HUMAN CAPITAL, INFORMALITY AND LABOUR MARKET OUTCOMES IN SUB-SAHARAN AFRICA

by

ANDREW KERR

A thesis submitted
for the degree of D.Phil. in Economics

ST ANTONY'S COLLEGE
UNIVERSITY OF OXFORD

Trinity Term 2011

Abstract

Human Capital, Informality and Labour Market Outcomes in sub-Saharan Africa.
Submitted in Trinity Term 2011 for the Degree of DPhil in Economics.
Andrew Kerr, St Antony's College.

In this thesis I explore three topics in labour economics, using micro data from South Africa and Tanzania.

South Africa suffers from extremely high income inequality, in part as a result of comprehensive Apartheid-era racial discrimination. The first topic explores possible explanations for the extremely large earnings differences across different types of employment for black South Africans, using the KwaZulu-Natal Income Dynamics Study data. I analyse the relative importance of individual ability and institutions, including public sector wage setting and trade unions, in determining earnings. My results suggest that human capital and individual ability explain much of the earnings differentials within the private sector, including union premiums, but cannot explain the large premiums for public sector workers.

Self-employment is very common in urban Tanzania but, unlike South Africa, survey data show that there are large overlaps in the distribution of earnings in private wage employment and self-employment. This suggests that self-employment represents a viable alternative to wage employment in small, low productivity firms for the majority of urban Tanzanians. In chapter three I build an equilibrium search model of the urban Tanzanian labour market that is inspired by these observations and that explains the choice of wage and self-employment and the variation in earnings across and within these sectors.

In the final topic I explore the effect of education on earnings in Tanzania. Estimating the returns to education has stimulated much recent work in applied econometrics as researchers advance their understanding of the effect of individual heterogeneity on the possibility of estimating the returns to education. In my attempt to purge estimates of the return to education of the influence of individual heterogeneity, I use an education reform in Tanzania as a natural experiment that provides exogenous variation in education. When using Ordinary Least Squares (OLS) I find high and strongly convex, increasing returns to education. My best attempt at separating out the effect of individual heterogeneity suggests that returns are still high but that they may actually be concave.

Acknowledgements

This thesis has only been possible with the support and encouragement of a large number of people. Dr Francis Teal has been a patient and wise supervisor. There is probably no other member of the Oxford economics department whose door is so consistently open to all. My meetings with Dr Teal were usually characterised by dry humour, as well as encouragement and an interest in whichever data I was working with. I thank him for his generosity with his time and approachability.

I thank Professor John Hart at the University of KwaZulu-Natal for encouraging me to take a broader view of the field. I am grateful to Professor Dori Posel for encouraging me to apply for graduate programs in economics, as well as for the solid training she provided as my supervisor and lecturer at UKZN. I am also grateful to them both, as well as Professor Jim Fairburn, Professor John Swart, Peter Habberton and Rod Capon for acting as referees for my application for the Rhodes Scholarship, which funded my first three years in Oxford.

The Centre for the Study of African economies has been an excellent place to conduct my DPhil research. I am grateful for the DPhil scholarship which provided funding for my fourth year in Oxford, as well as to the warm and friendly staff who keep the Centre a constant hive of activity.

Simon Quinn alerted me to the education reform I used in the final chapter of this thesis. He also obtained the Ministry of Education data in Dar es Salaam and allowed me to use it. The opportunity to work on this reform, particularly with Simon in the early stages, stimulated my interest in, and understanding of, some of the more challenging econometrics in this thesis, so I am very grateful to Simon. Fellow DPhil student Paolo Falco (Mr Ghana) was always willing to have a chat about survey data issues or tricky Stata routines and deserves my thanks. Alex Sienaert provided encouragement from the inside and our conversations about the KIDS data were extremely helpful. Rulof Burger read through this thesis and began several useful conversations about Card (2001), for which I am grateful. Members of the Economics Department 'A-side', Bet Caeyers, Kirk Bailey, Matt Collin, Naureen Karachiwalla and Sebastian Königs, have been an excellent source of conversation and fun and provided a welcome relief from some of the Manor Road awkwardness.

Members of the Oxford University Graduate Christian Union and St Andrews Church, Linton Road, provided encouragement and support as well as helping me to think more deeply about my own discipline and broaden my outlook beyond economics. Aaron Moore and Jimmy Copeland have been a source of laughter and adventure. Nishan de Mel and Donald Hay encouraged me to think about the relationship between my faith and economics. Susie and John Salmon made me feel at home in Oxford and were a source of a lot of the good books I read whilst a student there. Mee-yan

Judge and Charles Hampton helped me to see the world differently when I was struggling. I am grateful to them both.

Wellington Street, Jericho, became a home during my years in Oxford and I thank my housemates Daniel Cojocaru, Daniel Zoughbie, Jessica Leight, Joel Harrison, Nate Harper, Penny Grenfell and Peter Eckley for many good conversations, parties, dinners and bike lights discos!

Tim Gibbs, historian and fellow member of St Antony's College, provided an outsider's perspective on life in South Africa that was invaluable for me as I looked at home from afar. His friendship has been a fun and valuable one. Friendship with Nathan Rose was an honour, and I thank him for many helpful conversations, as well as for being instrumental in organising some fantastic trips around the UK. I am already missing our regular lunches at the Alternative Tuckshop. Jared Honeycutt (Sir Roger B) has been a great friend, as well as an adventurous running, cycling and swimming partner. I wish him well in his academic pursuits across the pond.

Denise Gray helped me to see that I was capable of finishing this thesis and has shown me patient love, for which I am deeply grateful. She also read through much of this thesis, correcting my grammar. I hope I am able to encourage her in her doctoral research half as well as she encouraged me in mine. I thank God for giving me the opportunity to undertake this thesis. I dedicate this thesis to my parents Nick and Barbara Kerr, who have consistently supported and encouraged me.

> "The kingdom of God is justice and peace, and joy in the Holy Spirit. Come Lord, and open in us the gates of your kingdom."
> -a song of the Taizé Community, France.

Word Count and Style

The total number of words in this thesis is 52360, excluding the bibliography. Tables and figures are placed at the end of each chapter, first tables and then figures. Citations follow the Harvard referencing system and the style of the American Economic Review. The thesis is typeset in $\LaTeX\,2_\varepsilon$ using double line spacing.

Declaration

Chapter two of this thesis is a substantially reworked version of my MPhil thesis, which I submitted in Trinity Term 2008 as part of the requirements for my MPhil degree. The title of my MPhil thesis was "Dynamics of the South African Labour Market".

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

The labour market plays a crucial role in determining the incomes earned by individuals and households, as most income earned by individuals comes from the income earned from their labour. In developing countries the functioning of the labour market may play an important role in determining whether individuals experience poverty, through either unemployment or low productivity employment. Understanding how labour markets work in developing countries is thus a vital part of any policy strategy seeking to raise incomes in these countries.

In this thesis I analyse the workings of labour markets in two African countries, Tanzania and South Africa. Tanzania is a low income country characterised by high levels of self-employment and low levels of inequality. It has a state bureaucracy with very limited capacity to improve the lives of its citizens. South Africa is an upper middle income country plagued by a level of inequality amongst the highest in the world as well as a very high unemployment rate. It also has a relatively well-financed state bureaucracy that is determined to implement policies that will reverse the damaging effects Apartheid had, and continues to have, on many of its citizens.

In the first chapter of this thesis I explore different explanations for the large earnings differentials across sectors that are observed in survey data from South Africa. In the second chapter I build a model to explain the choice of wage and self employment, as well as earnings determination, in urban Tanzania. In the final chapter I estimate the returns to education in Tanzania. In all three chapters I explore the role of individual heterogeneity in earnings determination in different ways.

The discovery of the existence of pervasive individual heterogeneity has profoundly influenced applied microeconometrics and labour economics over the last fifty years. This discovery has come about as researchers began to collect and use micro data to evaluate the success of the welfare state in influencing the lives of its citizens for the better (Heckman 2001). This research programme has been adopted in developing countries, with researchers attempting to evaluate, as well as suggest, policies that aim to raise incomes and decrease poverty. A central theme of this thesis is that an understanding of the influence of individual heterogeneity on labour market outcomes is crucial if sound conclusions are to be drawn from the available micro data about how labour markets function.

1.2 Chapter Outlines

South Africa suffers from extremely high income inequality, as a result of comprehensive Apartheid-era racial discrimination in educational funding and labour market legislation. Recent estimates of inequality have shown, though, that inter-racial inequality is declining whilst intra-racial inequality is rising (van der Berg and Louw 2004). In the first chapter of this thesis

I explore possible explanations for the extremely large earnings differences across different types of employment for black South Africans, using the KwaZulu-Natal Income Dynamics Study panel. I analyse the importance of individual ability and human capital, relative to labour market earnings institutions, including public sector wage policy and trade unions, in determining earnings.

Self-employment is very common in urban Tanzania, as in many African economies, which is indicative of a low level of development. The data from Tanzania I use in my second chapter show that, unlike South Africa, there are large overlaps in the distribution of earnings in private wage employment and self-employment, as well as substantial variation in earnings within these sectors. This suggests that self-employment represents a viable alternative to wage employment in small, low productivity firms for the majority of urban residents in Tanzania. Inspired by these observations, in the second chapter I build a theoretical model of the urban Tanzanian labour market to explain the choice of wage and self-employment and the distribution of earnings across, and within, these sectors.

In the final chapter I explore the effect of education on earnings in Tanzania. Estimating the returns to education has stimulated much recent work in applied econometrics, as researchers have advanced their understanding of the effect individual heterogeneity has on their ability to estimate the returns to education. In much of the research estimating the returns to education in Africa, however, there has been little acknowledgement of the difficulties this entails. In my attempt to purge estimates of the return to education of the influence of individual heterogeneity, I use an education reform in Tanzania as a natural experiment that provides exogenous variation in education.

1.3 Common themes

Despite the three chapters dealing with different topics, two different countries and using different tools there are some common themes across the chapters.

All of the chapters make use of earnings functions in some form. These are central in chapter two, where I make use of earnings functions to look at earnings differentials across different sectors in the South African labour market, and in chapter four, where I estimate the returns to education in Tanzania. The use of earnings functions dates back to Mincer (1958), who estimated the effects of education on earnings for white men in the USA. Mincer's (1958) contribution, along with Schultz (1958), was to analyse education as a form of capital (human capital), which could be invested in and which generated a return, similar to the way economists had long thought about physical capital. This way of thinking about education has become the norm, both in economics, and in society more generally, and has spawned a huge research agenda which is an important part of this thesis.

Heckman's (2001) Nobel Prize lecture focused on the diversity of economic life that micro data have revealed to those who have collected and analysed these data. In all three topics I explore the role of individual heterogeneity in determining earnings and also the choice of which sector to work in, using tools that have been developed to analyse and understand micro data.

Chapters two and three are both inspired by models of labour market segmentation and research that examines evidence for and against this segmentation. In a segmented labour market it is argued that formal sector institutions prevent the higher wages paid in the formal sector from being bargained down by those who wish to obtain employment in this sector at the going wage. This creates differences in the wages earned in the formal and informal sectors by otherwise identical workers. In chapter two I show that access to public sector employment is indeed an important determinant of earnings in South Africa. But I also argue that the KIDS data suggest that individual heterogeneity does explain much of the vast differentials across different kinds of private sector employment.

In chapter three I argue that what is most interesting about Tanzania is not that there are large average earnings differences between sectors, as is true in South Africa, and as emphasised in the segmentation literature, but rather the overlap of earnings between wage and self-employment. I take a step away from the segmentation literature by building a model in which the choice of wage and self-employment, and the earnings in each of these, are determined by individual productivity differences in self-employment, rather than by labour market institutions, as in a segmented labour market.

A question related to the segmentation question, that is touched on in both of the first two topics, is whether the labour market is competitive. In the second chapter I compare the competitive model with human capital with a segmented labour market model. I argue that my results suggest there are features which imply segmentation between the public and private sectors, but not within the private sector. In chapter three, despite moving away from a segmented labour market model, I do not model the labour market as competitive. Instead, following the equilibrium search literature (Mortensen and Pissarides 1999), I assume frictions limit the ability of workers to match with wage and self-employment opportunities and also give employers monopsony power that allows them to pay wages that are below the marginal product of the employees they hire.

1.4 The Contribution of this Thesis

The major contribution of this thesis is to apply recent developments and insights in econometrics and economic theory to explore earnings determination in two African labour markets, as well as to provide evidence on, and analyse, how these labour markets function. I answer questions about the importance of trade unions, the public sector and human capital in earnings determination, as well as the influence of self employment in a labour market when self-employment is prevalent.

In chapter two I analyse the workings of the South African labour market by using a panel dataset that allows me to explore the role of unobserved heterogeneity in earnings determination. Current research has focused on whether the labour market is segmented into the formal and informal sectors (Badaoui et al. 2008). I challenge the relevance of a simple binary dichotomy between the formal and informal sectors and move beyond the current body of research by exploring the interaction of individual heterogeneity and institutions, including the public sector and trade unions, in determining earnings.

In chapter three I develop what I believe to be the first matching model of an African labour market, incorporating self-employment as an alternative to matching with a firm. In doing so I emphasise the role of individual heterogeneity in earnings determination and the choice of whether to work in wage or self-employment. I also link my model to micro data from Tanzania, and generate earnings variation in both sectors, unlike other recent models of labour markets in developing countries (Albrecht et al. 2009).

In chapter four I estimate the returns to education in Tanzania using nationally representative cross sectional data, which has not been undertaken before. I compare my estimates for the country with earlier estimates

(Söderbom et al. (2006), Knight and Sabot (1990)), linking my findings to recent education policy in Tanzania. In this I embed my work in the treatment effects literature, using this as a basis to understand what it is possible to estimate, when there is heterogeneity in the returns to education, and the assumptions required to do so.

Chapter 2

The Role of Institutions in Determining Sectoral Earnings Differentials: Panel Data Evidence from South Africa

2.1 Introduction

In 2004 black self-employed workers in KwaZulu-Natal province, South Africa, had average hourly earnings six times lower than black employees in private sector firms, who themselves earned half the average hourly wage of black workers in the public sector. In this chapter I seek to explain the vast differences between the average earnings of the self-employed, at one end of the informal/formal divide in South Africa, and unionised public sector workers at the other.

Human capital theory provides one potential explanation for earnings differentials. This theory posits that workers' productivities explain their earnings, and sector choices, and implies that these sector choices are utility maximising, despite the possible existence of sectoral earnings differentials. Extensions and modifications of the Harris and Todaro (1970) model provide an alternative explanation, suggesting that institutional features of the labour market prevent formal sector earnings from equalising the demand and supply of labour in this sector. Different features that are present or absent in particular sectors of the labour market have been argued to generate wage differentials between the sectors for otherwise identical workers, which is described as segmentation. These include union wage bargaining, minimum wage laws and public sector pay policies.

In this chapter I use a panel data set that allows me to explore the role of observed and unobserved heterogeneity in determining earnings outcomes in the South African labour market, as well as the relative contributions of two key formal sector institutions, public sector pay policy and trade unions, to sectoral earnings differentials. As a result of the legacy of Apartheid racial classifications continue to be commonly used in surveys and censuses and in analyses of the labour market. I abstract from racial differentials in labour market outcomes in this chapter by confining my analysis to black South Africans only. This decision mainly reflects data constraints: the panel data set I use only contains information on black and Indian households. Given the large differences in labour market outcomes in South Africa across all racial groups, however, I have chosen to limit my analysis to black workers.

The structure of the chapter is as follows: Section 2.2 reviews evidence on possible explanations of earnings differentials in South Africa. The

KwaZulu-Natal Income Dynamics Study is described in Section 2.3. Section 2.4 tests the ability of both human capital theory and versions of the segmentation hypothesis to explain earnings, and explores the effects of allowing for unobserved heterogeneity to influence earnings. The robustness of the results to corrections for attrition is considered in Section 2.5 and Section 2.6 concludes.

