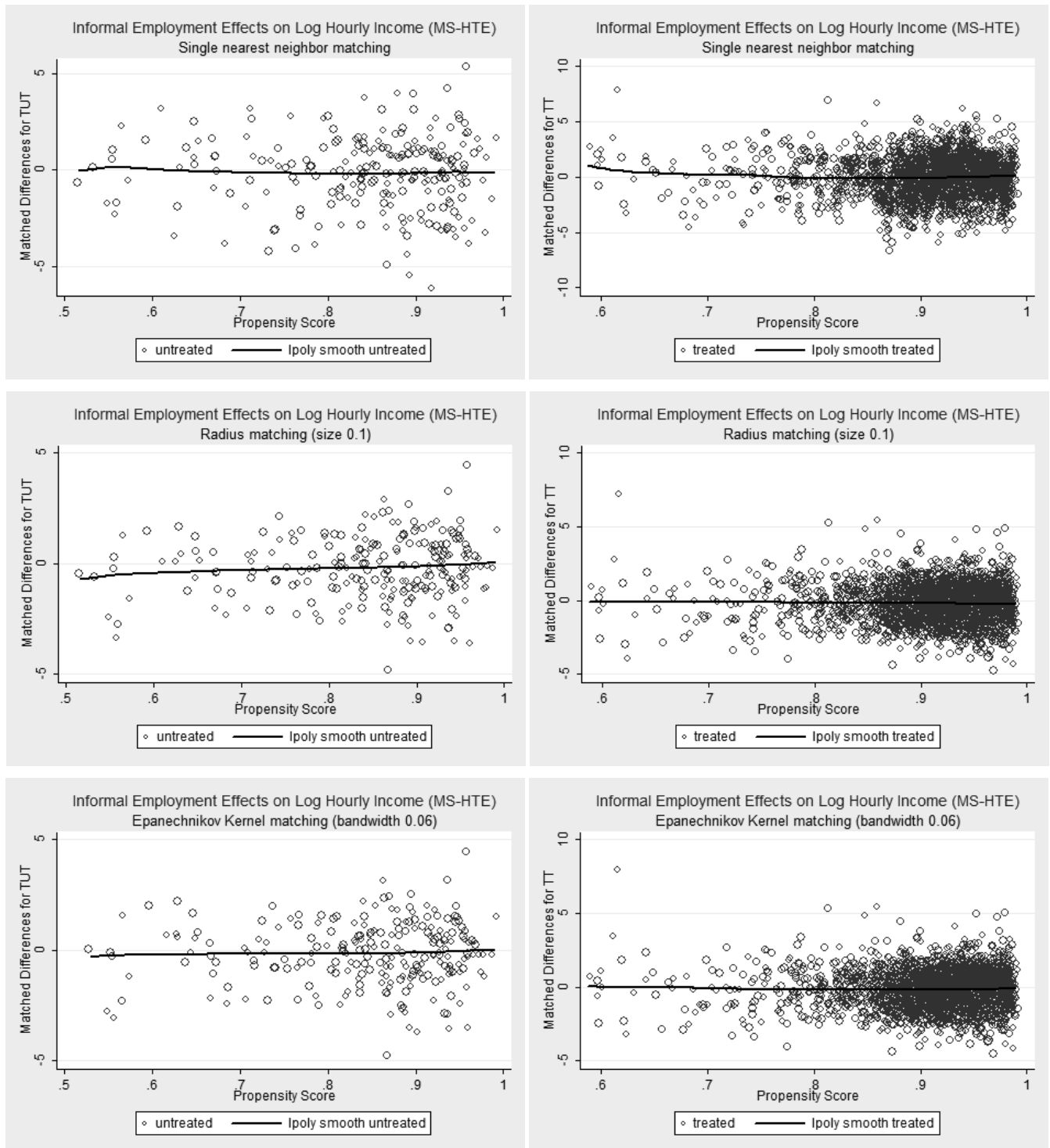
**Figure 4.2: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), men in paid employment**

*Source:* Tanzania ILFS 2006. *Note:* Paid-employed men aged 15 and older. Sampling weights are accounted for. Only observations in the common support region are used for calculating treatment effects. The figure plots the matched differences between treated and control units (right-hand side graphs) and between control and untreated units (left-hand side graphs) against continuous representation of the propensity score, and fits nonparametric smoothed curves using different matching estimators. We perform kernel-weighted local polynomial regressions to construct the curves. We select the Epanechnikov kernel function in calculating the weighted local polynomial estimate, and we choose a one-degree polynomial for smoothing. A small amount of spherical random noise is added to each point before graphing in order to reduce overplotting.



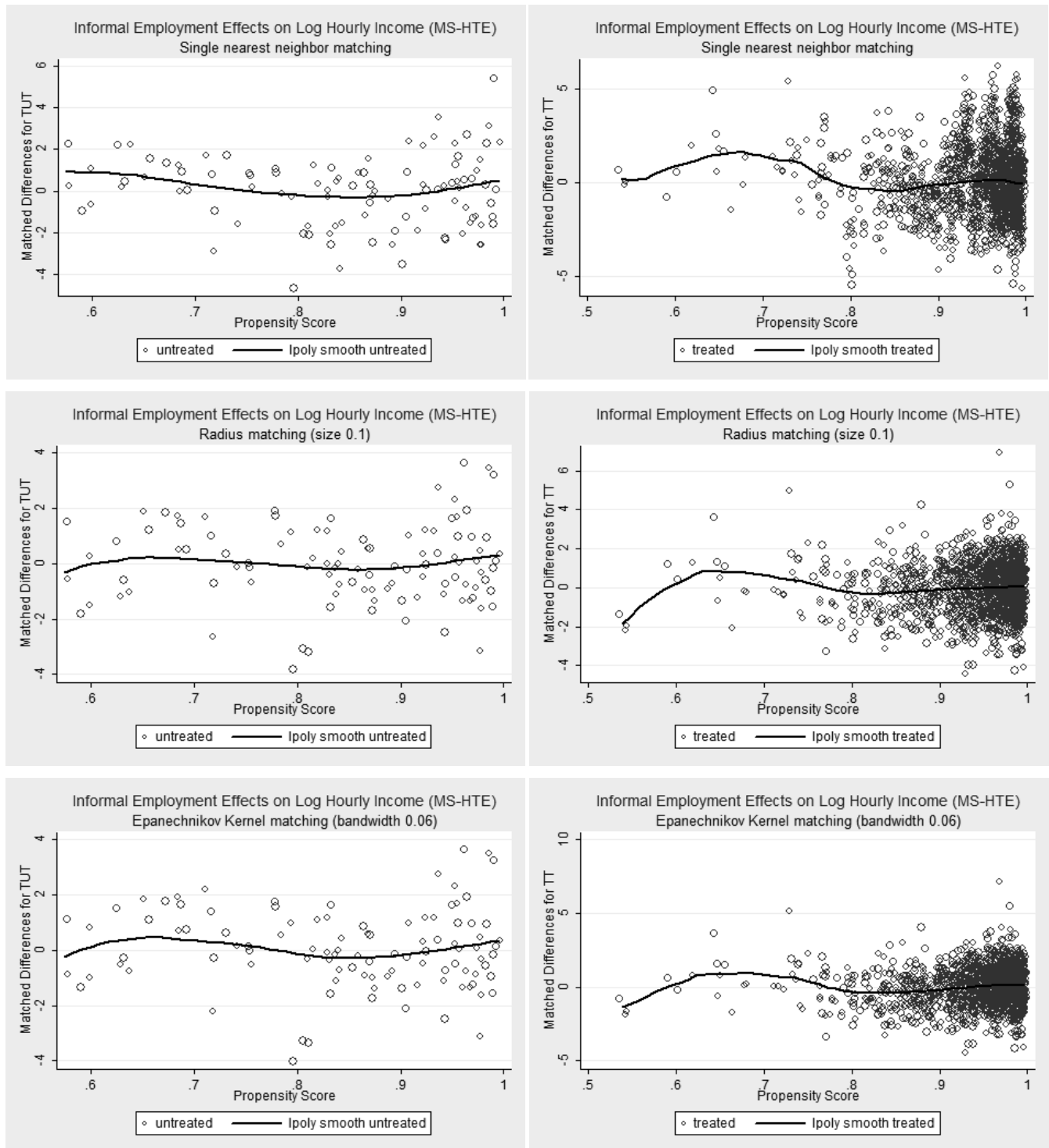
**Figure 4.3: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), women in paid employment**

*Source:* Tanzania ILFS 2006. *Note:* Paid-employed women aged 15 and older. Sampling weights are accounted for. Only observations in the common support region are used for calculating treatment effects. The figure plots the matched differences between treated and control units (right-hand side graphs) and between control and untreated units (left-hand side graphs) against continuous representation of the propensity score, and fits nonparametric smoothed curves using different matching estimators. We perform kernel-weighted local polynomial regressions to construct the curves. We select the Epanechnikov kernel function in calculating the weighted local polynomial estimate, and we choose a one-degree polynomial for smoothing. A small amount of spherical random noise is added to each point before graphing in order to reduce overplotting.



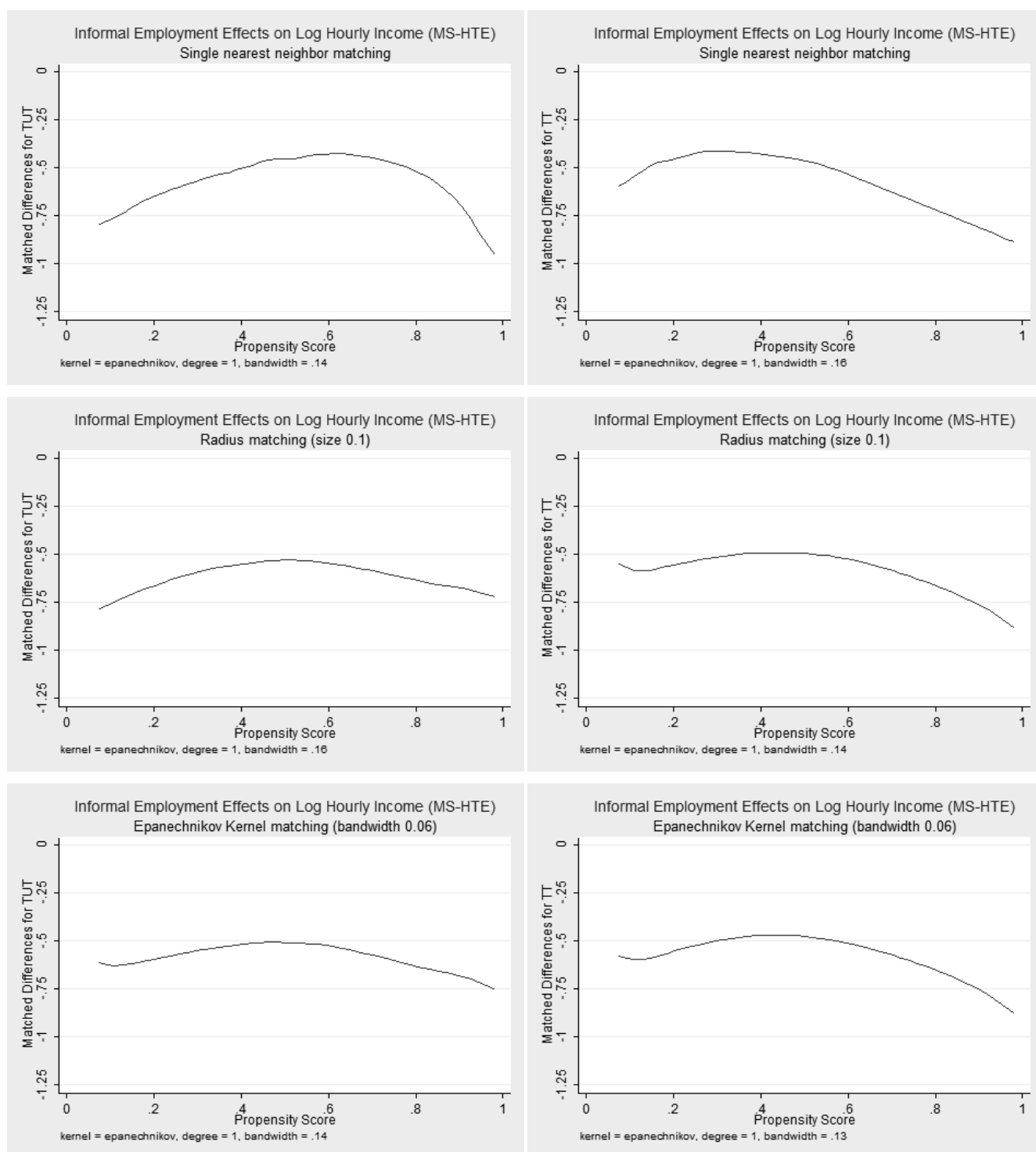
**Figure 4.4: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), men in self-employment**

*Source:* Tanzania ILFS 2006. *Note:* self-employed men aged 15 and older. Sampling weights are accounted for. Only observations in the common support region are used for calculating treatment effects. The figure plots the matched differences between treated and control units (right-hand side graphs) and between control and untreated units (left-hand side graphs) against continuous representation of the propensity score, and fits nonparametric smoothed curves using different matching estimators. We perform kernel-weighted local polynomial regressions to construct the curves. We select the Epanechnikov kernel function in calculating the weighted local polynomial estimate, and we choose a one-degree polynomial for smoothing. A small amount of spherical random noise is added to each point before graphing in order to reduce overplotting.