2.2 Identifying Sources of Earnings Differentials

In this section I review explanations for wage differentials between jobs that differ in their level of formality. The segmented labour market hypothesis challenges the assumptions of the neo-classical model of labour supply and a single competitive labour market in which workers are paid their marginal product. Instead, it emphasises institutional factors such as union power, wage setting in the public sector, minimum wage legislation and other forms of regulation, as well as firm behaviour in the private sector as having a significant impact both on an individual's choice of where to work or to search for employment and of the wage paid. It is argued that these institutional factors result in wage differentials across sectors which are inconsistent with the competitive model.

2.2.1 The Competitive Framework with Human Capital

The traditional neo-classical framework emphasises the productivity of individuals as the primary driver of wages in a competitive labour market. In the simplest competitive model there is no distinction between jobs that differ in their degree of informality or between unobserved, productivity-enhancing individual abilities. In this model any wage differentials between sectors or jobs would be competed away through workers moving to sectors where wages are highest and the assumption of decreasing marginal productivity of labour. Individual heterogeneity can easily be incorporated into human capital theory in a number of ways, however, resulting in a richer model with more power to explain outcomes observed in the labour market.

When incorporating human capital investment, workers choose the amount of education and training that maximises their expected lifetime utility, they are paid their marginal product and the market outcome is the efficient level of wages and employment. The work of Becker (1962) and Mincer (1973) underpins much of human capital theory, as well as the exploration of some of its empirical implications. In this literature productivity is assumed to be a function of education, experience in the labour market, and general and specific training. The estimation of earnings functions as a method of testing human capital theory has rapidly become a prominent feature of labour economics (Card 1999). Wages are generally observed to increase with education, training and experience (Mincer 1973), evidence in favour of the basic tenants of human capital theory. I explore human capital investment and the impact of education on earnings in much greater detail in chapter four of this thesis. In this chapter I assume that individuals have one kind of ability that affects productivity in all sectors, in addition to their observable human capital.

In this chapter I focus on earnings differentials across different sectors of employment. What has been called "occupational choice" has been incorporated into modern human capital theory (cf Boskin (1974) and Fleisher (1970)). In this extension individuals choose their education and training levels, as well as their sector of employment, to maximise their lifetime utility. This utility is based on their (possibly unobserved) productive capabilities in each sector and the benefits and costs of education and training required to enter each sector. Workers should thus be observed in occupations where their expected future lifetime utility is highest. In this competitive paradigm any wage differentials between individuals should reflect only observed and unobserved differences in productive capabilities and differences in job characteristics across sectors. All other wage differentials are assumed to be competed away as individuals move to sectors where wages are higher.

An early example of occupational choice predating a human capital approach is the work of Roy (1951). In his two-sector assignment model, Roy (1951) explains sector choice by unobserved differences in individuals' productive capabilities in each sector. Education and training, and the amount of investment in these, are not modeled. In his model workers choose the sector in which they have a comparative advantage. This leads to wage differences between sectors, but these simply reflect differences in unobserved ability between workers in the two sectors. In my empirical investigation I do not allow for more than one type of unobserved ability. I do take this further in the next chapter of this thesis, where I develop an explicit theoretical model of sectoral choice, ability and earnings in the context of the Tanzanian labour market.

2.2.2 Segmented Labour Market Theories

Despite the success in explaining wages of the competitive model that incorporates human capital, some economists have emphasised that the competitive model was not sufficient to explain wage differentials between individuals (Leontaridi 1998). This line of thought has been argued to begin with the American Institutionalist school of thought (Dunlop (1957), Kerr (1954)). It was further developed in the 1960s as social scientists grappled with the new micro data available to them, and argued that similar workers were observed to earn substantially different amounts depending on which sector they worked in (Leontaridi 1998). Instead of explaining these outcomes within the neoclassical paradigm, as occupational choice did, some economists emphasised the importance of labour market institutions that prevented the bargaining down of wages in high wage sectors by workers in low wage sectors and which uncoupled the productivity-wage relationship central to human capital theory. It was argued that institutional constraints led to different wage setting mechanisms in different sectors in the market and to workers with similar skills and experience being paid different wages, depending on which part of the market they were able to access. This was labeled segmentation.

Various explanations were given for the existence of institutional constraints and wage setting mechanisms differing between sectors. In the developing country context the Harris-Todaro model of migration (Harris and Todaro 1970) posited the existence of a minimum wage in the urban sector that was set above the market clearing wage, due to minimum wage legislation, and that was insulated from the forces of supply and demand. The rural labour market was assumed competitive. The wage in the rural agricultural sector and unemployment in the urban sector were determined by the equalising of the wage in the rural sector with the expected wage in the urban sector. The expected wage was the urban sector wage multiplied by the probability of obtaining urban sector employment. The existence of a minimum wage meant that rural workers or the unemployed could not bid the urban wage down to a level that equilibrated actual, as opposed to expected, wages in the rural and urban sector. If a worker managed to find urban sector employment he would thus be paid more than if he worked in the agricultural sector.

A very important development in the analysis of labour markets in developing countries was the recognition of the existence of the informal sector. Attention was first drawn to the informal sector in the 1972 International Labor Organization (ILO) mission to Kenya (International Labour Office 1972) and anthropological work by Hart (1973) in urban Ghana. This highlighted that many urban residents were making a living beyond the influence of minimum wage laws and other employment regulations, both in self-employment and in small, unregistered and often family-owned firms.

These new developments led to much new empirical and theoretical research on the informal sector. An important contribution was that of Fields (1975), who extended the Harris-Todaro framework to include an urban informal sector, which he called the murky sector, along with the urban formal sector and rural agriculture. The murky sector consisted of informal self-employment and casual employment in informal enterprises, and was considered a free entry sector. As a result anyone was assumed to be able to find employment in this sector, although the wage paid was assumed to be even lower than that paid in the agricultural sector. Individuals were willing to work for this low wage in the model because murky sector work allowed them to search for a high wage formal sector job more efficiently than if they worked in rural agriculture. Employment in this sector was assumed to be a temporary experience while an individual searched for a job in the formal sector. Again, the existence of a minimum wage or union activity in the formal sector created sectoral wage differentials, despite the model specifying identical individual productivity. It also implied that formal sector employment was lower than the equilibrium level, and thus productive employment opportunities that would have been created in the absence of fixed formal sector wage, were not created. Fields (1975), along with the

Harris and Todaro (1970) model, can be seen as part of the basis for the subsequent empirical investigation of segmentation in developing country labour markets.

In the wake of the early work on the informal sector there has been a large volume of research on this sector in economics, anthropology and other social sciences. One part of this literature focuses on how informality should be defined. Initially the ILO Kenya report gave several characteristics of informal activities, including "ease of entry", "family ownership of enterprises" and "unregulated and competitive markets" (International Labour Office 1972). The ILO has subsequently given numerous definitions of the informal sector with the most recent definition being expanded to include informal workers in registered enterprises, in addition to workers in unregistered or small enterprises, and those in enterprises that did not register some of their employees (Guha-Khasnobis et al. (2006), Muller (2003), Heintz and Posel (2008)). Informal workers are then defined by the ILO as those individuals who are "in law or in practice, not subject to national labour legislation, income taxation, social protection or entitlement to certain employment benefits" (Hussmans 2004).

In all these views the distinction between formality and informality is a binary one. An employed person is thought of as being either in a formal job or an informal job. Instead of a binary dichotomy Chen (2006) has argued that there is actually a continuum of the formality of jobs, with some being more formal than others. In this chapter I take this approach. Unionised, public sector work in South Africa could be considered highly formalised, unionised regular work in the private sector less so, and self-employment or casual employment even less formal. The segmented labour market hypothesis would suggest that jobs with higher levels of formality would be higher paying than less formal jobs, even after controlling for observed and unobserved ability. I test the relative importance of labour market institutions, compared to individual heterogeneity and human capital in my empirical analysis below.

2.2.3 Explaining Earnings Differentials in South Africa

In a competitive market, human capital theory predicts that earnings are determined by individuals' productive characteristics such as their education and labour market experience. There is much evidence in South Africa that human capital is an important determinant of earnings. The returns to education from ordinary least squares earnings regressions have been shown to be strongly convex. Tertiary education is associated with marginal returns of over 50 percent a year when using OLS and age earnings profiles suggest large returns to labour market experience (Keswell and Poswell 2004). Kingdon and Knight (2004) also show that higher levels of education increase expected wages through decreasing the probability of unemployment. There are also several studies that hint at the importance of unobserved human capital for earnings. Badaoui et al. (2008) show that unobserved ability partly explains a formal sector premium for employees, whilst Cornwell and Inder (2008) show that the ability to speak English (not captured in the KIDS data I use below) is a key determinant of earnings differentials in cross-sectional data.

There are many studies, however, showing that human capital is not the only determinant of earnings. The fundamental assumption in the Fields (1975) extension of the Harris and Todaro (1970) model is that the wage in the urban formal sector is fixed at a level higher than the market clearing wage. Without this there would be no segmentation, as the wage would be bargained down by the unemployed and those in the informal sector until wages in the urban formal and informal sectors were equalised. At the aggregate level, Casale et al. (2004) show that the broad unemployment rate in South Africa rose from 29 percent to 43 percent between 1995 and 2003, but that between 1997 and 2003 real wages for employees in registered businesses, one of the possible definitions of the formal sector, have been roughly constant. Interestingly, these authors also show that the earnings of those employed in unregistered businesses and the profits of owners of unregistered businesses declined by 36 percent over the same period, whilst employment in this category increased by more than 2.7 times[1].

These results point to a degree of inflexibility in parts of the South African labour market, with a number of institutional features that could explain this inflexibility being suggested by different authors. A key part of the South African labour market legislation is that which pertains to Bargaining Councils. These can be created by agreement of trade unions representing 50 percent of workers in any industry and firms employing 50 percent of workers in the same industry in a particular region. Bargaining Councils have the power to set minimum wages in the industry, even for firms in the industry who are not part of the Bargaining Council (Bhorat et al. 2007). Magruder (2009) shows that industries with Bargaining Council agreements have wages between 10-21 percent higher than industries in neighbouring areas without these agreements, and argues that the Bargaining Council system results in lower employment and a smaller average firm size of firms in regions and industries covered by Bargaining Council agreements.

Moll (1996) explores the impact of Industrial Councils, the forerunners

[1]The authors note that some of this increase in employment is the result of better capturing of informal employment in more recent surveys.

of Bargaining Councils, and provides some descriptive evidence that these created wage premia even for non-union workers covered by an Industrial Council agreement during the early 1990s. Butcher and Rouse (2001) investigate this effect more rigorously and find a small but significant premium using the 1995 October Household Survey. More recent work using the 2005 Labour Force Survey (Bhorat et al. 2007) has shown the public sector Bargaining Council is associated with large wage premia. This accords with work by Heintz and Posel (2008), who show that public sector workers earn 40 percent more than formal private sector workers after controlling for a range of observable characteristics. Thus there is some evidence that Bargaining Councils contribute to raising wages and lowering employment in South Africa, and that this effect is particularly large within the public sector.

Trade unions are often claimed to be one of the institutional factors that prevents the formal sector wage from clearing the labour market. This is argued to occur through their role in firm-level bargaining, in addition to their role in facilitating Bargaining Council agreements. Union premiums of 60 percent in the 1993 PSLSD data (rising to 145 percent for the lowest decile of black workers), documented by Schultz and Mwabu (1998), suggest that unionisation is an important factor in wage determination. Butcher and Rouse (2001) argue, however, that these large premiums are exaggerated as a result of not controlling for industry of employment and that this reduces the premium to around 20 percent, which is comparable with premiums in OECD countries. Their conclusion is that unions are not responsible for raising wages or for contributing to South Africa's extremely high unemployment rate.

Several studies specifically set out to investigate whether the South

African labour market is segmented by exploring whether formal sector employment, usually defined as those working in or owning registered enterprises, is associated with a substantial earnings premium. Heintz and Posel (2008) find large premia for formal sector workers using cross-sectional data from the 2004 Labour Force Survey. Cichello et al. (2005) show that obtaining formal employment is a key determinant of earnings growth between the first two waves of the KIDS data. There is thus some evidence for segmentation in these studies.

Results from cross-sectional data that suggest the existence of formal sector earnings premia may, however, be explained by individual differences in unobserved ability or unobserved school quality. The only study that takes unobserved heterogeneity into account in South Africa is Badaoui et al. (2008). The Labour Force Survey panel is used to show that there is no formal sector premium for black males once unobservable differences between individuals are taken into account, where the definition of formality used by the authors is whether the employee is employed in a registered enterprise.

Several important points should be made in relating the work of Badaoui et al. (2008) to my undertakings in this chapter. Firstly, these authors define a very narrow field of investigation. The authors look only at private employees, leaving out public sector workers and the self-employed. The authors thus only attempt to explore whether there is an earnings premium for employees in registered enterprises, compared to unregistered enterprises. In finding no formal sector premium, Badaoui et al. (2008) say nothing about *how* this result is obtained or how the institutions or regulations present in the South African context may have contributed to this situation. In comparison, my work in this chapter attempts to explore the effects of different institutions on wage setting both *within* formal employment and in less formal employment.

Secondly, though Badaoui et al. (2008) make use of the benefits of panel data in controlling for individual effects, they do not address the additional problems panel data bring, specifically attrition bias and attenuation bias caused by measurement error (Deaton 1997). My work allows for important distinctions within formal employment *and* it addresses the possible role of unobserved heterogeneity in wage outcomes, in addition to concerns about measurement error and attrition in panel data, using the KwaZulu-Natal Income Dynamics Survey, which I discuss in the next section.

2.3 Description of the KIDS Data and Survey Methodology

This section provides a brief description of the KwaZulu-Natal Income Dynamics Study (KIDS), a three-wave panel conducted in 1993, 1998 and 2004 in KwaZulu-Natal province, which was the largest province by population when the survey was undertaken[2] [3]. The 1993 wave was part of the first nationally representative household survey in South Africa: the Project for Statistics on Living Standards and Development (PSLSD). A decision was then taken to resurvey only the households in KwaZulu-Natal to create the KIDS panel data set (May et al. 2000). Although the PSLSD surveyed

[2]Subsequently Gauteng has become the most populous province, largely as a result of in-migration.

[3]The KwaZulu-Natal Income Dynamics Study (KIDS) was a collaborative project between researchers at the University of KwaZulu-Natal, the University of Wisconsin, London School of Hygiene and Tropical Medicine, International Food Policy Research Institute (IFPRI), the Norwegian Institute of Urban and Regional Studies and the South African Department of Social Development. In addition to support from these institutions, the following organizations provided financial support: Department for International Development - South Africa (DFID-SA); the United States Agency for International Development (USAID); the Mellon Foundation; and National Research Foundation/Norwegian Research Council grant to the University of KwaZulu-Natal.

all four main racial groups within South Africa (black, coloured, Indian and white), given the small number of white and coloured households surveyed in KwaZulu-Natal, a decision was taken to exclude coloured and white households from future survey waves (May et al. 2007). Given the fundamental differences in education and labour market outcomes for different racial groups and the small sample of Indian households, I have further confined my analysis to black households only.

The system of migrant labour that developed during Apartheid meant that individuals retained links with a rural home but spend most of the year in the urban area in which they worked. This system resulted from restrictions on setttlement in urban areas, particularly on migrants' family members (Posel and Casale 2006). The household was the primary unit of analysis in the PSLSD and a broad concept of the household was used to capture all those associated with the household in the household roster, including migrants. In section 2.A.1 in the appendix to this chapter, I set out how the questionnaire defined household membership and how households were tracked in the 1998 and 2004 follow-up surveys.

2.3.1 Describing Employment, Earnings and Education

Formal and informal employment have been distinguished in a variety of ways, as I discussed above. Statistics South Africa uses the enterprise criterion to distinguish the formal and informal sectors (Muller 2003), so that if the business the individual works for or owns is registered with a government agency or is paying VAT, then the individual is considered to be in the formal sector (Heintz and Posel 2008).

The questions asked in KIDS do not allow for the categorisation of jobs according to the ILO or Statistics South Africa definitions[4]. The KIDS data do allow me to explore the effects of public sector wage setting and unionisation on earnings, however, which can be argued to be two important labour market institutions. I thus describe and explain the differences in earnings between six mutually exclusive employment types that can be identified in the KIDS: four types of regular employment (private non-union, private union, public union and public non-union), as well as casual employment and self-employment.

The structure of employment in South Africa reflects a major difference between it and most other African countries: self-employment in informal enterprises and other types of informal employment make up only a small part of total employment (Kingdon and Knight 2004). This difference is evident in Table 2.1 for the first two waves of the KIDS data, where regular employment predominates. By 2004, however, casual employment had become the most common form in this survey, with a dramatic decline in regular employment, possibly reflecting non-random attrition from the panel. Most of the decline in regular employment is in private, non-union regular employment, and is concentrated amongst those classified as labourers (not shown). Table 2.2 shows that labour force participation was roughly constant across the three waves of the survey, whilst employment is estimated to have decreased by about ten percent between 1993 and 2004, resulting in an increase in the measured broad unemployment rate[5].

Table 2.1 also shows the large hourly earnings differences between regular, casual and self-employment that I highlighted at the start of this chapter

[4]Further details of the relevant labour market questions asked in the KIDS are provided in section 2.A.1 in the appendix to this chapter.

[5]The number of employed and the number of earners in the sample differs because some individuals did not provide earnings. I have not imputed earnings where these were missing.

and which demand an explanation. The hours worked variable is constructed using a question on the number of hours worked in an average day for those in casual and regular employment. The earnings variable is constructed from net earnings for those who reported paying tax and gross earnings for those who reported not having paid tax. For those in self-employment the questions were about the total hours and profits in the business, and households were allowed to list up to three members involved in the business, but there was no question asking the amount of time each member worked in the business. There is thus some discretion required as to how to allocate the time and profit from these businesses to each individual. I have assumed an equal share of time and profit for each individual listed as helping with the business.

The median wage for public sector workers in unions was 17 times higher than earnings in self-employment in 2004, a gap which had increased threefold since 1993. This increase in the earnings gap over the eleven year period covered by the surveys was generated by a decline in earnings by a third for the self-employed and an increase of roughly 65 percent for unionised public sector employees. There were substantial increases in real wages between 1993 and 1998, as noted in Cichello et al. (2005), with the largest percentage increases for public sector workers. This may be a selection effect due to attrition. It may also reflect the increase in the wages of black individuals in the public sector of the old KwaZulu homeland to the level of their white counterparts as a result of the new democratic government that took power in 1994, or the effects of the formation of the Public Service Co-ordinating Bargaining Council, which occurred in 1997 (Bhorat et al. 2007). Table 2.1 also shows that between 1998 and 2004 real wages remained constant or decreased slightly. Interestingly, Table 2.3 shows that, on average, those working in casual or self-employment have similar weekly working hours to those in regular employment. Thus despite working similar hours, self-employed and casual workers earned significantly less per month than employees with a regular job, and thus also had significantly lower hourly wages.

The expansion of educational opportunities for all South Africans since the end of Apartheid was one of the South African government's major challenges. Significant educational expansion between 1993 and 2004 is evident in the KIDS data, as shown in Table 2.3. Public sector employees have the highest levels of education across the three waves, and those in self-employment have the lowest. Large earnings differences across the sectors seem to be correlated with large differences in educational attainment, which is consistent with human capital theory. This will be given further attention in the subsequent analysis.

There are indications from the descriptive statistics I have presented that non-random attrition is a concern. The KIDS panel attempted to track households that moved, and refreshed the sample by surveying households of the children of core members[6]. The large attrition rate is problematic and may be the reason some of the changes in employment and earnings over the three waves do not accord with those from nationally representative cross-sectional surveys. The attrition rate is similar to other comparable surveys in developing countries (May et al. 2007), however, and techniques do exist to correct for some types of attrition, which I employ in Section 2.5. I now describe the attrition in KIDS in greater detail.

[6]Core members are defined in the Appendix to this chapter.

2.3.2 Attrition

In any panel survey attrition is a potential cause for concern, as it can generate attrition bias in parameter value estimates. One can employ methods to correct for this potential source of bias, however. The first two waves of KIDS were five years apart and the second and third waves were six years apart, meaning that higher levels of attrition could be expected than in panels with survey waves at shorter intervals. The survey methodology did, however, involve attempting to track those households that moved. Entire households exited the sample in waves subsequent to 1993 in four main ways. The first is that some households refused to be re-interviewed. This is generally less of a concern in developing countries, and occurred in only four black households in 1998. The figures for 2004 are not given in May et al. (2007) and are not available in the publicly released survey data. The second possibility was that the survey team found evidence that a household had moved, but not enough detail about where the household had moved to, to allow tracking of the household. The third possibility was that the survey team could find no-one in the community in which the household was living who recognised the names of the members on the household roster from the previous wave. The fourth possibility was that the survey team ascertained that all the core members had died by the next wave, meaning that no-one from the household was interviewed. Some of those households of which the survey team found no trace would probably also fall into this fourth category.

In addition to the attrition of households from the sample, it was also possible for individuals to attrite from households that were re-interviewed. This happened, for example, when an individual listed on the household roster in a previous round was described as not living more than 15 days out of the last month in the household. No further questions were asked about such individuals and they are considered to have exited.

Of the 1139 black households interviewed in 1993, 964 were re-interviewed in 1998. This figure differs slightly from the one given in May et al. (2000) due to the later discovery of the fabrication of 39 household interviews in 1993 and 1998, reported in May et al. (2007). Since some households had split, a total of 1000 households containing core members were interviewed in 1998, and of these 752 were successfully re-interviewed in 2004. Again, because of split, next generation and foster households, the total number of completed questionnaires in 2004 was 1211, including 185 foster households and 268 next generation households. Thus at least one core member of the 1132 households first interviewed in 1993 was successfully re-interviewed in 721 households in 2004, representing household attrition of 36 percent over eleven years.

As noted in May et al. (2007), this underestimates the individual attrition in the sample, which is higher because of the additional attrition of individuals that belonged to households that were re-interviewed. Of the 8258 individuals in households interviewed in 1993, 971 exited from the sample because they belonged to households that were not re-interviewed, and another 953 exited from households that were re-interviewed in 1998. Of the 7786 individuals that were listed as members of households that were interviewed in 1998, 1278 exited because they belonged to one of the 248 households that were not re-interviewed in 2004. Another 1410 individuals exited from households that were re-interviewed. This implies that of the 8258 individuals interviewed in 1993, 5397 were re-interviewed in 2004. Whilst these are large rates of attrition, May et al. (2007, pg. 638) show that they are comparable to attrition rates in other developing countries such as the LSMS Cote d'Ivoire panel survey conducted in the late 1980s and the Lima LSMS from Peru and argue that KIDS "continues in most of its important demographic characteristics to remain broadly representative of the population of the province."

Table 2.4 show some correlates of household attrition. In both 1998 and 2004 it seems that smaller households are more likely to exit the sample. Whether the household lived in a rural area or whether the household's dwelling was owned by a member does not, however, seem to be correlated with attrition. Households with higher per capita income seem more likely to exit in 2004, but not in 1998. Table 2.5 shows some correlates of individual attrition (all individuals older than 15, including those who exit the sample due to household attrition). Individuals that exit are older, more likely to be male, come from larger households and have slightly higher levels of education. Whether the individual is a core member or not is an important predictor of attrition in both 1998 and 2004: as a result of the survey methodology, core household members were tracked if they moved, whilst other household members were not. Table 2.5 shows the same characteristics for the sub-sample of earners, which I use in the next section. In this sample men and those who are married seem more likely to exit. Younger earners also seem more likely to exit the sample.

Having given a description of attrition in the KIDS I turn to explaining wage differentials. I return to attrition in section 2.5, when I explore the effects of attrition on the results I obtain in the following sections.

2.4 Explaining Wage Differentials in South Africa

In this section I explore the extent to which the competitive model, within a human capital framework, can explain the large average wage differentials across employment of different levels of formality documented above, estimating earnings functions using Ordinary Least Squares (OLS). Previous research suggests that controlling for unobserved heterogeneity, including unobserved comparative advantage, will be an important step forward in analysing these earnings differentials.

2.4.1 Can Human Capital Theory Account for Earnings Differentials?

I begin testing explanations for the large wage differentials shown above by exploring to what extent they are due to observable human capital, estimating earnings functions using OLS. Table 2.7 shows basic OLS earnings functions for the three waves of the panel, with the log of real hourly wage being the dependent variable[7]. Missing earnings or hours data, occurring because individuals refused to answer or the respondent did not know, are not imputed. The education variables used are dummies indicating whether the individual had some primary, completed primary, some secondary, matric (completed secondary) or some tertiary education. The omitted category is no education.

Across all three waves the results show that there are large returns to education and labour market experience. The strongly convex returns to education are consistent with other findings for South Africa (Keswell and Poswell 2004) and for other African countries (Bennell 1996), and contradicts the assertion of Psacharopoulos (1994) that returns to education are generally concave in developing countries. The age-earnings profile, which is a proxy for rewards to labour market experience, suggest that 20 extra

[7]The interpretation of the regression coefficients is the percentage change in the hourly wage for a unit change in a regressor. For dummy variables the percentage effect of a change from zero to one is calculated by $e^{\beta} - 1$. Table 2.6 in the Appendix to this chapter shows basic summary statistics for the variables included in the regressions.

years of labour market experience would result in a 40-year-old gaining a 55 percent increase in earnings compared to a 20-year-old, holding other variables constant. The negative but insignificant married dummy in 1998 is not consistent with the finding that married individuals earn higher wages on average (Casale and Posel 2010).

The vast raw wage differences across various dimensions of formality, noted in section 2.3, are substantially reduced in the OLS earnings functions for each of the three waves, conditioning on a range of covariates. In the pooled data the six-fold raw difference in average earnings between the highest and lowest paying sectors is reduced to roughly a four-fold difference when controlling for observable human capital. Interestingly, most of this occurs through reduced differentials within the private sector, with the large raw public sector premium hardly declining. It would seem that observable human capital can explain some of the earnings differentials within the private sector, but not the differentials between the public and private sectors. The union premium in the private sector is estimated to be around 65 percent, but drops to roughly 30 percent when I control for industry and occupation, shown in the last column of Table 2.7. This is consistent with the analysis of Butcher and Rouse (2001) that I discussed in section 2.2[8].

2.4.2 Modeling Unobserved Heterogeneity

OLS estimation fails to allow for factors such as unobserved ability and preferences for different job characteristics. Panel estimators can be used to control for unobserved individual time invariant heterogeneity. It is then possible to estimate the change in earnings as an individual moves between

[8]With the KIDS data I am only able to explore the effect of union membership and not on the broader role unions play in raising minimum wages for all workers through the bargaining council system, which I discussed in Section 2.2.

union and non-union employment, or into or out of the public sector. There is a significant amount of movement between sectors, as shown in Tables 2.8 and 2.9, which is required for identification of differences in earnings across different levels of formality. In particular, anticipating my results below, there is extensive movement out of, and into, private unionised employment.

I use fixed effects and first and second difference estimators in my estimation. The fixed effects model is

$$y_{it} = \alpha_i + x_{it}^{'}\beta + \varepsilon_{it}. \tag{2.1}$$

The α_i terms are the individual specific effects. x_{it} and β are K x 1 vectors, and I assume the ε_{it} are iid [0, σ^2]. The fixed effects estimator is obtained by subtracting the time-averaged model from the original model, giving

$$y_{it} - \bar{y_i} = (x_{it} - \bar{x_i})'\beta + (\varepsilon_{it} - \bar{\varepsilon_i}). \tag{2.2}$$

Wooldridge (2002) shows that this is a consistent and unbiased estimator if either N$\rightarrow \infty$ or T $\rightarrow \infty$ and assuming strict exogeneity of the ε_{it} (E$[\varepsilon_{it}|x_{it}, \alpha_i] = 0$). Alternatively, one can estimate the fixed effects model using the first difference estimator, subtracting the model lagged one period, to obtain

$$y_{it} - y_{i,t-1} = (x_{it} - x_{i,t-1})'\beta + (\varepsilon_{it} - \varepsilon_{i,t-1}). \tag{2.3}$$

This is also unbiased and consistent under the assumptions above, but is less efficient than the fixed effects estimator when ε_{it} are serially uncorrelated. Longer differencing periods can reduce the effects of attenuation bias due to measurement error and help determine how serious the bias is in the first differenced results (Griliches and Hausman 1986). I explore this further below.

2.4.3 The effects of unobserved heterogeneity

Table 2.10 reports results from the fixed effects estimator, as well as the first and second difference estimators. Second differencing is the maximum difference allowed by the data, since KIDS is a three wave panel. The first column shows the results of the fixed effects estimator. They suggest that a significant premium remains for those in public sector employment, relative to private regular employment. Those in unionised, public sector employment earn around 90 percent more than those in private, regular, non-union employment, controlling for time invariant unobservables and time varying observables. Private sector, regular employment is more lucrative than casual or self-employment, with those in casual employment earning 27 percent less, and those in self-employment 9 percent less than those in regular, private, non-union employment, although the self-employment coefficient is not significantly different from zero at the 5 percent level.

There is also a statistically significant[9] two-and-a-half-fold difference between the highest and lowest paying sectors. The fixed effects results suggest that controlling for unobserved heterogeneity further reduces the differentials within the private sector. The public sector premium, however, is hardly reduced at all.

The results from the first difference estimator are similar to those from the fixed effects estimator. Wooldridge (2002) notes that this suggests the strict exogeneity assumption is correct. The first differenced results also indicate that a significant premium remains for those in public sector employment, relative to private, regular employment, and that the private sector

[9]Table 2.11 contains the results of pairwise F tests for the equality of the coefficients on the different types of employment for the fixed effects regression.

union premium is much lower than in the OLS regressions, where it was estimated to be around 65 percent, and is not significantly different from zero.

The small union premium in the private sector is an important finding, as it suggests that the larger union premium in the cross-sectional results may actually be a quality effect, and that union members earn higher wages than non-union members in private employment because they are (unobservably) more productive. In this interpretation, unions are playing a positive role, allowing higher quality workers to capture returns to their human capital. It is also possible, however, that firms are responding to unionisation and a concomitant rise in wage rates, by increasing their capital stock, lowering their labour usage and hiring the higher quality workers necessary to utilise the increased capital. This is a decidedly less positive interpretation, but unfortunately it is not possible to distinguish between these explanations using the KIDS data.

I noted above that there is substantial movement into and out of union membership, which should allow for the identification of the union effect in both the fixed effects and differenced estimators. This requirement of movement means that panel estimators rely on those observed more than once to identify the regression coefficients. Those only observed once are excluded. As a simple robustness check, it is helpful to determine whether the pooled cross-sectional results also hold for the sub-sample of individuals observed more than once. The last column in Table 2.12 shows the key OLS results are indeed similar in the reduced sample, and does not suggest that the fixed effects or first differenced estimates are a result of the smaller sample of earners observed more than once. I now explore some other potential sources of bias in my estimates.

2.4.4 Measurement Error and Panel Data

Measurement error is of concern in any data set, and this can be more problematic in panel data (Deaton 1997). A well-known result from econometric theory is that measurement error in the independent variable results in attenuation bias, meaning that coefficients of the independent variables are biased towards zero. Following Cameron and Trivedi (2005, pg. 905), the effects of measurement error in panel data can be illustrated using a model with a scalar regressor:

$$y_{it} = \alpha_i + x^*_{it}\beta + \mu_{it}. \tag{2.4}$$

α_i is the individual unobserved fixed effect, x^* is measured with error and I observe x, with $x_{it} = x^*_{it} + v_{it}$. If I use a first difference estimator then

$$\Delta y_{it} = \beta\Delta x^*_{it} + \Delta\mu_{it} \tag{2.5}$$

$$= \beta\Delta x_{it} + \Delta\mu_{it} - \beta\Delta v_{it}. \tag{2.6}$$

If I define ρ=Cor$(x^*_{it}, x^*_{i,t-1})$, then it can be shown that

$$plim\hat{\beta} = \beta + (plim\frac{1}{N}\sum_{i=1}^{N}\Delta x_{it}{}^2)^{-1} plim\frac{1}{N}\sum_{i=1}^{N}(\Delta x_{it}\Delta\mu_{it} - \beta\Delta x_{it}v_{it}) \tag{2.7}$$

$$= \beta - \frac{\beta\sigma^2_v}{(1-\rho)\sigma^2_{x^*} + \sigma^2_v}. \tag{2.8}$$

This inconsistency is larger than in the cross-sectional case when $\rho > 0$. It is also clear from equation (2.8) that as $\rho \to 1$ the inconsistency becomes large. KIDS has longer periods between waves than many other panel surveys, with five years between the first two waves and six years between the second and third waves. This would help to alleviate attenuation bias, if the independent variable changes by a larger amount compared to panels with shorter periods between waves, as this would lower the value of ρ.