**Figure 4.5: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), women in self-employment**

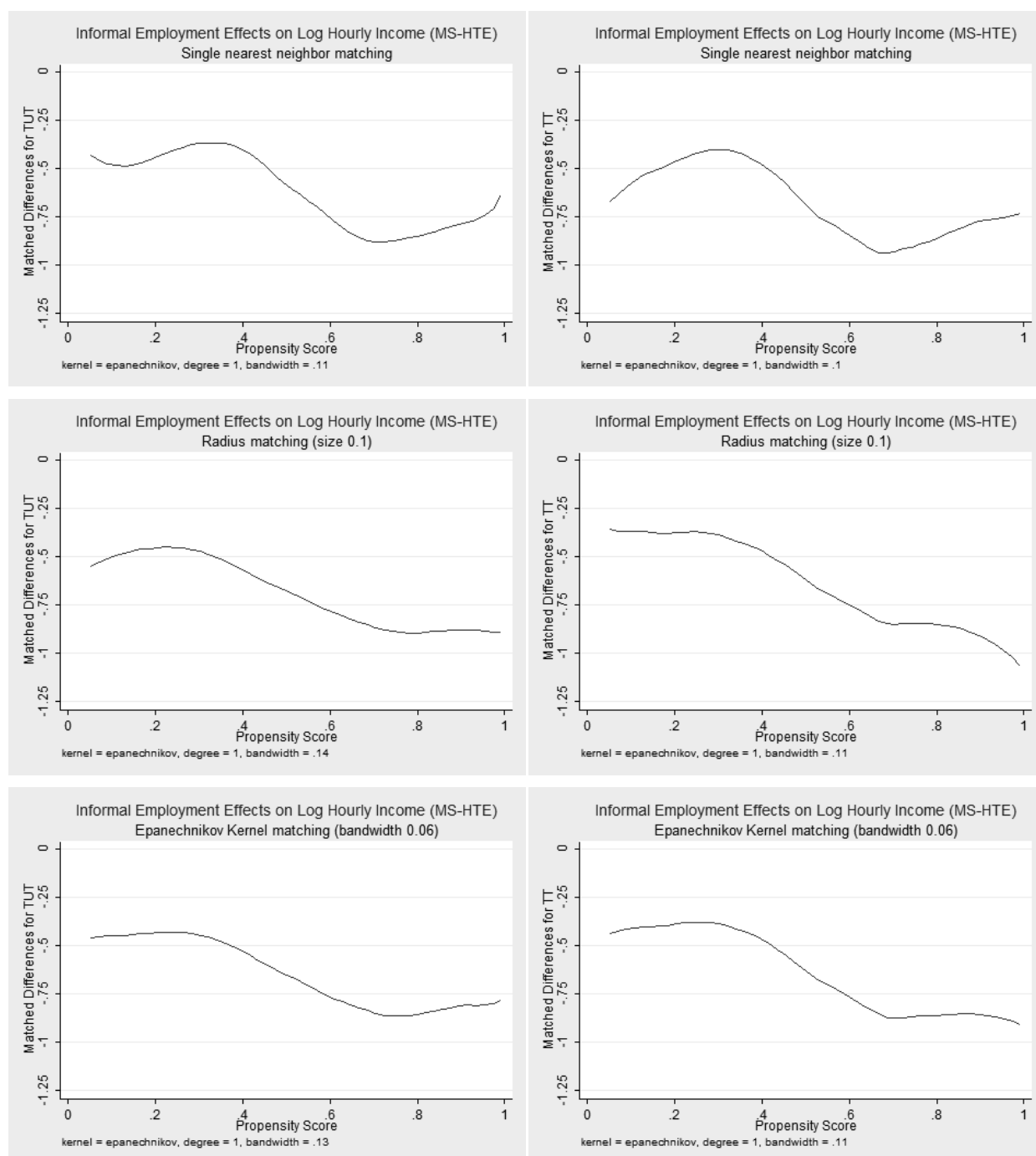
*Source:* Tanzania ILFS 2006. *Note:* self-employed women aged 15 and older. Sampling weights are accounted for. Only observations in the common support region are used for calculating treatment effects. The figure plots the matched differences between treated and control units (right-hand side graphs) and between control and untreated units (left-hand side graphs) against continuous representation of the propensity score, and fits nonparametric smoothed curves using different matching estimators. We perform kernel-weighted local polynomial regressions to construct the curves. We select the Epanechnikov kernel function in calculating the weighted local polynomial estimate, and we choose a one-degree polynomial for smoothing. A small amount of spherical random noise is added to each point before graphing in order to reduce overplotting.



**Figure 4.6: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), graphs of the nonparametric smoothed curves without plotted points, men in paid employment**

*Source:* Tanzania ILFS 2006. *Note:* This figure only differs from Appendix D Figure 4.2 (see the footnote) in that it suppresses superimposing a scatter-plot over the smooth (this option is very useful in our case because the number of plotted points is so large as to clutter the graphs) and shortens the range of y-axis values considerably, in order to better assess the potential non-linearity of the heterogeneity pattern of informal employment returns.





**Figure 4.7: Heterogeneous effects of informal employment on log hourly income from main job (MS-HTE), graphs of the nonparametric smoothed curves without plotted points, women in paid employment**

*Source:* Tanzania ILFS 2006. *Note:* This figure only differs from Appendix D Figure 4.3 (see the footnote) in that it suppresses superimposing a scatter-plot over the smooth (this option is very useful in our case because the number of plotted points is so large as to clutter the graphs) and shortens the range of y-axis values considerably, in order to better assess the potential non-linearity of the heterogeneity pattern of informal employment returns.

## **Chapter 5**

### **Conclusion**

The objective of this research was to contribute to our better understanding of the main factors behind large and persistent gender disparities in Africa's labour markets, using the 2005 Ethiopia Labour Force Survey and the 2006 Tanzania Integrated Labour Force Survey. The most important findings of this work are summarized as follows.

In Chapter 2, we saw that women, in Ethiopia, face substantially lower wages compared to men. The pay gap with men tends to decrease as women get older and are better-paid, and when they hold formal jobs, especially in the public sector.

Job-related characteristics, in particular selection across industries and occupations, are the main contributors of the gender wage differential, followed by human capital endowments (education, potential work experience and training). The gap attributable to non-observable factors (i.e., discriminatory practices, gender specific preferences, cultural and social norms, and other unobservables) still represents an important proportion (between 15 and 23 per cent).

Job attributes appear to be, in turn, significantly driven, not only by human capital endowments, but also by gender status, which either picks up a form of sex-based discriminatory sorting/segmentation, and/or gender specific preferences.

It is among youths and low-wage earners that the gender wage gap is the more influenced by job characteristics and the less by unobservable factors. As regards human capital, it explains a larger share of the wage differential among the elderly, low-paid workers and public sector employees. In the private sector, informal workers differ from formal workers in that their observable characteristics (human capital, job and other attributes) contribute much more to the gender wage gap.