I noted above the large decline in the union premium when using the fixed effects and first difference estimators and gave two possible economic explanations. This could also be attributed to measurement error in sector of employment, however, and the resultant attenuation bias. As a further robustness check, longer differencing periods can reduce the effects of attenuation bias and help to determine how serious the bias is in the first difference results (Griliches and Hausman 1986). The KIDS is a three-wave panel, and hence the maximum differencing possible allows for a second difference estimator. If measurement error was driving my results then I would expect to see larger coefficients in the second difference estimator than in the first difference. Table 2.10 shows the sector coefficients in the second difference results are generally not significant, most likely as a result of the small sample, but that they are not universally larger than the fixed effects and first difference coefficients. In fact, the private sector union premium is *lower* in the second difference regression than the first difference regression, and is in fact negative, implying that my low estimate of the premium is unlikely to be the result of measurement error[10].

2.4.5 Asymmetry of Sectoral Movement Effects

When using panel estimators, the earnings change of moving into sector A from sector B and of moving into sector B from sector A are constrained to be the same. Movement between sectors is assumed to be exogenous and

[10]Butcher and Rouse (2001) show that ignoring controls for industry *inflates* the union premium. My estimates do not control for industry and can thus be thought of as the *maximum* estimate of the union premium, assuming I have controlled for other potential biases.

not to be determined by individual characteristics. There may, however, be asymmetric effects resulting from movement that is not exogenous. I focus here on the large public sector earnings premium I found even after controlling for unobserved individual time invariant heterogeneity. The first column in Table 2.13 uses a first difference estimator to estimate earnings differentials but excludes those who move out of the public sector. These results are similar to the full sample, shown in column two of Table 2.10. Column two shows the results excluding those who enter the public sector, which are radically different, with no significant public sector premium[11]. Columns three, four and five repeat the analysis combining unionised and non-unionised public sector workers, and the same result obtains.

These dramatic results suggest a simple informal model of public sector employment. One could assume two broad types of individuals in the population, low and high ability types, who both have some chance of gaining entry to public sector employment, perhaps because of poor screening of applicants or nepotism. If public sector pay is not determined by ability and if being fired from the public sector is very difficult then low ability workers may manage to enter public employment, knowing that finding better paying employment in the private sector is not possible. As a result low ability workers would stay in the public sector for as long as possible, earning a premium while they did so. High ability workers' earnings could be matched or exceeded if they received a job offer in the private sector, meaning they are the only type who leave public sector employment. This simple model would explain the lack of a drop in earnings for those exiting the public sector, as well as the large earnings premium for those who enter the public sector from private employment.

[11] A similar analysis for the union effect in the private sector produces no asymmetric effect.

2.4.6 The effects of the tax system

Up until this point I have focused on individuals' net earnings. The KIDS asked all those in regular employment to report both net and gross pay (the self-employed and those classified as casual workers both in the survey and in this chapter were only asked about gross pay or profit). Tables 2.14 and 2.15 show OLS, and fixed effects and differenced estimates respectively, using gross pay for those who were asked about it and reported it. The results are similar to those using net pay in Tables 2.7 and 2.10, though there are now larger differences between self and casual employment and the other categories where gross and net pay differ. The union premium in the private sector is much larger and significant, however, when using gross pay in the regressions that control for unobserved heterogeneity. This suggests that unions increase the cost to private firms more than they increase returns for workers.

2.4.7 Further Robustness Checks

As a final set of robustness checks I report OLS and fixed effects regressions dividing the sample by gender. Table 2.16 shows that the key results when separating the sample into men and women are preserved. The ranking of sectors is roughly the same for the full sample, although women are seen to be much worse off in self-employment compared to men. The union premium in the private sector is again dramatically lowered when estimated using the fixed effects panel estimator for both men and women.

2.5 Addressing Attrition

Non-random attrition may be a source of bias when estimating earnings differentials using panel data. In this section I model attrition at an individual level and then use two different attrition correction methods to correct the results from the pooled cross section and the difference estimators. The pooled cross section can be corrected using Inverse Probability Weighting (IPW) and the difference estimator regressions can be corrected using a Heckman selection correction, following Wooldridge (2002).

2.5.1 Modeling Attrition

I model attrition as a binary variable s_{it} that takes a value of one if the individual exits the sample at time t, after appearing in the sample at time $t-1$, and is zero otherwise. s_{it} is a function of individual, household and community characteristics that are represented by the vector w_{it}:

$$s_{it} = 1[w_{it}\delta + v_{it} > 0]. \tag{2.9}$$

Tables 2.18 and 2.19 in the Appendix to this chapter[12] show the results of three probit models of individual attrition. The first column uses 1993 characteristics of individuals themselves, as well as characteristics of their households and communities, to predict whether individuals exit the sample in 1998, the second wave of the panel. The second uses the same characteristics to predict attrition in 2004. The third column uses characteristics of all those present in 1998 (including new household members not present in 1993) to predict attrition in 2004.

The results in the first column suggest that both elderly household mem-

[12] The second table reports the last half of the regression results from the first table due to the large number of explanatory variables.

bers and the very young, the omitted category, are least likely to exit in 1998. Attrition is also found to be less likely for resident household members and for males. Those in private sector regular unionised employment were also less likely to exit the sample. No-one in any other category of employment or the unemployed were significantly more likely to exit compared to the omitted category, which was those not in the labour force. Core members are also significantly less likely to exit, as the descriptive analysis in Section 2.3 also showed, and is indicative of the tracking methodology by which households that split were tracked if these contained core members. Individuals in larger households were less likely to exit, possibly reflecting lower mobility and broader ties to the community of larger households (Maluccio 2000). Per capita income was not a significant predictor of household attrition, as being in any of the highest four income quintiles in 1993 did not make one more likely to have attrited by 1998 than individuals in the lowest quintile. Individuals whose households were located in the former homeland of KwaZulu were significantly less likely to exit than those who lived in the former province of Natal.

The results from the third column, modelling individual attrition in 2004 using 1998 individual, household and community characteristics, are similar to those using 1993 characteristics to predict attrition in 1998. Again, younger members are less likely to exit, as are resident members, core members and individuals in larger households. The income level of the individual's household does not seem to influence the likelihood of attrition, and neither does the individual's level of education. Unlike attrition between the first two waves, the third column indicates that males are more likely to exit than females. Community level variables that might be thought to influence attrition, for example high levels of reported HIV/AIDS or violence, seem not to affect the probability of an individual exiting the sample.

Despite this, community dummies (not shown) are found to be strongly significant in many instances across all three models of attrition, indicating possible unobserved community level shocks that influenced large numbers of households to move from some communities. Some examples of this are listed in Maluccio (2000), who describes two communities located on private farms in the former province of Natal, which both seem to have had high attrition rates, in one case as a result of the farm going bankrupt. The community dummies are introduced to control for these types of unobserved community-wide shocks or characteristics.

The second column of Tables 2.18 and 2.19 shows a model of attrition by 2004 on the characteristics of those observed in 1993. Notable differences are that those in the highest four household per capita income quintiles are more likely to attrite compared to the bottom quintile, with the higher income quintiles associated with a higher probability of attrition. Individual education levels are positively associated with attrition. This may suggest that the process generating attrition is linked to movement associated with income opportunities over the longer term. Again, the very young and old are least likely to exit the sample.

2.5.2 Correcting for Possible Attrition Bias

In Section 2.4 individual earnings were modeled as

$$y_{it} = \alpha_i + x_{it}^{'}\beta + \varepsilon_{it}. \tag{2.10}$$

In correcting for possible attrition bias in the pooled cross section Inverse Probability Weighting relies on the so-called "ignorability" assumption (Wooldridge

2002):

$$P(s_{it} = 1|y_{it}, x_{it}, w_{i1}) = P(s_{it} = 1|w_{i1}). \tag{2.11}$$

Inverse Probability Weighting requires using the predicted probabilities of exit from the sample from a model of attrition, for example the attrition probit estimated above, and then weighting observations so that those that have a higher predicted probability of attriting, but nevertheless stay in the sample, are given a higher weight. These results are shown in the first column of Table 2.12, and are not dramatically different from the uncorrected results. An optimistic interpretation of the results is that attrition is not a serious concern. It could also indicate, however, that the ignorability assumption has been violated, given the fairly low predictive power of the variables in the attrition model.

In correcting for possible attrition bias in the differenced estimation, I follow Wooldridge (2002) and assume joint normality of $\triangle\varepsilon_{it}$ and ν_{it}, which are the errors in the attrition and differenced earnings functions respectively, strict exogeneity of x_{it} and that selection does not depend on Δx_{it} once w_{it} has been controlled for. It can then be shown that

$$E(\Delta y_{it}|\Delta x_{it}, w_{it}, s_{it} = 1) = \beta\Delta x_{it} + \rho\lambda(w_{it}\delta), \tag{2.12}$$

where $\lambda(w_{it}\delta)$ is the inverse mills ratio from the attrition probit.

Table 2.17 shows the first and second differenced equations corrected for attrition. The results again do not differ dramatically to the uncorrected results. The key element of formality is again seen to be access to public sector employment, with union membership in the private sector not associated with a significant premium compared to non-union workers.

A final robustness check on the results that correct for unobserved heterogeneity is to use the first two waves and then the last two waves. This is shown in the final two columns of Table 2.17. The results are similar to the uncorrected and attrition corrected results.

2.6 Conclusion

In this chapter I have shown, using the KIDS data, that average wage differentials for black South Africans across sectors of differing levels of formality are vast. Observable human capital explains about 35 percent of the variance of earnings, and substantial differences in earnings across different types of employment remain after controlling for this human capital. There is a sizeable earnings premium for more formal employment, along the public sector and unionisation dimensions of formality, when using OLS. Public sector workers earn the most, with those in unions at an added advantage. Those in private sector, regular employment are next in the ranking, with union members again earning a premium. Those in casual and self-employment are shown to be at a significant disadvantage compared to those in regular employment.

The crucial question that I am able to answer using the panel dimension of KIDS is whether the large earnings differentials between jobs that differ across various dimensions of formality are the result of unobservable differences between individuals. Once I use panel estimators, to control for unobservable time invariant heterogeneity, the differential between the highest and lowest paid sectors is reduced to a two-and-a-half-fold difference. The only substantial earnings premium found is for those in public sector employment, the most formal of the employment categories I consider. Interestingly, I find that this premium is driven by those entering the public sector, rather than those leaving it, and I outlined a very simple model of ability driving movement into and out of the public sector to explain this result. I also find that the union premium in the private sector is reduced substantially, suggesting that either unobservably higher quality individuals select into union membership, or that firms in which unions are active attempt to mitigate higher wages by selecting employees more carefully. I am unable to distinguish between these explanations with the KIDS data, although I have shown that union membership raises the cost of an employee to a firm more than it raises the net benefit to the employee. Attrition and measurement error can bias results when using panel data but I have shown my results are robust to these concerns.

My results suggest that the explanation for the large earnings differences across different types of employment in South Africa is not found only in human capital or only in institutions that create a segmented labour market. Human capital differentials do provide an important part of the explanation for earnings differentials, particularly within the private sector. I have also shown, however, that access to employment in the public sector creates a substantial earnings premium, even after unobserved heterogeneity is accounted for. Trade unions are often claimed to be an institution that raise wages and contribute to unemployment in South Africa. My conclusions in this chapter are more equivocal, however, and I cannot rule out that trade unions may play a more positive role in the labour market.

Table 2.1: Median Hourly Earnings

	1993		**1998**		**2004**	
	Median	**N**	**Median**	**N**	**Median**	**N**
Employment Categories						
Public Union	13.90	61	24.35	116	23.39	138
	(9.73)		(56.12)		(77.80)	
Public non-Union	10.77	100	18.27	96	17.96	41
	(14.14)		(42.16)		(49.89)	
Private Union	9.94	168	12.15	152	11.44	66
	(10.90)		(32.89)		(38.19)	
Private non-Union	4.04	459	6.55	300	8.59	126
	(15.62)		(16.65)		(28.53)	
Casual Employment	2.19	91	5.16	167	3.70	299
	(4.58)		(18.70)		(8.73)	
Self Employment	2.32	156	3.44	88	1.02	66
	(29.11)		(33.57)		(5.26)	
Frequency	1081		961		789	

Standard Deviations in Parentheses. Earnings expressed in constant 2004 Rand.
Average Rand/ US Dollar exchange rate was 3.26 in 1993, 5.52 in 1998 and 6.46 in 2004
(International Monetary Fund 2009)

Table 2.2: Labour Force Participation and Unemployment

	1993	**1998**	**2004**
Employed	1171	1065	1028
Labour Force Participants	1990	2118	1968
Unemployment Rate	41%	50%	48%

Source: Own calculations from KIDS.

Table 2.3: Mean Weekly Hours Worked and Educational Attainment (Years)

	1993		1998		2004	
	Hours	**Educ**	**Hours**	**Educ**	**Hours**	**Educ**
Employment Categories						
Public Union	43.90	10.36	37.36	11.2	38.62	12.4
	(10.97)	(3.98)	(16.97)	(2.94)	(16.16)	(3.32)
Public non-Union	44.96	9.79	38.23	9.3	36.76	9.5
	(14.83)	(4.13)	(18.44)	(4.30)	(18.49)	(4.51)
Private Union	47.24	7.51	43.15	8.38	45.36	8.34
	(11.10)	(3.36)	(13.28)	(3.27)	(15.19)	(3.96)
Private non-Union	45.43	5.88	42.56	6.69	42.91	8.5
	(15.57)	(3.85)	(14.89)	(3.95)	(18.76)	(3.76)
Casual Employment	38.18	5.25	34.74	7.57	40.78	7.97
	(17.35)	(4.04)	(19.35)	(3.82)	(20.20)	(4.04)
Self Employment	35.46	5.62	48.16	6.16	39.05	8.28
	(25.28)	(3.79)	(27.39)	(3.87)	(24.01)	(3.79)
Frequency	1081		961		789	

Standard Deviations in Parentheses. Source: own calculations from KIDS.

Table 2.4: Household Attrition Descriptive Statistics

	1993		1998		2004
	Resurveyed	**Attrited**	**Resurveyed**	**Attrited**	**Full Sample**
Correlates					
HH Size	7.56	5.55	8.65	5.15	6.96
	(4.17)	(3.46)	(4.54)	(3.22)	(4.07)
Per cap Income	178.25	186.27	291.88	523.64	459.14
	(241.87)	(254.85)	(516.24)	(719.70)	(1280.90)
Rural	0.73	0.81	0.77	0.65	0.77
	(0.44)	(0.39)	(0.42)	(0.48)	(0.42)
Own house	0.84	0.77	0.92	0.69	0.90
	(0.36)	(0.42)	(0.27)	(0.47)	(0.30)
Frequency	964	175	752	248	1211

Standard Deviations in Parentheses. Source: own calculations from KIDS.