In Chapter 3, we highlighted important disadvantages that women face in Ethiopia in terms of working hours. Our findings indicate the coexistence of two phenomena: a strong gender-based division of labour and a double work burden on women, which are exacerbated in rural and remote areas where usually access to basic infrastructure is limited, tribal law prevails, and traditional gender roles are the more deeply rooted.

We observed that women (men) participate more and spend longer hours than men (women) in household (market) work. However, despite the fact that market work is predominantly masculine, the proportion of women working in the labour market and the time they spend in it are important and much higher than the incidence and the average duration of housework for men. Thus, compared to men, who generally focus only on market work, women tend to accumulate both types of work, and thereby are double-burdened. These findings are observed not only on average, but also at all points of the population distribution.

In urban areas, there is substitutability in time allocation decisions. While housework time appears to be quite insensitive and inelastic to market work hours, especially among men, the amount of hours devoted to productive activities seems to be, in turn, constrained and conditioned somehow by the amount of hours allocated to household chores.

In rural areas, the same holds for men, with the nuance that the adverse impact of housework hours on market work time is less strong in the countryside than in cities. For women, there is no evidence that the time spent working at home affects and constrains the time spent working in the labour market, and vice versa.

Our results also indicate that the effect of education on labour supply is strongly gendered. First, education has generally a negative influence on the time allocated to housework, especially among women. Second, education is usually a negative predictor of the time devoted to market work by men, and a positive predictor of that devoted by women.

In Chapter 4, we looked at the impact of informality on wages and showed that women face a significantly higher informal employment wage penalty than men. To explain this result, we

conjecture that the exclusion hypothesis is more prevalent among women. According to this hypothesis, individuals are denied access to formal jobs due to the disproportionate constraints they face and have no choice but to work in informal employment, which is mostly a residual where activities are being pursued in the absence of other options.

Our results argue in favour of the exclusion hypothesis, indicating that informal paid-employed workers, and in particular women, are affected by a range of constraints, such as, for instance, the burden of household responsibilities and the lack of adequate infrastructure, that constitute important obstacles which may explain why they fall back in informal employment. The decision to work informally is thus probably not the mere result of a rational choice in which people weigh the expected economic returns against the costs based solely on economic factors. This decision is very likely to be governed also by cultural and social norms, circumstances and constraints.

This research has also recognized some of the limitations and shortcomings of the data used, highlighting possible areas of improvement for further data collection.

In Chapter 2, we were unable to extend the analysis to self-employment, which comprises the vast majority of workers in Ethiopia, because the survey only collected information on earnings for paid-employees.

Besides, we could neither account for seasonality, given that survey data collection took place in a very short period of time, nor for some interesting variables that are very likely to affect labour market outcomes such as workers' actual experience in the labour market, ethnicity, religion and language, simply because they were missing from the survey.

In our attempt to apply estimation procedures that differentiate participation decision from wage determination, the little information available in the survey did not allow us to find valid instruments and we had no choice but to base the model identification solely upon the non-linearity in the functional form.

We also tried to implement a recent decomposition technique proposed by Appleton *et al.* (1999) that takes into account sectoral structures between genders. We finally did not pursue as we were unable with data at hand to accurately model participation in each wage employment sector.

In Chapter 3, the main obstacle encountered was that no surveys using time-diary methods had been carried out in Ethiopia, which forced us to base our results on stylized (questionnaire-based) estimates, while they are clearly less accurate and detailed than diary-based estimates.

In the regression analysis, we insistently tried to deal with endogeneity issues and account for the simultaneity of time allocation decisions, but we did not manage to do it because the data did not provide variables which could reasonably be used as valid instruments for market and household work hours.

In addition, the lack of some crucial information in the survey made it difficult to investigate the determinants of time allocation decisions. For instance, variables on access to basic infrastructure resources were missing, while they undoubtedly are important determinants of the time allocated by individuals to market and household work.

As regards Chapter 4, the data available did not allow us to extent the analysis to self-employment. Observed characteristics failed in explaining most of the variance of self-employment income across individuals in Tanzania, suggesting that our model was not properly specified and that some relevant variables were omitted. Self-employment income is typically generated at the household level, and is usually derived not only from human capital but also from physical capital. Our data offered little information on enterprise characteristics and no information on workers' actual experience, while they arguably constitute crucial determinants of income dispersion. As a consequence, we could not draw any conclusions regarding self-employment.

Notwithstanding these shortcomings, we believe that our results, based on the analysis of the 2005 Ethiopia Labour Force Survey and the 2006 Tanzania Integrated Labour Force Survey, have complemented, in a sizeable manner, the thin existing literature on the subject and could contribute to add to the policy debate on the way to address the most pressing gender inequality in Africa.

While not principally a research on policies to reduce gender inequality, our empirical analysis touches indeed upon policies issues at various points. One important policy implication is that progress towards gender equity in education in Africa is essential to reduce gender disparities in labour market outcomes, as education is a strong determinant of women's access to productive activities, and thus an important factor to support women's economic empowerment (Chapter 3). Moreover, reducing gender inequality in access to education is important to improve women's wages (Chapter 2).

However, we also find that achieving gender equality in education will not be enough to address the barriers that women face in the labour market. A substantial part of the wage gap in Ethiopia remains unexplained (Chapter 2), while in Tanzania, women face a substantially higher wage penalty by working informally compared to men (Chapter 4).

All in all, this suggests that, besides achieving parity in education, other interventions would be needed. Relevant interventions include, for instance, the development of basic infrastructure services to allow women to spend more time in productive activities and less in household chores, and the implementation of information campaigns and other awareness-raising efforts in support of national legislation's anti-discriminatory provisions, to compensate for the adverse impact of discriminatory practices, gender specific preferences, social and cultural norms, and other unobserved factors, that contribute to women's disadvantage in several labour market outcomes.

Looking forward, and on the basis that new and more comprehensive data could become available in Africa, it would be interesting to extend this research by looking at the way social and cultural norms, access to basic infrastructure services, and ethnicity interact with gender to explain some of the most pressing inequalities in Africa's labour markets.