Table 2.5: Individual Attrition Descriptive Statistics

	1993 Resurveyed	1993 Attrited	1998 Resurveyed	1998 Attrited	2004 Full Sample
Full Sample					
Age	23.31	22.50	26.53	24.34	24.26
	(18.94)	(16.14)	(19.30)	(16.45)	(18.68)
Male	0.47	0.47	0.46	0.49	0.47
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Married	0.24	0.25	0.19	0.17	0.15
	(0.43)	(0.43)	(0.39)	(0.38)	(0.36)
Year of Educ	4.62	5.14	5.53	6.17	6.38
	(4.06)	(4.17)	(4.25)	(4.31)	(4.42)
Frequency	6334	1924	6507	2688	8318
Earner Sample					
Age	37.80	34.15	39.89	37.57	38.13
	(11.97)	(11.69)	(11.40)	(10.69)	(11.16)
Male	0.52	0.48	0.47	0.55	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Married	0.60	0.57	0.49	0.53	0.42
	(0.49)	(0.50)	(0.50)	(0.50)	(0.49)
Year of Educ	6.54	6.62	7.54	8.10	8.93
	(4.18)	(4.03)	(4.09)	(3.93)	(4.25)
Frequency	737	239	526	354	750

Standard Deviations in Parentheses. Source: own calculations from KIDS.

Table 2.6: Earner Sample Summary Statistics

	1993	**1998**	**2004**
Age	36.90	38.96	38.13
	(12.00)	(11.17)	(11.16)
Male	0.51	0.50	0.49
	(0.50)	(0.50)	(0.50)
Married	0.60	0.50	0.42
	(0.49)	(0.50)	(0.49)
No Educ	0.16	0.09	0.07
	(0.37)	(0.29)	(0.25)
Some Primary	0.30	0.25	0.19
	(0.46)	(0.43)	(0.39)
Completed Primary	0.10	0.09	0.07
	(0.30)	(0.29)	(0.26)
Some Secondary	0.31	0.33	0.34
	(0.46)	(0.47)	(0.47)
Completed Secondary	0.09	0.19	0.20
	(0.28)	(0.39)	(0.40)
Frequency	976	880	750

Standard Deviations in Parentheses. Source: own calculations from KIDS.

Table 2.7: OLS EARNINGS FUNCTIONS

	93	98	04	pooled	IO
	(1)	(2)	(3)	(4)	(5)
Age	0.064	0.039	0.065	0.062	0.056
	(0.018)***	(0.023)*	(0.019)***	(0.01)***	(0.01)***
Age^2	-.0006	-.0003	-.0007	-.0006	-.0006
	(0.0002)***	(0.0003)	(0.0002)***	(0.0001)***	(0.0001)***
Incomplete prim educ	0.26	0.527	0.222	0.302	0.227
	(0.091)***	(0.166)***	(0.193)	(0.085)***	(0.073)***
Complete prim educ	0.442	0.791	0.382	0.504	0.4
	(0.117)***	(0.178)***	(0.258)	(0.107)***	(0.095)***
Incomplete sec educ	0.691	1.073	0.558	0.738	0.547
	(0.104)***	(0.18)***	(0.211)***	(0.104)***	(0.076)***
Complete sec educ	0.884	1.422	0.822	1.040	0.73
	(0.18)***	(0.216)***	(0.219)***	(0.134)***	(0.108)***
Tertiary educ	1.515	1.731	1.429	1.537	1.063
	(0.154)***	(0.255)***	(0.253)***	(0.127)***	(0.116)***
Married	0.178	-.131	0.113	0.061	0.053
	(0.075)**	(0.085)	(0.09)	(0.047)	(0.047)
Male	0.439	0.316	0.124	0.299	0.252
	(0.075)***	(0.069)***	(0.076)	(0.042)***	(0.039)***
Public Union	0.615	0.977	0.947	0.896	0.69
	(0.16)***	(0.116)***	(0.142)***	(0.09)***	(0.081)***
Public non-Union	0.444	0.808	0.869	0.664	0.486
	(0.111)***	(0.144)***	(0.177)***	(0.083)***	(0.069)***
Private Union	0.517	0.504	0.456	0.506	0.278
	(0.109)***	(0.134)***	(0.141)***	(0.088)***	(0.067)***
Casual Employment	-.452	-.147	-.325	-.283	-.246
	(0.13)***	(0.13)	(0.093)***	(0.065)***	(0.062)***
Self employment	-.449	-.144	-.849	-.463	-.375
	(0.159)***	(0.193)	(0.177)***	(0.099)***	(0.098)***
Const.	-.605	-.080	-.141	-.537	-.353
	(0.338)*	(0.521)	(0.449)	(0.227)**	(0.222)
Obs.	976	880	750	2606	2606
R^2	0.347	0.36	0.428	0.381	0.421

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. All variables are dummies except age variables. Reference categories are no education and private non-union employment. The IO column adds industry and occupation controls to the pooled data.

Table 2.8: Earner Movement 93-98

	regprivatenonunion	regpubunion	regpubnonunion	regprivateunion	casearner	selfearner	Total
regprivatenonunion	101	10	6	38	29	12	196
regpubunion	3	12	8	2	2	1	28
regpubnonunion	8	19	12	6	7	3	55
regprivateunion	24	15	7	31	14	7	98
casearner	7	1	6	3	11	4	32
selfearner	8	4	1	5	10	38	66
Total	151	61	40	85	73	65	475

An individual's 1993 sector is on the left and 1998 sector is on the top. Source: own calculations from KIDS.

Table 2.9: Earner Movement 98-04

	regprivatenonunion	regpubunion	regpubnonunion	regprivateunion	casearner	selfearner	Total
regprivatenonunion	20	6	1	10	40	6	83
regpubunion	3	30	8	5	2	1	49
regpubnonunion	3	15	1	3	6	0	28
regprivateunion	8	4	2	11	10	6	41
casearner	7	1	3	4	18	2	35
selfearner	1	2	1	0	13	30	47
Total	42	58	16	33	89	45	283

An individual's 1998 sector is on the left and 2004 sector is on the top. Source: Own calculations from KIDS.

Table 2.10: Fixed Effects and Differenced earnings functions

	Fixed Effects	First Difference	Second Difference
	(1)	(2)	(3)
Age	0.083	0.064	-.369
	(0.042)**	(0.043)	(0.32)
Age^2	-.001	-.0009	0.003
	(0.0005)**	(0.0005)*	(0.003)
Public Union	0.654	0.488	0.391
	(0.167)***	(0.157)***	(0.359)
Public non-Union	0.484	0.352	0.753
	(0.16)***	(0.166)**	(0.401)*
Private Union	0.169	0.125	-.033
	(0.095)*	(0.099)	(0.233)
Casual Employment	-.207	-.271	-.407
	(0.116)*	(0.122)**	(0.268)
Self employment	-.108	-.129	-.274
	(0.166)	(0.179)	(0.269)
1998 Year Dummy	0.411	0.547	
	(0.104)***	(0.109)***	
2004 Year Dummy	0.289		
	(0.194)		
Const.	0.184	-.098	-.793
	(0.9)	(0.137)	(0.264)***
Obs.	1154	619	125
R^2	0.109	0.089	0.103

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Reference employment category is private non-union employment.

Table 2.11: F tests of equality of sectoral coefficients from Fixed Effects Regression

Labour Market Status	regpubNU	regprivU	casual	selfemp
regpubU	0.99	8.58	22.65	13.23
	(0.32)	(0.003)	(0.000)	(0.000)
regpubNU		4.09	14.05	7.75
		(0.04)	(0.000)	(0.005)
regprivU			9.05	2.81
			(0.002)	(0.09)
casual				0.16
				(0.686)

Notes: p values are in parenthesis. NU indicates non-union and U indicates union.

Table 2.12: IPW and OLS earnings functions

	IPW	Stayers
	(1)	(2)
Age	0.073	0.067
	(0.016)***	(0.019)***
Age^2	-.0008	-.0007
	(0.0002)***	(0.0002)***
Incomplete prim educ	0.306	0.374
	(0.078)***	(0.096)***
Complete prim educ	0.56	0.611
	(0.105)***	(0.123)***
Incomplete sec educ	0.677	0.811
	(0.08)***	(0.098)***
Complete sec educ	1.015	1.196
	(0.102)***	(0.125)***
Tertiary educ	1.474	1.471
	(0.158)***	(0.159)***
Married	0.092	0.043
	(0.057)	(0.066)
Male	0.241	0.285
	(0.05)***	(0.061)***
Public Union	0.881	0.927
	(0.103)***	(0.114)***
Public non-Union	0.732	0.743
	(0.091)***	(0.111)***
Private Union	0.523	0.473
	(0.064)***	(0.087)***
Casual Employment	-.336	-.224
	(0.07)***	(0.088)**
Self employment	-.439	-.138
	(0.114)***	(0.094)
Const.	-.674	-.624
	(0.299)**	(0.388)
Obs.	2229	1131
R^2	0.372	0.353

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. The first column corrects for attrition in OLS using IPW. The second shows the OLS results for the sub-sample of earners who are observed in more than one wave.

Table 2.13: First Differenced earnings functions excluding movers into, and out of, Public Sector Employment

	Exclout	Exclinto	Pub	PubExclout	PubExclinto
	(1)	(2)	(3)	(4)	(5)
△ age	0.045	0.044	0.048	0.046	0.043
	(0.042)	(0.044)	(0.042)	(0.042)	(0.044)
△ agesq	-.0007	-.0005	-.0007	-.0007	-.0005
	(0.0004)*	(0.0004)	(0.0004)*	(0.0004)*	(0.0004)
△ pubunionD	0.423	0.155			
	(0.183)**	(0.209)			
△ pubnonunionD	0.462	0.039			
	(0.195)**	(0.196)			
△ publicD			0.32	0.335	0.031
			(0.122)***	(0.157)**	(0.174)
△ privunionD	0.107	0.144			
	(0.116)	(0.12)			
△ casearner1	-.296	-.170	-.304	-.342	-.215
	(0.119)**	(0.119)	(0.11)***	(0.113)***	(0.113)*
△ selfearner	-.155	-.181	-.177	-.182	-.204
	(0.159)	(0.171)	(0.156)	(0.158)	(0.169)
cons	-.019	-.222	-.084	-.002	-.202
	(0.147)	(0.148)	(0.139)	(0.146)	(0.148)
Obs	585	570	628	585	570
R^2	0.076	0.062	0.079	0.072	0.059

Notes: *, **, *** denote significance at 10%, 5% and 1% levels.

Table 2.14: OLS gross hourly earnings

	93	98	04	Pooled
	(1)	(2)	(3)	(4)
Incomplete prim educ	0.396	0.534	0.256	0.381
	(0.143)***	(0.18)***	(0.2)	(0.108)***
Complete prim educ	0.53	0.789	0.361	0.535
	(0.168)***	(0.186)***	(0.266)	(0.121)***
Incomplete sec educ	0.836	1.088	0.62	0.824
	(0.142)***	(0.18)***	(0.225)***	(0.12)***
Complete sec educ	1.072	1.526	0.819	1.149
	(0.228)***	(0.227)***	(0.24)***	(0.163)***
Tertiary educ	1.720	1.840	1.472	1.659
	(0.18)***	(0.265)***	(0.314)***	(0.159)***
Married	0.268	-.067	0.113	0.113
	(0.095)***	(0.094)	(0.092)	(0.05)**
Male	0.422	0.32	0.221	0.323
	(0.087)***	(0.074)***	(0.086)**	(0.05)***
Age	0.07	0.073	0.042	0.071
	(0.023)***	(0.031)**	(0.023)*	(0.015)***
Age^2	-.0007	-.0008	-.0004	-.0008
	(0.0003)***	(0.0004)**	(0.0003)	(0.0002)***
Public Union	0.776	0.982	1.280	1.084
	(0.179)***	(0.131)***	(0.156)***	(0.099)***
Public non-Union	0.652	0.92	0.818	0.799
	(0.136)***	(0.151)***	(0.226)***	(0.098)***
Private Union	0.602	0.56	0.462	0.579
	(0.114)***	(0.136)***	(0.141)***	(0.094)***
Casual Employment	-.359	-.317	-.512	-.389
	(0.146)**	(0.146)**	(0.085)***	(0.081)***
Self employment	-.429	-.318	-1.083	-.541
	(0.172)**	(0.197)	(0.187)***	(0.11)***
Const.	-.878	-.586	0.362	-.779
	(0.429)**	(0.625)	(0.478)	(0.313)**
Obs.	916	821	632	2369
R^2	0.332	0.41	0.501	0.418

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Reference categories are no education and private non-union employment.

Table 2.15: FIXED EFFECTS AND DIFFERENCED EARNINGS FUNCTIONS USING GROSS EARNINGS

	Fixed Effects	First Difference	Second Difference
	(1)	(2)	(3)
Age	0.107	0.112	-.391
	(0.047)**	(0.04)***	(0.371)
Age^2	-.001	-.001	0.003
	(0.0005)**	(0.0006)**	(0.004)
Public Union	0.769	0.613	0.659
	(0.138)***	(0.158)***	(0.365)*
Public non-Union	0.513	0.334	0.969
	(0.161)***	(0.189)*	(0.494)**
Private Union	0.307	0.319	0.56
	(0.106)***	(0.161)**	(0.33)*
Casual Employment	-.212	-.236	-.173
	(0.15)	(0.176)	(0.279)
Self employment	-.158	-.128	0.214
	(0.179)	(0.206)	(0.503)
1998 Year Dummy	0.479	0.853	
	(0.127)***	(0.179)***	
2004 Year Dummy	0.238		
	(0.2)		
Const.	-.337	-.238	-1.200
	(0.949)	(0.183)	(0.322)***
Obs.	1035	550	102
R^2	0.141	0.123	0.106

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Reference employment category is private non-union regular employment.

Table 2.16: Male and female regressions

	OLS Female	OLS Male	Fixed Effects Female	Fixed Effects Male
	(1)	(2)	(3)	(4)
Age	0.053	0.069	0.073	0.138
	(0.015)***	(0.015)***	(0.043)*	(0.042)***
Age^2	-.0005	-.0007	-.0005	-.002
	(0.0002)***	(0.0002)***	(0.0005)	(0.0005)***
Incomplete prim educ	0.142	0.465		
	(0.095)	(0.096)***		
Complete prim educ	0.482	0.525		
	(0.12)***	(0.121)***		
Incomplete sec educ	0.708	0.775		
	(0.094)***	(0.097)***		
Complete sec educ	0.994	1.078		
	(0.114)***	(0.115)***		
Tertiary educ	1.567	1.399		
	(0.142)***	(0.165)***		
Married	0.007	0.069		
	(0.058)	(0.068)		
Public Union	0.971	0.831	0.631	0.676
	(0.118)***	(0.1)***	(0.258)**	(0.214)***
Public non-Union	0.75	0.557	0.523	0.419
	(0.112)***	(0.106)***	(0.227)**	(0.225)*
Private Union	0.411	0.536	0.104	0.16
	(0.098)***	(0.08)***	(0.183)	(0.152)
Casual Employment	-.244	-.353	-.328	-.083
	(0.077)***	(0.08)***	(0.155)**	(0.168)
Self employment	-.528	-.339	-.047	-.206
	(0.08)***	(0.094)***	(0.242)	(0.194)
Const.	-.301	-.431	-.372	-.792
	(0.291)	(0.292)	(0.879)	(0.853)
Obs.	1300	1306	546	623
R^2	0.42	0.328	0.093	0.077

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Year dummies are included but not reported. Reference employment category is private non-union regular employment.

Table 2.17: Attrition Corrected Differenced earnings functions

	First Diff	Second Diff	FirstDiff 93 98	FirstDiff 98 04
	(1)	(2)	(3)	(4)
Age	0.055	-.434	0.05	-.010
	(0.029)*	(0.318)	(0.062)	(0.044)
Age ^{2}	-.0007	0.004	-.0009	0.00005
	(0.0004)**	(0.003)	(0.0005)*	(0.0005)
Public Union	0.438	0.473	0.392	0.641
	(0.158)***	(0.343)	(0.181)**	(0.242)***
Public non-Union	0.294	1.067	0.133	0.778
	(0.169)*	(0.437)**	(0.207)	(0.257)***
Private Union	0.097	0.028	0.088	0.251
	(0.13)	(0.235)	(0.138)	(0.166)
Casual	-.265	-.281	-.138	-.345
	(0.133)**	(0.314)	(0.211)	(0.131)***
Self employed	-.176	-.421	-.147	-.121
	(0.187)	(0.243)*	(0.228)	(0.319)
1998 Year Dummy	0.619			
	(0.158)***			
lambda9898	-.205			
	(0.106)*			
lambda0404	0.328			
	(0.092)***			
lambda93930404		-.390		
		(0.6)		
Obs.	594	110	413	215
R^2	0.098	0.15	0.03	0.088

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. All variables are dummies except age variables. Reference employment category is private non-union regular employment.