## References

- Amemiya, T. (1979) 'The Estimation of a Simultaneous-Equation Tobit Model', *International Economic Review*, Vol. 20, number 1, pp. 169-181.
- Amemiya, T. (1974) 'Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal', *Econometrica*, Vol. 42, pp. 999-1021.
- Angel-Urdinola, D., and Q. Wodon (2010) 'Income Generation and Intra-Household Decision Making: A Gender Analysis for Nigeria', J.S. Arbache, A. Kolev and E. Filipiak (eds), *Gender Disparities in Africa's Labour Markets*, Washington, DC: The World Bank.
- Appleton, S., J. Hoddinott and P. Krishnan (1999) 'The Gender Wage Gap in Three African Countries', *Economic Development and Cultural Change*, 47 (2): 289-312.
- Appleton, S., J. Hoddinott, P. Krishnan and M. Kerry (1995) 'Does the Labour Market Explain Low Female Schooling? Evidence from Three African Countries', Queen Elizabeth House Development Studies Working Paper No.83, University of Oxford.
- Apps, P. F., and R. Rees (2002) 'Household Production, Full Consumption and the Costs of Children', *Labour Economics*, Vol. 8, pp. 621-648.
- Apps, P. F., and R. Rees (1988) 'Taxation and the Household', *Journal of Public Economics*, Vol. 35, pp. 155-169.
- Arbache, J.S., A. Kolev, and E. Filipiak (2010) 'Gender Disparities in Africa's Labour Markets', Washington, DC: The World Bank.
- Backiny-Yetna, P., and Q. Wodon (2010) 'Gender Labour Income Shares and Human Capital Investment in the Republic of Congo', J.S. Arbache, A. Kolev and E. Filipiak (eds), *Gender Disparities in Africa's Labour Markets*, Washington, DC: The World Bank.
- Bardasi, E., and Q. Wodon (2010) 'Working Long Hours and Having No Choice: Time Poverty in Guinea', *Feminist Economics*, Vol. 16, number 3, pp. 45-78.
- Bardasi, E., and Q. Wodon (2006) 'Measuring Time Poverty and Analysing Its Determinants: Concepts and Application to Guinea', *Economics Bulletin*, Vol. 10, number 10, pp. 1-7.
- Barr, A. and A. Oduro (2002) 'Ethnic Fractionalization in an African Labour Market', *Journal of Development Economics*, 68 (2): 355-79.
- Beblo, M., D. Beninger, A. Heinze and F. Laisney (2003) 'Methodological Issues Related to the Analysis of Gender Gaps in Employment, Earnings and Career Progression', Final Report, Project carried out for the European Commission Employment and Social Affairs DG.
- Becker, G. S. (1965) 'A Theory of the Allocation of Time', *Economic Journal*, Vol. 75, number 299, pp. 493-517.
- Bhalotra, S. (2003) 'Child Labour in Africa', OECD Social, Employment and Migration Working Papers, No. 4, OECD Publishing.



- Bigsten, A., T. Mengistae and A. Shimeles (2007) 'Mobility and Earnings in Ethiopia's Urban Labour Markets, 1994–2004', Policy Research Working Paper Series 4168, World Bank.
- Blackden, C. M., and E. Morris-Hugues (1993) 'Paradigm Postponed: Gender and Economic Adjustment in Sub-Saharan Africa', Technical Note 13, Poverty and Human Resources Division, Technical Department, Africa Region, Washington, DC: The World Bank.
- Blackden, C. M., and Q. Wodon (2006) 'Gender, Time Use, and Poverty in Sub-Saharan Africa', Working Paper 73, Washington, DC: The World Bank.
- Blinder, A.S. (1973) 'Wage Discrimination: Reduced form and Structural Estimates', *Journal of Human Resources*, 8: 436–55.
- Blinder, A., and Weiss, Y. (1976) 'Human Capital and Labour Supply: A Synthesis', *Journal of Political Economy*, Vol. 84, number 3, pp. 449–72.
- Bourguignon, F., and P. A. Chiappori (1992) 'Collective Models of Household Behaviour: An Introduction', *European Economic Review*, Vol. 36, number 2–3, pp. 355–364.
- Bourguignon, F., M. Fournier and M. Gurgand (2007) 'Selection Bias Corrections Based on the Multinomial Logit Model: Monte Carlo Comparisons', *Journal of Economic Surveys*, 21 (1): 174–205.
- Brand, J. (2010) 'Civic Returns to Higher Education: A Note on Heterogeneous Effects', *Social Forces*, Vol. 89, number 2, pp. 417–433.
- Brand, J., and D. Davis (forthcoming 2011) 'The Impact of College Education on Fertility: Evidence for Heterogeneous Effects', *Demography*.
- Brand, J., and Y. Xie (2010) 'Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education', *American Sociological Review*, Vol. 75, number 2, pp. 273–302.
- Browning, M., and P. A. Chiappori (1998) 'Efficient Intra-Household Allocations: A General Characterization and Empirical Tests', *Econometrica*, Vol. 66, number 6, pp. 1241–78.
- Buchinsky, M. (2001). 'Quantile Regression with Sample Selection: Estimating Women's Return to Education in the US', *Empirical Economics*, 26(1): 87–113.
- Burda, M., D. S. Hamermesh, and P. Weil (2007) 'Total Work, Gender and Social Norms', Working Paper 13000, National Bureau of Economic Research, Washington, DC.
- Bussolo, M., R. E. De Hoyos, and Q. Wodon (2009) 'Higher Prices of Export Crops, Intra-Household Inequality and Human Capital Accumulation in Senegal', M. Bussolo and R. E. De Hoyos (eds), *Gender Aspects of the Trade and Poverty Nexus: A Macro–Micro Approach*, Washington, DC: World Bank and Palgrave Macmillan.
- Castells, M., and A. Portes (1989) 'World Underneath: The Origins, Dynamics and Effects of the Informal Economy', A. Portes, M. Castells and L.A. Benton (eds), *The Informal Economy: Studies in Advanced and Less Developed Countries*, Johns Hopkins University Press, Baltimore, MD, pp. 11–40.

- Charmes, J. (2006) 'A Review of Empirical Evidence on Time Use in Africa from UN-Sponsored Surveys', C. M. Blackden and Q. Wodon (eds), *Gender, Time Use, and Poverty in Sub-Saharan Africa*, World Bank Working Paper No. 73, Washington, DC: The World Bank.
- Chiappori, P.-A. (1988) 'Rational Household Labour Supply', *Econometrica*, Vol. 56, pp. 63-90.
- CSA – Central Statistical Agency (2006) 'Report on the 2005 National Labour Force Survey', Statistical Bulletin 365, CSA, Addis Ababa, Ethiopia.
- Cotton, J. (1988) 'On the Decomposition of Wage Differentials', *The Review of Economics and Statistics*, 236–43.
- Dahl, G.B. (2002) 'Mobility and the Returns to Education: Testing a Roy Model with Multiple Markets', *Econometrica*, 70: 2367–420.
- Dawson, D. (2007) "A Gendered Analysis of 'Time Poverty' – The Importance of Infrastructure", Global Poverty Research Group Working Paper 78, Economic and Social Research Council, Swindon, UK.
- Dercon, S. and P. Krishnan (2000) 'Vulnerability, Seasonality and Poverty in Ethiopia', *Journal of Development Studies*, 36 (6): 25–53.
- De Vreyer, P., and F. Roubaud (2012) 'Urban Labour Markets in Sub-Saharan Africa', Washington, DC: The World Bank and the French Development Agency (Forthcoming).
- Dubin, J.A. and D.L. McFadden (1984) 'An Econometric Analysis of Residential Electric Appliance Holdings and Consumption', *Econometrica*, 52: 345–62.
- Ellis, A., M. Blackden, J. Cutura, F. MacCulloch, and H. Seebens (2007) 'Gender and Economic Growth in Tanzania – Creating Opportunities for Women', Washington, DC: The World Bank.
- Fafchamps, M., M. Söderbom, and N. Benhassine (2008) 'Wage Gaps and Job Sorting in African Manufacturing', ESRC Global Poverty Research Group.
- Fafchamps, M., M. Soderbom and N. Benhassine (2006) 'Job Sorting in African Labour Markets', Working Paper Series 2006-02, Centre for the Study of African Economies, University of Oxford.
- Fenstermaker Berk, S., and A. Shih (1980) "Contributions to Household Labour: Comparing Wives' and Husbands' Reports", S. Fenstermaker Berk (ed.), *Women and Household Labour*, Beverly Hills, CA: Sage.
- Fields, G.S. (2005) 'A Guide to Multisector Labour Market Models', Social Protection Discussion Paper Series No. 0505, Washington, DC: The World Bank.
- Fields, G.S. (1990) 'Labour Market Modelling and the Urban Informal Sector: Theory and Evidence', D. Turnham, B. Salomé and A. Schwarz (eds), *The Informal Sector Revisited*, Development Centre Seminars, Paris: The OECD.
- Flood, L., and U. Grasjo (2001) 'A Monte Carlo Simulation Study of Tobit Models', *Applied Economics Letters*, Vol. 8, pp. 581-584.
- Folbre, N. (2004) 'A theory of the Misallocation of Time', N. Folbre and M. Bittman (eds), *Family Time: the Social Organization of Care*, London: Routledge.