Table 2.18: Individual Attrition Probit regressions

	Attrition 93-98	Attrition 98-2004	Attrition 93-2004
	(1)	(2)	(3)
Age 13-19	0.116	0.324	0.114
	(0.059)*	(0.059)***	(0.052)**
Age 22-30	0.305	0.282	0.07
	(0.068)***	(0.067)***	(0.057)
Age 31-50	0.141	0.231	0.014
	(0.08)*	(0.072)***	(0.069)
Age 51-64	0.104	0.172	-.099
	(0.102)	(0.113)	(0.092)
Age 65+	-.178	0.057	-.290
	(0.1)*	(0.099)	(0.09)***
Male	-.061	0.044	0.025
	(0.032)*	(0.032)	(0.03)
Married	0.203	0.126	0.08
	(0.066)***	(0.071)*	(0.055)
Resident Dummy	-.177	-.255	-.088
	(0.069)**	(0.062)***	(0.063)
Employed	0.089	-.032	0.05
	(0.083)	(0.085)	(0.076)
Public Union	-.322	0.077	-.326
	(0.238)	(0.157)	(0.199)
Public non-Union	-.225	0.336	0.003
	(0.189)	(0.169)**	(0.157)
Private Union	-.334	0.137	-.077
	(0.156)**	(0.141)	(0.131)
Private non-Union	-.013	0.17	0.01
	(0.111)	(0.114)	(0.102)
Casual Employment	0.003	0.194	0.152
	(0.14)	(0.134)	(0.132)
Self employment	-.187	0.067	-.045
	(0.13)	(0.137)	(0.106)
Incomplete prim educ	-.051	-.107	0.135
	(0.046)	(0.046)**	(0.039)***
Complete prim educ	-.100	-.027	0.191
	(0.078)	(0.073)	(0.067)***
Incomplete sec educ	0.012	0.006	0.159
	(0.062)	(0.064)	(0.055)***
Complete sec educ	-.037	0.021	0.194
	(0.09)	(0.084)	(0.082)**
Tertiary educ	-.036	0.068	0.308
	(0.189)	(0.178)	(0.163)*

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Marginal effects are reported. Regression is continued in Table 2.19.

Table 2.19: Individual Attrition Probit regressions continued

	Attrition 93-98	Attrition 98-2004	Attrition 93-2004
	(1)	(2)	(3)
Core hh member	-.478	-.563	-.406
	(0.072)***	(0.079)***	(0.066)***
Log(residents)	-.301	-.456	-.399
	(0.077)***	(0.059)***	(0.034)***
Log(non-residents)	-.019	-.075	-.044
	(0.073)	(0.059)	(0.031)
Age of hh Head	0.003		0.0005
	(0.003)		(0.001)
HH head is female	0.011	0.025	-.018
	(0.079)	(0.07)	(0.035)
Max educ level in hh	-.002	-.006	-.031
	(0.015)	(0.018)	(0.007)***
HH has flush toilet	0.323	-.234	0.097
	(0.221)	(0.238)	(0.118)
2nd income quintile	0.116	-.097	0.142
	(0.108)	(0.094)	(0.044)***
3rd income quintile	0.082	-.067	0.108
	(0.116)	(0.102)	(0.047)**
4th income quintile	0.095	-.046	0.117
	(0.121)	(0.105)	(0.052)**
5th income quintile	0.191	-.090	0.193
	(0.136)	(0.132)	(0.065)***
Community violence Dummy	0.102		0.199
	(0.127)		(0.056)***
Rural area in 93	0.609	0.356	0.003
	(0.266)**	(0.472)	(0.178)
Tarred road Dummy	0.189		-.211
	(0.246)		(0.176)
Old KwaZulu area	-.635	0.029	-.547
	(0.238)***	(0.608)	(0.255)**
HH owns its house	0.002		-.073
	(0.153)		(0.064)
Obs.	8258	7654	8258

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Marginal effects are reported. The omitted category is a single, female non-resident, non-labour force participant with zero years of education and who is not a core member in a household in the lowest income quintile, not owning the house the household lives in, without a flush toilet in a community with no reported violence, no tarred roads and that was located in an urban area in the old Natal province in the original 1993 survey.

2.A Chapter Appendix

2.A.1 KIDS questionnaire

Defining the Household and Household Members in KIDS

The KIDS household roster included questions on age, gender, education and relationships with other household members. To be considered part of the household the criteria were that an individual

1. live under this 'roof' or within the same compound/homestead/stand at least 15 days out of the past year,

2. shares food from a common source when members eat together (i.e. they cook and eat together)

3. contributes to or share in, a common resource pool.

For the rest of the survey the questions were further limited to those individuals who had lived "under this roof for more than 15 days of the last 30 days," in addition to criteria 2. and 3. above, to avoid any possible double counting.

The 1998 KIDS follow up survey attempted to track households, including those that moved between waves and those that split up, interviewing all members of any households that contained core members of the 1993 households. Core members were those that satisfied any of 4 criteria: (i) a self-declared head of household (from 1993)

(ii) resident spouse/partner of self-declared head of household (from 1993)

(iii) resident member in a three-generation household (from 1993) and all of the following were true:

child, child-in-law, or niece/nephew of self-declared head;

at least 30 years old;

have at least one child resident in the household

(iv) resident spouse/partner of individual satisfying criterion (iii).

A similar procedure was followed in the third wave in 2004 to track core members in households that had split. One important difference, however, was that in 2004 the panel was refreshed by also tracking next generation households (which did not happen in 1998). Next generation households were defined as those containing adult children of core members, who themselves have children. A second important difference was that children of core members, less than 18 years of age, who were being taken care of in other households were also tracked and interviewed. Household questionnaires for these foster households were less detailed, and so these households are excluded from my analysis.

Labour Market Module

The labour market section in all three surveys asked about employment status for all household members. Further questions were asked to all household members about regular, casual and various forms of self-employment. Regular employment was defined using the question "Does X have a regular job for which he/she earned a salary in the past week?" Self employed professionals (lawyers, doctors etc) were included in this category, but less than 5 were captured in each survey. Casual employment was asked about using the question "Did X do any casual or temporary work in the past month for which he/she was paid in some way, for example, by being given money or food?" Finally self-employment was asked about using the question "Apart from the work that I have already talked about, did anyone in this household do any other kind of work for an income in the past month? I am going to read from a list of possibilities. As I read each one, please indicate whether or not it was a source of income for the household in the past month." A substantial list of possibilities was included, such as shop keeper, selling things on the street, house building or driving a taxi.

Chapter 3

A Model of Comparative Advantage with Matching in the Urban Tanzanian Labour Market

3.1 Introduction

African labour market analysis has often emphasised average wage differentials across sectors. A high-paying "protected" or formal sector that is usually assumed to comprise unionised jobs in large firms and public sector employment is assumed to exist together with a low paying "sink", informal or "murky" sector comprised of self-employment and wage employment in unregistered small firms. A labour market with these characteristics has been described as "segmented", and much effort has been exerted in testing whether labour markets in a variety of countries are indeed segmented. In this chapter I suggest an alternative approach to modeling the Tanzanian labour market and focus on the distinction between wage and self-employment and the role of comparative advantage and individual heterogeneity in determining where to work and earnings in each of these sectors.

In this chapter I argue that describing the Tanzanian labour market as segmented does not adequately capture some important empirical features of this market. In Tanzania there are very large overlaps in the earnings distributions for jobs traditionally considered to be in the formal and informal sectors. There is also a large variance in the distribution of earnings *within* what has traditionally been considered the informal sector. These facts require an explanation not to be found in segmented labour market theories.

Economic theory suggests that a focus on the differences between self and wage employment may be important for understanding the workings of the labour market in low income countries like Tanzania with large amounts of self-employment. In the recent search theory literature both ex-ante wage posting and ex-post wage bargaining models suggest workers may earn less than their marginal product due to the monopsony power of firms. Self-employment, however, which is very common in urban Tanzania, does not involve interaction with a firm and thus the self-employed can capture all the returns to their labour and the return to any capital they have accumulated. This is an important distinction which has not been captured in models of African labour markets to date.

I use these theoretical and empirical insights to develop a matching model of the urban labour market in Tanzania. In this model the productivity of some individuals in wage employment is low relative to their self-employment productivity. Variation in self-employment productivity generates an opportunity cost of wage employment that is increasing in self-employment productivity, thus incorporating insights about comparative advantage in different sectors that date back to Roy (1951).

Section 3.2 reviews the literature on African labour markets, other developing countries and the recent literature on the role of sorting, comparative advantage and heterogeneity, which has been explored in the developed country context as well as in Latin America. Section 3.3 describes the data I use in this chapter. Section 3.4 explores the data and how they speak to different models of the Tanzanian labour market. In section 3.5 I develop a model which captures some of the features presented in the empirical exploration. Section 3.6 explores how the numerical solution to the model matches the empirical evidence and section 3.7 concludes.

3.2 Modeling African Labour Markets

Much of the recent work on African labour markets attempts to provide evidence on whether or not labour markets are segmented, inspired by Fields's (1975) extension of the Harris and Todaro (1970) model of migration. In the previous chapter I noted that public sector wage setting is an important institution that generates earnings differentials in South Africa. But I also argued that observed and unobserved human capital explains much of the earnings differentials in employment outside the public sector. In this chapter I use the insight that individual heterogeneity is an important determinant of labour market outcomes to model the effects of comparative advantage and sorting in the Tanzanian labour market, in the tradition of Roy (1951). In this section I review the debate about wage differentials that the Harris and Todaro (1970) model has generated as well as describing alternative models of comparative advantage and sorting.

3.2.1 Segmented Labour Markets

In the previous chapter I noted that the Harris and Todaro (1970) model and the extension by Fields (1975) have provided the basis for much of the empirical research on earnings determination in labour markets in Africa, Latin America and in developing countries in other regions over the last 40 years.

Heintz and Posel (2008) and Badaoui et al. (2008) are two recent examples of papers testing for segmentation in South Africa that are representative of part of the debate that has occurred in the literature. Heintz and Posel (2008) use cross-sectional evidence from South Africa and find large wage differentials between the formal and informal sectors, using a definition of informality suggested by the ILO that includes workers in informal enterprises as well as workers in informal jobs working for formal firms. This leads them to conclude that the labour market is segmented between the formal and informal sectors[1]. Badaoui et al. (2008), however, come to the opposite conclusion using similar data from South Africa and the same definition of informality, though their analysis is limited to men only, in order that their results are not influenced by female labour force participation decisions. The authors also exclude the self-employed and those in the public sector to focus on the effects of informality on employees of profit maximising firms. They find that, although there is a difference in average earnings in the formal sector and informal sector when using the cross-section and controlling for observable human capital, there is no difference once they take into account tax payments in the formal sector, as well as unobserved heterogeneity, using the panel element of the data.

[1]They also suggest there are different types of informal employment, some of which have higher barriers to entry, which means that *within* the informal sector there may be different segments paying different wages on average.

In the previous chapter I explored the determinants of earnings differentials in the KIDS data from South Africa. I analysed the role of unions and the public sector in generating the large earnings differentials across sectors in South Africa. I also argued that observed and unobserved heterogeneity did seem to play an important role in explaining earnings differentials. I use this insight in this chapter and thus in the following section I explore the more recent labour economics literature that emphasises the importance of heterogeneity and the role of individual and job attributes in determining earnings.

3.2.2 Moving away from the representative worker

In 2000 James Heckman was jointly awarded the Nobel Prize in economics with Daniel McFadden, the first time the award had been given to individuals working in microeconometrics. In the conclusion to his Nobel Prize lecture Heckman explained the rapid evolution of microeconometrics in noting that "[i]n the past half century, economics has been enriched by vast new resources of microeconomic data. The data have opened the eyes of economists to the diversity and heterogeneity of economic life" (Heckman 2001, pg. 734). The exploration of the implications of this heterogeneity for economic theory, econometrics, and analysis of public policy continues and much of Heckman's Nobel lecture was devoted to summarising the frontier of this research.

The importance of heterogeneity in economic life has stimulated research in many areas of economics. In econometrics the explosion of the treatment effects literature and the new tools developed to assess these effects has been a response to the heterogeneity of responses to treatment that plagues estimation of average treatment effects, which I explore in more detail in chapter four of this thesis. The heterogeneity in individuals' choices of whether to work and how much to work motivated Heckman's early econometric work on female labour supply (Heckman 1974). The low explanatory power of human capital regressions is suggestive of the importance of unobserved individual heterogeneity. This observation has also led researchers to develop models of the labour market which can generate wage differentials even with identical workers (Mortensen 2003), as I explore in more detail in the next section.

The focus on heterogeneity has also influenced the debate on segmentation in developing countries in several ways. As I discussed above, the importance of unobserved individual heterogeneity has been offered as an alternative explanation of sectoral earnings differentials to segmentation, partly as a result of the emergence of panel data in developing countries.

Bill Maloney has taken a slightly different tack, suggesting that the utility of different jobs is affected by the unobserved characteristics of different types of employment. Maloney (1999) argues that researchers should take account of, or at least be aware of these unobserved characteristics, but that in practice they have mostly failed to do this in studies arguing for or against segmentation on the basis of earnings differentials. Maloney's conclusion is that "traditional earnings differentials cannot prove or disprove segmentation in the developing-country context" (Maloney 1999, pg. 275). Maloney has also suggested an alternative method of testing for segmentation. He argues that in a segmented labour market, movement between the formal and informal sectors would be limited. If formal sector jobs could be held for life (barring any bad conduct), as has been argued to be the case in some Latin American countries, then there should be movement from the informal sector to the formal sector but very little in the opposite direction.

Instead of thinking about a segmented labour market, Maloney (2004) suggests that self-employment in the informal sector represents a viable alternative to salaried work in the formal sector. This is as a result of the low firm productivity in many developing countries and the benefit of being one's own boss. Maloney (2004) thus argues that workers' sector choices may represent the best outcomes possible given their low education levels. This suggests a model of comparative advantage, similar in spirit to Roy (1951), in which workers choose between salaried employment in a low productivity firm and self-employment. I use this as the basis for my model of the Tanzanian labour market.

As I emphasise in chapter four of this thesis, the returns to education literature has been an area in which individual heterogeneity has been acknowledged to be a crucial factor in assessing the importance of education for individual earnings. This has led to the development of new econometric techniques that attempt to generate consistent estimates of the returns to education in the presence of individual heterogeneity.

As a result of the importance of the returns to education within labour economics, and as a key parameter for policy makers, there is a body of literature attempting to estimate the returns to education in African economies. Very little of this work, however, acknowledges any of the difficulties that individual heterogeneity creates in estimating an accurate return to education. In a review of previous studies of the estimates of the return to education in South Africa, for example, Keswell and Poswell (2004) do not include any studies that use instrumental variables as a result of education being endogenous and only one of the studies they review attempted to deal with ability bias (Moll 1998) and then only by including a measure of cognitive ability in the earnings function.

There has been some research emphasising the role of individual heterogeneity in determining individual earnings in African economies. Sandefur et al. (2007) show that the earnings distribution *within* different sectors has a large variance and also that there is a substantial overlap of the sectoral earnings distributions for Ghana, Tanzania and Ethiopia. Falco et al. (2011) have subsequently shown that this overlap still holds even when controlling for observable human capital. This analysis uses the Urban Panel Surveys for Ghana and Tanzania, though this analysis is still in the context of exploring sectoral differentials. There is also some research using manufacturing firm and worker data that attempts to explore the role of individual, as well as firm, heterogeneity in wage determination. Fafchamps et al. (2009) use data on matched manufacturing firms and workers from 11 African countries to explore whether sorting across firms and occupations can explain the return to education found in earnings regressions. Firm and occupational controls lower the estimated returns to education leading the authors to conclude that sorting occurs; through individuals with higher education levels both matching with more productive firms and being in more productive occupations.

Despite the existence of some empirical research on the role of heterogeneity I have not come across a single theoretical model of African labour markets that takes into account heterogeneity and comparative advantage. This chapter attempts to address this lacuna and the following section explores the class of models I use to do this.

3.2.3 Equilibrium Search Models

In addition to the acknowledgement of the "diversity of economic life", another outcome of the collection of large amounts of micro data was the realisation that the frictionless competitive model of the labour market could not explain many of the features these data suggested were important (Mortensen and Pissarides 1999). This led to the emergence of equilibrium search and matching models that incorporated labour market frictions and the necessity of search by both firms and workers. Important contributions to this literature have come from Peter Diamond, Dale Mortensen and Christopher Pissarides, who were jointly awarded the 2010 Nobel Prize in economics for their work.

Mortensen and Pissarides (1999) have noted there are two main areas of theoretical research within the equilibrium search framework that have traditionally been employed to explain different sets of stylised facts. Matching models with search and recruiting frictions have been used to explore the flows of workers between employment, unemployment and non-participation as well as job creation and destruction flows (Mortensen and Pissarides 1994). The second main area is the "wage posting" literature, which has been used to explain the distribution of wages in an economy as the outcome of a strategic game where firms post wages in a labour market with search frictions (Burdett and Mortensen 1998)[2].

In both matching and wage posting models, the labour market is assumed to be characterised by search frictions for both workers and firms. Thus information about profitable opportunities is not instantly available to all agents in the economy, unlike in the frictionless competitive model. It takes time, therefore, for workers to find firms with vacant jobs and for firms to fill these jobs. In matching models the rate at which firms and workers meet is modeled using a matching function, an aggregate representation

[2]Equilibrium search models with wage posting can generate different wages paid to identical individuals (Mortensen 2003). Interestingly, Fields (2004) sees this reflecting the influence of the earlier segmentation literature.

of how many workers and firms match at any point in time (Cahuc and Zylberberg 2004). In matching models, search frictions imply that firms have monopsony power over the workers they meet, and can exploit this to pay workers less than their marginal product. Firms and workers are assumed to bargain over the wage based on the value of the match created by the firm and worker, where the value may vary across matches if firms or workers are heterogeneous.

Recent research has used the equilibrium search framework in a number of ways to explain features of developing country labour markets. Satchi and Temple (2009) develop a macro matching model of the Mexican labour market where individuals choose employment in either the rural, urban formal or urban informal sectors. The model allows for the presence of frictions in the formal sector and assumes their absence in the informal sector, which helps to generate a realistically sized informal sector. They also assume, however, that the level of informal sector productivity is similar to the formal sector wage, which they argue is backed up in microeconometric studies using Mexican survey data. This assumption has the effect of making the frictionless informal sector more attractive and thus also larger. Because theirs is a macro model, Satchi and Temple (2009) focus on average wage differentials between sectors, as well as the relative size of the sectors. Individuals, however, are still homogeneous and as a result wages are also identical within sectors, implying that the model does not allow for any of the heterogeneity Heckman (2001) has noted is so important. Thus whilst it provides an explanation for the size of the informal sector, it cannot help to explain the distribution of earnings.

Albrecht et al. (2009) build a two-sided search model of the Mexican labour market based on the Mortensen and Pissarides (1994) matching model but allow for two sectors: the formal and informal sectors. This model does help to explain the observed earnings distribution. Workers differ in their formal sector productivity, which generates a distribution of formal sector earnings, however, informal sector productivity and earnings are assumed to be the same for all individuals. Informal sector productivity is calibrated to be 70 percent of the average level of formal sector productivity, with 65 percent of the population assumed to be more productive in the formal sector. The model thus sets up variation in the opportunity cost of informal employment. Some individuals have a low opportunity cost of informal sector employment as a result of their low productivity in the formal sector. These individuals take up informal employment whilst more productive individuals only accept formal employment opportunities. As a result earnings are higher in the formal sector, which contradicts the assertions of Satchi and Temple (2009), mentioned above, that earnings are roughly similar in the two sectors.

The assumption of wage bargaining in the formal sector means individuals are not paid their marginal product. The model incorporates ideas from Roy (1951) about comparative advantage in the labour market, but with ability varying in only one of the sectors. I use this as a basis for the model I build of the Tanzanian labour market but address some of its weaknesses, including the lack of a distribution of earnings in one of the sectors.

3.2.4 Differing Conceptions of the Source and Characteristics of the Informal Sector

The divergent models of segmentation and comparative advantage in developing countries have very different conceptions of the source and characteristics of the informal sector. As regards characteristics, in Fields's (1975) model the informal sector offers both low wages and hence low utility to individuals who are unable to obtain a formal sector job. Wages are set by bargaining with unions or by non-profit maximising public sector firms in the formal sector. No individuals would prefer the low paying informal sector in these models. By contrast models of comparative advantage assume that there are different abilities required in each sector and that some individuals are better at work in one sector than others. This means that most people prefer to work in either one sector or the other.

In Satchi and Temple (2009) the informal sector and the formal sector also differ in the levels of frictions. A justification for the assumed difference in sectoral frictions is not provided by the authors. Zenou (2008) argues that the informal sector is mostly comprised of self-employment where there can be no matching with a firm and informal work for friends or relatives, in which case co-ordination failures are much less likely to occur. He concludes that frictions are likely to be much lower in the informal sector and that this can be approximated by assuming there are no frictions in the informal sector.

In the Albrecht et al. (2009) model, the formal sector differs from the informal sector because there is wage bargaining in the formal sector, whereas informal earnings are simply the common informal sector productivity. In my model I take these distinctions to their logical conclusion by calling the two sectors wage and self-employment, with workers choosing between them based on their comparative advantage. Clearly, earnings determination differs between these two kinds of employment and the levels of frictions may also differ. In the following section I show that a distinction between wage and self-employment helps to make sense of the empirical evidence from Tanzania.

3.3 Data Sources

I use two sources of data from Tanzania in this chapter. The first is the nationally representative 2001 and 2006 Integrated Labour Force Surveys (ILFS) collected by the Tanzanian National Bureau of Statistics (NBS). I also make use of the 2004 and 2005 rounds of the Tanzanian Urban Panel Survey (TUPS) to explore individual movement between different types of employment. The ILFSs collected detailed individual member and household information from households in both urban and rural Tanzania. Population weights were calculated by NBS for each of the surveys but since the survey weights for 2001 have not been released publicly by NBS I have not weighted any of these data.

The TUPS data were collected by members of the Centre for the Study of African Economies (CSAE) at the University of Oxford. They were collected in the main urban centre of Tanzania and provide a small but representative sample of workers from urban Tanzania. The 2004 round was an urban subsample of those individuals surveyed by NBS for the 2000-2001 Household Budget Survey (HBS) and the 2005 round attempted to re-interview those interviewed a year before.

The Tanzanian UPS (along with a similar survey undertaken by CSAE in Ghana) provide one of the few individual panel surveys in Africa. Information was collected on both wage and self-employment, and thus the TUPS provides evidence on individual movement between wage and self-employment as well as between large and small firms and the public sector. The survey did not track those who had moved or could not be found, resulting in a 25 percent rate of attrition between 2004 and 2005, and implying the sample may not be representative of the 2005 urban population. I do not explore the effects of attrition in this chapter.

3.3.1 Measurement error in ILFS earnings data

Self-employment income has been considered difficult to measure and is thus often not collected in household surveys (Deaton 1997). Some recent work using the TUPS has, however, used self-employment income to explore the effects of experience, education and sector of employment on earnings (Rankin et al. 2010). In this chapter I also use of self-employment earnings data in Tanzania from the ILFSs and the TUPS.

In the Tanzanian ILFSs each survey first asked about gross income and the expenses incurred to earn it in the period in question (either a week or a month) for those engaged in self-employment. The surveys then asked about net income for the business.

Subtracting expenses from gross income produces a net income value in 2001 identical to the net income question asked directly, suggesting that the Tanzanian National Bureau of Statistics (NBS) forced the equivalence of these two measures in its post-survey cleaning. In 2006, however, the direct question on net income yields radically different answers from those calculated by taking gross income and subtracting expenses. The direct income question yields self-employment incomes that are on average six times higher than the measure using gross income and subtracting expenses.

The data from this income question also have some other problematic features, such as the amounts reported for weekly earnings being higher on average than those for monthly earnings, even before adjustment to a common period. In addition, when comparing the 2001 earnings data with inflation adjusted earnings in 2006 the direct income question in 2006 yields average increases in self-employment income over 5 years that are not matched by increases for employees in either the private or public sector. Further investigation revealed that NBS had again forced the equivalence of net income and gross income minus expenses by altering the net income measure but without taking into account whether individuals reported the net income or expenses figures as daily, weekly or monthly. This meant that if individuals reported weekly net income but monthly gross income and costs, which the survey instrument allowed, NBS forced the weekly net income to equal the monthly gross income. This means that when I convert net weekly income to a monthly figure I obtain an income figure much higher than actual income. As a result I use gross income minus expenses as the measure of self-employment income in both surveys.

Inflation makes comparison between the 2001 and 2006 surveys more difficult. A consumer price index (CPI) is collected monthly by the Tanzanian National Bureau of Statistics but there have been concerns raised about the accuracy of this index. I discuss deflating incomes in section 3.A.1 in the Appendix to this chapter.

3.4 Evidence for distinguishing between Models of the Tanzanian Labour Market

In this section I present empirical evidence from the Tanzanian labour market in an attempt to discriminate between competing ways of modelling this market. It is possible to think about the labour market as segmented, using Fields's (1975) update of the Harris and Todaro (1970) model, or to emphasise a comparative advantage model as first suggested by Roy (1951) and recently explored by Albrecht et al. (2009) in a developing country context.

3.4.1 Employment

I begin by exploring the distribution of employment in urban Tanzania in Table 3.1, using the ILFS data for 2001 and 2006. Table 3.1 shows that self-employment is the most common type of employment in urban Tanzania. Over 80 percent of the self-employed are own account workers with the remainder mostly employing a small number of employees: in 2001, conditional on having paid employees, 85 percent of the self-employed had 5 or fewer employees. Employment in the public sector is common and constituted 60 percent of urban wage employment in 2001, which is consistent with a picture of a segmented labour market in which there is a substantial amount of employment in firms that pay wages that are not profit maximising (though there is evidence that public sector employment declined relative to the size of most other kinds of employment between 2001 and 2006). Table 3.1 also shows that agriculture is still fairly important even within urban Tanzania.

The data show a substantial increase in the proportion of large firm employees between 2001 and 2006. Part of this may be explained by the change in the way firm size was captured in the two surveys. The 2006 ILFS survey only asked individuals whether the firm in which they worked had less than 10 employees or 10 or more. Thus the measure of firm size used is a crude one and I must assume any firm with 10 or more employees is a "large" firm. The 2001 ILFS asked if an individual worked in a firm with 1-5 employees, 5-10 and more than 10. I treat large firms in 2001 as those with more than 10 employees. Thus the definitions in 2001 and 2006 are not exactly comparable. This by itself does not seem to be a sufficient explanation for the jump in the number of individuals reported to be working for large firms between 2001 and 2006, however. It might be partly explained by different weightings used in each of the surveys but since I have been unable to access the 2001 population weights I cannot confirm this. My analysis suggests there are some issues of comparability between the surveys but in the absence of detailed reports on these surveys and how they were conducted, as well as surveys weights, it is difficult to explore this in any further detail.

Table 3.1 also shows the extent to which those who have obtained high levels of education obtain public sector employment. Over half of those who are working and have completed O levels (11 years of education under the new education system, 12 under the old) are working in the public sector. Those unable to obtain O levels seem to be disadvantaged as regards obtaining public sector employment, which I show in the next section is well paid relative to other forms of employment in Tanzania.

3.4.2 Earnings

The main focus in testing for segmentation has been on whether earnings differentials exists between sectors. Table 3.2 shows that median earnings in the ILFS are low for those in self-employment with no employees and in small firms, and are higher in large firms and the public sector, which seems consistent with a segmented labour market hypothesis. Wage employees were asked about their gross income in both ILFSs. It is not clear what impact the non-payment of taxes would have on sectoral income differentials. It is likely, though, that public sector workers paid tax but that small firm employees and the self-employed did not. Median earnings are also much higher for entrepreneurs with employees, who own what seem to be mostly informal businesses, with one third of these employers having written accounts and only 15 percent having accessed credit in the last year. Table 3.2

also shows low or negative earnings growth across the sectors between 2001 and 2006. This is surprising because this occurred at a time when Tanzania recorded real GDP growth of around 6 percent per year. It is consistent, however, with the Household Budget Survey analytical report (National Bureau of Statistics, Tanzania 2007), which suggests there was no decline in poverty between 2000 and 2007 as recorded in the HBSs in these years.

Differences in median earnings tell only part of story however. Figure 3.1 shows the real earnings densities in wage employment for the same categories as in Table 3.2 from the pooled 2001 and 2006 ILFS data. There is clearly much more variance in earnings in self-employment compared to wage employment[3]. Despite having similar medians, the density of earnings for the self-employed without employees has more mass in the upper tail of the density, compared to the density for employees in small firms, which has much more mass in the centre and roughly the same in the lower tail.

There are also large overlaps between what are generally considered formal and informal sector jobs. Around a third of employees in firms that have more than 10 employees earn less than the median in self-employment for those without employees. The differences in median earnings suggests that a focus on average earnings differentials is important but the large variation within sectors suggests a model of earnings determination should also account for this heterogeneity in earnings.

One obvious explanation for earnings variation is that there are differences in observable human capital. I explore this possibility in Figure 3.2 which plots the density of the unexplained part of earnings after a regression of log earnings that includes basic human capital controls (age, education,

[3]This may partly be explained by measurement error, since the self-employed are asked about and must therefore calculate their profits, whereas wage employees simply report their earnings, as explained above.

tenure), as well as gender, which is reported in Table 3.9 in the appendix to this chapter. Figure 3.2 is similar to Figure 3.1, although the variance of earnings in each sector is lower. This similarity is perhaps suprising, given that education seemed important in obtaining access to public sector employment, but makes more sense when considering that even in the public sector, human capital can only explain about 25 percent of the variance in earnings (see Table 3.9 in the Appendix to this chapter). This analysis suggests that the substantial overlap between wage and self-employment is not due to differences in observable human capital.

Figure 3.3 makes the distinction between wage and self-employment in graphing the earnings densities but includes public employment separately. There is clearly a large premium for those in the public sector, and I have shown above that this seems difficult to obtain without at least having obtained O levels. Within the private sector, however, there is not much difference between median earnings in self and wage employment, though there is substantially higher variance in self-employment, as I noted above.

A final piece of evidence is shown in Table 3.3, which gives earnings regressions with detailed sectoral controls, and controls for firm and own enterprise size using the TUPS. The TUPS is much smaller than the ILFS but allows a more substantial set of controls than the ILFS. As shown in Table 3.9 in the Appendix to this chapter using the ILFS data, and as discussed above, education, experience and gender are all important determinants of earnings. Table 3.3 also shows the large premia for public sector earnings and that those in larger firms earn more. Probably most interesting is the fact that individuals in the smallest firms actually earn less on average than those in self-employment. I return to this point when discussing the results of my model.

3.4.3 Employment Transitions

In the segmentation literature, following the tradition of the Harris and Todaro (1970) model, individuals are assumed not to move from the formal sector to the informal sector, usually as a result of labour legislation or union strength that prevent workers being fired. Maloney (1999) shows that in Mexico, transitions between the formal sector and the informal sector occur in both directions, thus questioning whether the Mexican labour market is segmented.