- Foster, J. E., J. Greer, and E. Thorbecke (1984) 'A Class of Decomposable Poverty Indices', *Econometrica*, Vol. 52, number 3, pp. 761–66.
- Gannon, B., R. Plasman, F. Rycx and I. Tojerow (2007) 'Inter-industry Wage Differentials and the Gender Wage Gap: Evidence from European Countries', *The Economic and Social Review*, 38 (1): 135–55.
- Ghez, G., and Becker, G. S. (1975) 'The allocation of Time and Goods over the Life Cycle', NY: Columbia University Press.
- Glick, P. (1999) 'Simultaneous Determination of Home Work and Market Work of Women in Urban West Africa', *Oxford Bulletin of Economics and Statistics*, Vol. 61, number 1, pp. 57-84.
- Goodin, R., J. M. Rice, A. Parpo, and L. Eriksson (2008) 'Discretionary Time: A New Measure of Freedom', Cambridge: Cambridge University Press.
- Gronau, R. (1977) 'Leisure, Home Production, and Work – the Theory of the Allocation of Time Revisited', *Journal of Political Economy*, Vol. 85, pp. 1099-1124.
- Guarcello, L., S. Lyon, and F. Rosati (2006) 'The Twin Challenges of Child Labour and Youth Employment in Ethiopia', Understanding Children's Work Project Working Paper 18.
- Halvorsen, R., and R. Palmquist (1980) 'The Interpretation of Dummy Variables in Semi-Logarithmic Equations', *American Economic Review*, Vol. 70, pp. 474-475.
- Hart, K. (1973) 'Informal Income Opportunities and Urban Employment in Ghana', *The Journal of Modern African Studies*, Vol. 11, number 1, pp. 61-89.
- Heckman, J. (1979) 'Sample Selection Bias as a Specification Error', *Econometrica*, 47 (1): 153–62.
- Heckman, J. (1978) 'Dummy Endogenous Variables in a Simultaneous Equation System', *Econometrica*, Vol. 46, pp. 931-960.
- Heckman, J., S. Urzua, and E. Vytlacil (2006) 'Understanding Instrumental Variables in Models with Essential Heterogeneity', *The Review of Economics and Statistics*, Vol. 88, pp. 389-432.
- Heckman, J., and E. Vytlacil (2005) 'Structural Equations, Treatment Effects, and Econometric Policy Evaluation', *Econometrica*, Vol. 73, issue 3, pp. 669-738.
- Heckman, J., and E. Vytlacil (2001) 'Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling' *The Review of Economics and Statistics*, MIT Press, Vol. 83, number 1, pp. 1-12.
- Hoddinott, J., and L. Haddad (1995) 'Does Female Income Share Influence Household Expenditures? Evidence from Côte d'Ivoire', *Oxford Bulletin of Economics and Statistics*, Vol. 57, number 1, pp. 77–96.
- Husmanns, R. (2007) 'Measurement of Employment, Unemployment and Underemployment – Current International Standards and Issues in their Application', Bureau of Statistics, Geneva: International Labour Office.
- Husmanns, R. (2004) 'Measuring the Informal Economy: From Employment in the Informal Sector to Informal Employment', Policy Integration Department, Bureau of Statistics, Working Paper No. 53, Geneva: International Labour Office.

- Hyder, A. and B. Reilly (2005). 'The Public Sector Pay Gap in Pakistan: A Quantile Regression Analysis', PRUS Working Paper No. 33.
- ILD – Instituto Libertad y Democracia (2005) 'Program to Formalize the Assets of the Poor of Tanzania and Strengthen the Rule of Law: The Diagnosis', Vol. 1, Lima: ILD.
- ILO – International Labour Organization (2010) 'Decent Work Country Profile Tanzania (mainland)', Dar es Salaam and Geneva: International Labour Organization.
- ILO – International Labour Organization (2002) 'Decent Work and the Informal Economy', Report of the Director-General, International Labour Conference, 90<sup>th</sup> Session, Report VI, Geneva: International Labour Organization.
- Jones, A. M. (1992) 'A Note on Computation of the Double-Hurdle Model with Dependence with an Application to Tobacco Expenditure, *Bulletin of Economic Research*, Vol. 44, number 1, pp. 67-74.
- Juster, F. T., and F. P. Stafford (1991) 'The Allocation of Time: Empirical Findings, Behavioural Models, and Problems of Measurement', *Journal of Economic Literature*, Vol. 29, pp. 471–522.
- Jütting, J.P., and J.R. de Laiglesia (2009) 'Is Informal Normal? Towards More and Better Jobs in Developing Countries', Development Centre Studies, Paris: The OECD.
- Kabubo-Mariara, J. (2003). 'Wage Determination and the Gender Wage Gap in Kenya: Any Evidence of Gender Discrimination?', Research Paper 132, African Economic Research Consortium, Nairobi.
- Kan, M. Y. (2006) "Measuring Housework Participation: The Gap Between 'Stylized' Questionnaire Estimates and Diary-based Estimates", GeNet Working Paper 20, Institute for Social and Economic Research, University of Essex, United Kingdom.
- Kang, J.-H. (2007) 'The Usefulness and Uselessness of the Decomposition of Tobit Coefficients', *Sociological Methods Research*, Vol. 35, number 4, pp. 572-582.
- Kennedy, P. (1981) 'Estimation with Correctly-Interpreted Dummy Variables in Semi-Logarithmic Equations', *American Economic Review*, Vol. 71, p.801.
- Klasen, S. (1999) 'Does Gender Inequality Reduce Growth and Development? Evidence from Cross-Country Regressions', Development Research Group, Poverty Reduction and Economic Management Network, Policy Research Report on Gender and Development, Working Paper Series No. 7, Washington, DC: The World Bank.
- Klasen, S., and F. Lamanna (2008) 'The Impact of Gender Inequality in Education and Employment on Economic Growth in Developing Countries: Updates and Extensions', EUDN Working Paper No.10, Namur, Belgium: European Development Research Network.
- Kolev, A. and P. Suárez Robles (2010a). 'Addressing the Gender Pay Gap in Ethiopia: How Crucial is the Quest for Education Parity?', *Journal of African Economies*, Vol. 19, number 5, pp. 718-767.
- Kolev, A. and P. Suárez Robles (2010b). 'Exploring the Gender Pay Gap through Different Age Cohorts: The Case of Ethiopia', J.S. Arbache, A. Kolev and E. Filipiak (eds), *Gender Disparities in Africa's Labour Markets*, Washington, DC: The World Bank.

- Krishnan, P., T.G. Selassie and S. Dercon (1998) 'The Urban Labour Market During Structural Adjustment: Ethiopia 1990–1997', Centre for the Study of African Economies, Institute of Economics and Statistics, University of Oxford.
- Kuepie, M., C.J. Nordman and F. Roubaud (2009) 'Education and Earnings in Urban West Africa', *Journal of Comparative Economics*, 37 (3): 491–515.
- Lachaud, J.P. (1997) *Les Femmes et le Marché du Travail Urbain en Afrique Subsaharienne*. Editions l'Harmattan, 233pp.
- Lee, L.F. (1983) 'Generalized Econometric Models with selection', *Econometrica*, 51 (2): 507–12.
- Lee, Y. S., and L. J. Waite (2005) "Husbands' and Wives' Time Spent on Housework: A Comparison of Measures", *Journal of Marriage and Family*, Vol. 67, number 2, pp. 328–36.
- Lewis, W.W. (2004) 'The Power of Productivity: Wealth, Poverty and the Threat to Global Stability', University of Chicago Press, Chicago, Ill.
- Lim, L.L. (2002) 'Female Labour Force Participation', Background Paper for the United Nations Population Division, Expert Group Meeting on Completing the Fertility Transition, March.
- Maloney, W.F. (2004) 'Informality Revisited', *World Development*, Vol. 32, number 7, pp. 1159-1178.
- Manser, M., and M. Brown (1980) 'Marriage and Household Decision Making', *International Economics Review*, Vol. 21, pp. 31-44.
- Marini, M. M., and B. A. Shelton (1993) 'Measuring Household Work: Recent Experience in the United States', *Social Science Research*, Vol. 22, number 4, pp. 361–82.
- McDonald, J. F., and R. A. Moffitt (1980) 'The Uses of Tobit Analysis', *The Review of Economics and Statistics*, Vol. 62, number 2, pp. 318-321.
- McElroy, M., and M. J. Horney (1981) 'Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand', *International Economic Review*, Vol. 22, pp. 333-349.
- McFadden, D.L. (1973). 'Conditional Logit Analysis of Qualitative Choice Behaviour', P. Zarembka (ed), *Frontiers in Econometrics*, New York: Academic Press.
- Medeiros, M., R. Guerreiro Osorio, and J. Costa (2007) 'Gender Inequalities in Allocating Time to Paid and Unpaid Work: Evidence from Bolivia', Working Paper 34, International Poverty Centre, Brasilia, Brazil.
- Mengistae, T. (2001) 'Indigenous Ethnicity and Entrepreneurial Success in Africa: Some Evidence from Ethiopia', Policy Research Working Paper Series 2534, World Bank.
- Mincer, J. (1974) 'Schooling, Experience and Earnings', New York: The National Bureau of Economic Research.
- Mincer, J. (1962) 'Labour Force Participation of Married Women', *Aspects of Labour Economics*, Universities-National Bureau Conference Series No. 14, NY: NBER, pp. 63-105.
- Ministry of Finance and Economic Development – MoFED (2005) 'Ethiopia: The Millennium Development Goals (MDGs) Needs Assessment – Synthesis Report', Development Planning and Research Department, Addis Ababa.

- Montenegro, C. (2001) 'Wage Distribution in Chile: Does Gender matter? A Quantile Regression Approach'. Working Paper Series No. 20, Development Research Group, World Bank.
- Morrison, A., D. Raju, and N. Sinha (2007) 'Gender Equality, Poverty and Economic Growth', Poverty Reduction and Economic Management Network, Gender and Development Group, Policy Research Working Paper No. 4349, Washington, DC: The World Bank.
- Moser, C. (1978) 'Informal Sector or Petty Commodity Production: Dualism or Dependence in Urban Development?', *World Development*, Vol. 6, issues 9-10, pp. 1041-1064.
- Nawata, K. (2007) 'A Monte Carlo Analysis of the Type II Tobit Maximum Likelihood Estimator When the True Model is the Type I Tobit Model', *Economics Bulletin*, Vol. 3, number 54, pp.1-10.
- NBS – National Bureau of Statistics (2007) 'Analytical Report for Integrated Labour Force Survey (ILFS), 2006', NBS, Dar es Salaam, Tanzania.
- NBS – National Bureau of Statistics (2002) 'Integrated Labour Force Survey 2000/01 – Analytical Report', NBS, Dar es Salaam, Tanzania.
- Neumark, D. (1988) 'Employers Discriminatory Behaviour and the Estimation of Wage Discrimination', *Journal of Human Resources*, 23: 279–95.
- Newey, W.K. (1999), 'Two Step Series Estimation of Sample Selection Models', Massachusetts Institute of Technology (MIT), Department of Economics, Working Papers:99-04.
- Ñopo, H., N. Daza, and J. Ramos (2011) 'Gender Earnings Gaps in the World', IZA Discussion Paper Series No. 5736.
- Nordman, C.J., A.-S. Robilliard, and F. Roubaud (2011) 'Gender and Ethnic Earnings Gaps in Seven West African Cities', *Labour Economics*, doi:10.1016/j.labeco.2011.09.003
- Nordman, C.J., F. Rakotomanana and A.S. Robilliard (2010) 'Gender disparities in the Malagasy labour market', J.S. Arbache, A. Kolev and E. Filipiak (eds), *Gender Disparities in Africa's Labour Markets*, Washington, DC: The World Bank.
- Nordman, C.J. and F. Roubaud (2005). 'Reassessing the Gender Wage Gap: Does Labour Force Attachment Really Matter? Evidence from Matched Labour Force and Biographical Surveys in Madagascar', Working Papers DT/2005/06, DIAL.
- Nordman, C.J. and F.C. Wolff (2009). 'Gender Differences in Pay in African Manufacturing Firms'. Working Papers hal-00421227\_v1, HAL.
- Nordman, C.J. and F.C.Wolff (2008). 'Islands Through the Glass Ceiling? Evidence of Gender Wage Gaps in Madagascar and Mauritius', Working Papers, DT/2008/02, DIAL.
- Oaxaca, R. (1973) 'Male–Female Wage Differentials in Urban Labour Markets', *International Economic Review*, 14 (3): 693–709.
- Perry, G., W.F. Maloney, O. Arias, P. Fajnzylber, A. Mason, and J. Saavedra-Chanduvi (2007) 'Informality: Exit and Exclusion', Washington, DC: The World Bank.

- Piesse, J., and J. Simister (2003) 'Bargaining and Household Dynamics: The Impact of Education and Financial Control on Nutrition Outcomes in South Africa', *South African Journal of Economics*, Vol. 71, number 1, pp. 163–72.
- Pollak, R., and M. Wachter (1975) 'The Relevance of the Household Production Function and Its Implications for the Allocation of Time', *Journal of Political Economy*, Vol. 68, pp. 349-359.
- Press, J. E., and E. Townsley (1998) "Wives' and Husbands' Housework Reporting: Gender, Class and Social Desirability", *Gender and Society*, Vol. 12, number 2, pp. 188–218.
- Reimers, C.W. (1983) 'Labour Market Discrimination Against Hispanic and Black Men', *The Review of Economics and Statistics*, 65 (4): 570–79.
- Rosenbaum, P. R., and D.B. Rubin (1984) 'Reducing Bias in Observational Studies Using Sub-Classification on the Propensity Score', *Journal of the American Statistical Association*, Vol. 79, pp. 516-524.
- Rosenbaum, P.R., and D.B. Rubin (1983) 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', *Biometrika*, Vol. 70, pp. 41–55.
- Ryder, H. E., Stafford, F. P., and Stephan, P. E. (1976) 'Labour, Leisure and Training over the Life Cycle', *International Economic Review*, Vol. 17, number 3, pp. 651-74.
- Schneider, F. (2004) 'The Size of the Shadow Economies of 145 Countries All over the World: First Results over the Period 1999 to 2003', IZA Discussion Paper 1431, Institute for the Study of Labor, Bonn.
- Siphambe, H.K. and M. Thokweng-Bakwena (2001) 'The Wage Gap between Men and Women in Botswana's Formal Labour Market', *Journal of African Economies*, 10 (2): 127–42.
- Soto, H. de (2000) 'The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else', Basic Books, New York, NY.
- Soto, H. de (1989a) 'The Other Path: The Invisible Revolution in the Third World', Harper & Row, New York, NY.
- Soto, H. de (1989b) 'The Informals Pose an Answer to Marx', *Economic Impact*, number 67, pp. 56-66.
- Suárez Robles, P. (2010) 'Gender Disparities in Time Allocation, Time Poverty, and Labour Allocation across Employment Sectors in Ethiopia', J.S. Arbache, A. Kolev and E. Filipiak (eds), *Gender Disparities in Africa's Labour Markets*, Washington, DC: The World Bank.
- Temesgen, T. (2006) 'Decomposing Gender Wage Differentials in Urban Ethiopia: Evidence from Linked Employer–Employee (LEE) Manufacturing Survey Data', *Global Economic Review*, 35 (1): 43–66.
- Ting, L., and A. Malhotra (2005) 'Disorders of Sleep: An Overview', *Primary Care: Clinics in Office Practice*, Vol. 32, number 2, pp. 305–318.
- Tobin, J. (1958) 'Estimation of Relationships for Limited Dependent Variables', *Econometrica*, Vol. 26, pp.24-36.
- Tsai, S.-L., and Y. Xie (forthcoming) 'Heterogeneity in Returns to College Education: Selection Bias in Contemporary Taiwan', *Social Science Research*.

- UN-Habitat (2004) 'Addis Ababa, Ethiopia: Urban Inequities Report', Global Urban Observatory, Monitoring System Branch, Nairobi, Kenya.
- UNICEF (1999) 'Women in Transition', The MONEE Project Regional Monitoring Report No. 6, ICDC, Florence.
- Utz, R.J. (2008) 'Sustaining and Sharing Economic Growth in Tanzania', Washington, DC: The World Bank.
- Vickery, C. (1977) 'The Time-Poor: A New Look at Poverty' *Journal of Human Resources*, Vol. 12, pp. 27–48.
- Ward, J., B. Lee, S. Baptist, and H. Jackson (2010) 'Evidence for Action – Gender Equality and Economic Growth', London: Chatham House (The Royal Institute of International Affairs) and Vivid Economics.
- Warner, R. L. (1986) 'Alternative Strategies for Measuring Household Division of Labour: A Comparison', *Journal of Family Issues*, Vol. 7, number 2, pp. 179–95.
- WEF – World Economic Forum (2006) 'The Global Gender Gap Report', Geneva: WEF.
- Weichselbaumer, D. and R. Winter-Ebmer (2003) 'A Meta-analysis of the International Gender Wage Gap'. IZA Discussion Paper No. 906.
- Wodon, Q., and K. Beegle (2006) 'Labour shortages despite underemployment? Seasonality in time use in Malawi', C. M. Blackden and Q. Wodon (eds), *Gender, Time Use, and Poverty in Sub-Saharan Africa*, World Bank Working Paper No. 73, Washington, DC: The World Bank.
- Woldehanna, T., B. Tefera, N. Jones, and A. Bayrau (2005) 'Child Labour, Gender Inequality and Rural/Urban Disparities: how can Ethiopia's national development strategies be revised to address negative spill-over impacts on child education and well-being?', Young Lives Working Paper Series No. 20, London: Young Lives.
- World Bank (2011) 'World Development Report 2012 – Gender Equality and Development' Washington, DC: The World Bank.
- World Bank (2009a) 'Ethiopia: Unleashing the Potential of Ethiopian Women – Trends and Options for Economic Developments', Poverty Reduction and Economic Management, Washington, DC: The World Bank.
- World Bank (2009b) 'Youth and Employment in Africa – The Potential, the Problem, the Promise', Africa Development Indicators 2008/09's essay, Washington, DC: The World Bank.
- World Bank (2007) 'Urban Labour Markets in Ethiopia: Challenges and Prospects', Volume I: Synthesis Report, Volume II: Background Papers, Poverty Reduction and Economic Management Unit, Africa Region, Washington, DC: The World Bank.
- World Bank (2006) 'Gender and Transport Resource Guide', Module IV: Gender and Rural Transport Initiative, Washington, DC: The World Bank.



Xie, Y., J. Brand, and B. Jann (2011) 'Estimating Heterogeneous Treatment Effects with Observational Data', Institute for Social Research, Population Studies Center, Research Report 11-729, University of Michigan.

Xie, Y., and X. Wu (2005) 'Market Premium, Social Process, and Statisticism', *American Sociological Review*, Vol. 70, pp. 865-870.

Zhou, X., and Y. Xie (2011) 'Propensity-Score-Based Methods versus MTE-Based Methods in Casual Inference', Institute for Social Research, Population Studies Center, Research Report 11-747, University of Michigan.