Table 3.4 uses longitudinal data from the 2004 and 2005 rounds of the Tanzanian Urban Panel Survey to explore how individuals move between different types of employment. Two things stand out from this table. The first is that there does seem to be little movement from the formal to informal sectors, with public sector workers and workers from larger firms very unlikely to move into self-employment, as the segmented labour market hypothesis would suggest. The second is that there is also very little movement between self-employment and wage employment, though there is movement within these categories between self-employment with, and without, employees, as well as some between small and large firm wage employment. This suggests that comparative advantage in wage and self-employment may play an important role in decisions about which sector to work in. It is possible that individuals are choosing to work either for themselves or for a firm, and that this may depend on their comparative advantage in each of these options.

3.4.4 Summary of empirical results

I have noted that some of the of the evidence presented above suggests that the urban Tanzanian labour market fits broadly into the Fields (1975) view.

Wages are high in the formal sector, particularly the public sector, and are low in an informal sector comprised of wage employment in smaller firms and the self-employed. There is also almost no movement from employment in the public sector into self-employment or small firms. I have also shown, however, that there is a large diversity in earnings within self-employment and private wage employment and that those owning informal firms who employ others have relatively high earnings, comparable with the earnings of employees in larger firms and the public sector. In addition, wage employees in small firms earn less on average than the self-employed without employees. The Tanzanian UPS transitions data also suggest that there is very little movement in either direction between wage employment (in small or larger firms or the public sector) and self-employment.

I have also noted that there are obvious differences between wage and self-employment in the way earnings are determined in each. The empirical evidence above, as well as the conceptual differences in earnings determination, suggest that a more fruitful way of modelling the labour market may be to assume that individuals choose between wage and self-employment based on their comparative advantage, which I undertake in the following section.

3.5 A Model of the Tanzanian Labour Market

3.5.1 Model Outline

Much of the labour economics literature on developing countries divides the labour market into a formal and informal sector where earnings are higher in the formal sector. Some of the evidence I showed in the previous section, however, does not support this view. As a result of this evidence, as well as intrinsic differences in earnings determination for wage and self-employment, I build a two-sector model of comparative advantage in the Tanzanian labour market that also captures the difference in the way earnings are determined in wage and self-employment. The empirical analysis above suggested the public sector provides a large part of the incentive to acquire advanced secondary or tertiary education by rewarding those who acquire jobs in it with high earnings. I do not include it in my model, however, and instead focus on the choice between private wage employment and self-employment. The model I build generates outcomes that are broadly consistent with the empirical evidence from Tanzania.

I use a matching framework in a Roy-type model (Roy 1951) of comparative advantage, in which individuals choose either wage or self-employment opportunities based on their expected earnings in each sector. Frictions limit the rate at which workers and firms meet each other. Frictions are also modeled in finding self-employment, so that workers do not instantaneously find self-employment opportunities. Thus I do not follow Zenou (2008) or Satchi and Temple (2009) in assuming that one of the sectors has no frictions, but this does not affect my main conclusions. Individuals earn their marginal product in self-employment, though wage employees do not as they are assumed to bargain with the firms they meet.

Homogeneity in ability would mean earnings would only differ across sectors but not within sectors. As a result of the variation in earnings within sectors that I highlighted above, I assume that individuals vary in self-employment productivity but that there is constant productivity across all firms and workers in wage employment. Despite there being variation only in self-employment productivity, my model also generates earnings variation in the bargained wage, as a result of the variation in workers' outside options due to their differences in self-employment productivity.

This incorporates ideas from Roy (1951), who assumes individuals choose to work in the sector that maximises their income and that this income ability differs in each sector, as well as including ideas from Maloney (1999) and Albrecht et al. (2009) who argue that wage employment has a low opportunity cost for individuals with low self-employment productivity and that self-employment becomes a relevant alternative when firm productivity is low.

As a result of the differences in self-employment productivity, individuals with low ability take only wage employment, some individuals take both, and some take only self-employment. All experience unemployment when their jobs are destroyed, and they then go back to searching for either wage opportunities or self-employment opportunities or both[4].

A more complex model might perhaps include variation in firm productivity (Postel-Vinay and Robin 2002), ability differences in both sectors (Roy 1951) or productivity shocks to matches (Mortensen and Pissarides 1994). I have built what I think is the simplest model that can explain some of the empirical evidence from the Tanzanian labour market, however, particularly earnings dispersion within wage and self-employment as well as a lack of movement between wage and self-employment that suggests a role for comparative advantage. The basis for the model is the Albrecht et al.

[4]Some readers may find it odd that I include a job destruction rate and frictions in self-employment, since this is often assumed to be a residual sector that those who lose their wage employment move to until they find another wage job. I have not motivated my choice with empirical evidence but it is possible to think of situations in which this may indeed be a plausible assumption. For example, someone selling items to drivers stuck in traffic jams may find his customers disappear if the road is improved. An individual selling things from an illegally sited kiosk may find it destroyed by government officials. In both cases the individual concerned would have to look for a new job, in the sense of finding a site to sell goods at. This would take time and could be modeled by assuming there are frictions in self-employment. Frictions may well be lower in self-employment, as argued in Zenou (2008) and I allow for this possibility, exploring the effects of varying the level of frictions in section 3.6.

(2009) matching model of a Latin American labour market, but I modify the model in ways that generate earnings distributions that better fit the evidence presented above.

3.5.2 Individual Behaviour

I begin by describing individual behaviour in the model. Individuals can choose to work in either wage or self-employment. I allow only self-employment productivity to vary across individuals. I assume that the population is constant and that self-employment productivity is distributed over a population of individuals whose size is normalised to 1. In much of the labour economics literature both employment for workers and the filled jobs of firms are thought of as assets that bring an instantaneous return, in the form of wage or self-employment income for workers and profits for firms. The capital value of these assets may change in the future as a result of shocks to the match which result in it being destroyed. In my model the probability that there will be a change in the capital value of the asset is the probability that the match will end. For workers, the new value of the employment asset is assumed to be the value of unemployment that results from the job ending. For the firm, the new value is the (negative) value of holding a vacancy open.

An individual with productivity y in self-employment earns y but faces a probability q_s that this self-employment opportunity will be destroyed, that she will enter unemployment and that there will be a change in the value of the self-employment asset. This means that the asset equation for the self-employed is[5]

$$rV_s(y) = y + q_s(V_u(y) - V_s(y)) \tag{3.1}$$

[5] I present an intuitive explanation of the asset equations that will be used throughout the model in section 3.A.2 in the Appendix to this chapter.

where V_u is the value of unemployment, V_s is the value of self-employment and r is the interest rate[6].

Unemployed workers search for employment opportunities and whilst doing this receive unemployment income $z(y)$, which is a function of their self-employment ability, as in Wong's (2003) model of the US labour market[7].

$$rV_u(y) = z(y) + \alpha \max[(V_s(y) - V_u(y)), 0] + m(\theta) \max[(V_e(y) - V_u(y)), 0]. \quad (3.2)$$

α is the probability that an unemployed worker will find a self-employment opportunity. $m(\theta)$ is the rate at which wage employment opportunities arrive and is a function of θ, a measure of labour market tightness. $\theta = v/u$, where v is the number of vacancies and u is the number of unemployed persons. When the labour market is tight v is low relative to u, θ is thus low and workers will have a harder time finding firms to match with, implying that the matching rate will be lower. The matching rate in self-employment is assumed to be exogenous. When an individual finds a self-employment opportunity or a wage employment opportunity, if she matches with a firm, this results in a change in the capital value of the unemployment asset ($V_u(y)$) that an unemployed person is thought of as holding.

I noted above that variation in self-employment productivity creates an opportunity cost of taking wage employment that is increasing in self-employment productivity, as in the model of Albrecht et al. (2009). Thus workers with low self-employment productivity choose employment in the

[6]For reference I list and provide a brief explanation of all the model parameters and variables in Table 3.5.

[7]Assortative mating by income, education or assets is a possible explanation for unemployment earnings depending on productivity, with a spouse providing income in a period of unemployment, but neither the household nor education are taken into account explicitly in this model. Fafchamps and Quisumbing (2005) provide evidence that sorting on human capital occurs in the Ethiopian marriage market.

wage sector and reject potential self-employment options, because they would earn much less in self-employment and are better off searching for wage opportunities. Similarly, workers with high self-employment productivity reject wage employment offers. Leaving unemployment, and entering wage employment, is worse than staying in unemployment and waiting for future self-employment opportunities for individuals with high self-employment productivity. This is captured by the max operator in the last part of equation (3.2). Similarly, taking self-employment opportunities for unemployed individuals with low self-employment productivity, is worse than waiting for future wage employment opportunities, and again the max operator in the second to last part of equation (3.2) captures the fact that individuals do not take opportunities that leave them worse off than their current situation.

The asset equation for wage employees is

$$rV_e(y) = w(y) + q_e(V_u(y) - V_e(y)), \tag{3.3}$$

where the wage w depends on y and V_e is the value of wage employment. As in self-employment, wage employment is assumed to end with an exogenous probability (q_e in wage employment) in which case the asset value $V_e(y)$ changes and the worker returns to unemployment.

Individuals only differ in their ability in self-employment. Despite this, the model generates variation in earnings in both sectors, shown in the dependence of w on productivity y in equation (3.3). This dependence of the wage on ability is a result of variation in the outside options of workers who are assumed to bargain with the firms with which they match. I explain the bargaining process in Section 3.5.4 below.

3.5.3 Firms

Firms are assumed to have homogeneous productivity y_w and to employ only one worker. They match with workers at the rate $m(\theta)/\theta$. When a match occurs the firm produces a continuous stream of output y_w and pays a wage $w(y)$ to the worker it matches with. The wage paid depends on worker ability because ability affects the worker's outside option. Filled jobs can also be thought of as assets that firms hold, meaning that the asset equation for a filled job is

$$r\prod_e(y) = y_w - w(y) + q_e(\prod_v - \prod_e(y)). \tag{3.4}$$

Matches are assumed to end with exogenous probability q_e, which brings about a change in the capital value of the filled job asset for the firm $\prod_e(y)$. Firms whose matches end are assumed to return to search for workers who are also searching for employment. While this search occurs, firms are assumed to incur a cost of holding a job open, meaning that the asset equation for a vacant job is

$$r\prod_v = -c + (m(\theta)/\theta)E[\max[(\prod_e(y) - \prod_v, 0]]. \tag{3.5}$$

In equation (3.5), c is the cost of holding a vacant job open per unit of time and $m(\theta)/\theta$ is the rate at which firms match with workers. Following much of the literature, I also assume that firms are risk neutral.

3.5.4 Wage Determination

The match between a firm and a worker generates a gain for both parties compared to what they would have earned if the match did not occur. These gains are called rents and the sum of the rents of the firm and worker are termed the surplus of the match in the matching literature. When workers and firms match they are assumed to bargain over how this surplus is distributed between them. Setting up a simple non-cooperative bargaining game with some straightforward assumptions can generate a "surplus sharing rule" (Cahuc and Zylberberg 2004).

In the model I have set out, a worker's outside option is the value of unemployment and a firm's is the value of keeping a vacancy open. Thus the rents are $(V_e(y) - V_u(y))$ for a worker and $\prod_e - \prod_v$ for a firm. This means that the surplus is $S = \prod_e - \prod_v - (V_e(y) - V_u(y))$.

I assume that with free entry firms compete away expected profits, so that $\prod_v = 0$ and hence that the firm's outside option has an expected value of zero. Thus $S = \prod_e - (V_e(y) - V_u(y))$. Following the matching literature I assume that the bargaining outcome is the maximum of the Nash criterion (Nash (1953), Binmore et al. (1986)). Thus the negotiated wage, w(y), maximises

$$[V_e(y) - V_u(y)]^{\gamma}[\prod_e(y)]^{(1-\gamma)}. \tag{3.6}$$

The first order conditions from this problem give $\prod_e(y) = (1 - \gamma)S$ and $V_e(y) - V_u(y) = \gamma S$, the surplus sharing rule. Thus firms and workers share the surplus according to their bargaining power, which is determined outside the model. Substituting the value of $\prod_e$ from equation (3.4) it is possible to show that

$$w(y) = \gamma y_w + (1 - \gamma) r V_u(y)). \tag{3.7}$$

Equation (3.7) shows a dependence of the negotiated wage w on y because workers' outside options, $V_u(y))$, vary with self-employment ability y, despite identical productivity for all employees. There are two sources of this variation. The first is that unemployment income depends on self-employment ability y. The second is that even if unemployment income was constant across workers, the value of the unemployment asset $V_u(y)$) differs for different kinds of workers, as I explore in section 3.5.5.

3.5.5 Productivity Cut Offs

The model setup and the variation in self-employment ability implies an opportunity cost of taking wage employment that is increasing in self-employment ability. Workers with high self-employment productivity prefer unemployment to wage employment and will reject wage employment offers because taking these would mean foregoing high self-employment income. Workers with low self-employment ability prefer unemployment to self-employment and will reject self-employment opportunities because these mean foregoing higher income in wage employment. Some individuals will take both kinds of employment because the incomes they earn in each are similar. The equilibrium of this model will thus be one in which there are two cutoff values in ability, y^* and y^{**} with $y^* < y^{**}$. These correspond respectively to the self-employment productivity at which unemployment is no longer preferred to self-employment, and at which wage employment is no longer preferred to unemployment.

Individuals with ability $y^* < y < y^{**}$ take both self-employment and wage employment, whilst workers with ability $y < y^*$ do not take self-employment. This means that $y^* < y < y^{**}$ workers have an extra term in their unemployment asset equation (3.31), where their self-employment ability y enters directly. This affects the outside options of these workers and hence their wages have an extra source of wage variation compared to the $y < y^*$ workers , whose self-employment ability does not enter directly into their outside option in the wage bargain, since they are assumed never to take self-employment in this equilibrium.

I have thus injected a further degree of plausibility into the model by generating earnings dispersion in both sectors, despite only having one type of ability, in self-employment. In the next sections I follow Albrecht et al. (2009) in first solving the model, and then showing that the equilibrium is unique.

3.5.6 Solving the Model

Solving the model requires obtaining the cutoff values y^* and y^{**} and solving for the equilibrium value for θ, labour market tightness. I begin by finding an expression for y^* in terms of the exogenous model parameters and θ. Below y^*, $V_u(y) > V_s(y)$ (unemployment is preferred to self-employment). This explains why low ability individuals take only wage employment opportunities. Individuals with productivity above y^*, but below y^{**}, take both wage and self-employment. At y^*, $V_u(y^*) = V_s(y^*)$. This implies the max operator in the unemployment asset equation is now irrelevant for individuals with productivity y^*, because taking a wage job does not imply a decrease in the capital value of the unemployment asset. Thus equation (3.2) becomes

$$rV_u(y^*) = z(y^*) + m(\theta)(V_e(y^*) - V_u(y^*)). \tag{3.8}$$

It is then possible obtain $V_e(y^*)$ in terms of $V_u(y^*)$ from equation (3.3), the wage employment asset equation:

$$V_e(y^*) = \frac{w(y^*) + q_e V_u(y^*)}{r + q_e}, \tag{3.9}$$

and then solve for $V_u(y^*)$:

$$rV_u(y^*) = \frac{z(y^*)(r + q_e) + m(\theta)w(y^*)}{r + q_e + m(\theta)}. \tag{3.10}$$

Since $V_u(y^*) = V_s(y^*)$ at $y = y^*$, equation (3.1), the self-employment asset equation, then implies $rV_s(y^*) = y^*$. Thus using equations (3.10) and (3.7) I can express y^* in terms of θ and the exogenous parameters of the model:

$$y^* = \frac{z(y^*)(r + q_e) + m(\theta)(\gamma y_w + (1 - \gamma)y^*))}{r + q_e + m(\theta)}. \tag{3.11}$$

In a similar fashion I can also find the value of y^{**}, the level of productivity above which individuals take only self-employment. Above y^{**}, $V_u(y) > V_e(y)$, meaning unemployment is preferred to wage employment. At y^{**}, $V_u(y^{**}) = V_e(y^{**})$ and hence equation (3.2), the unemployment asset equation, can be written as

$$rV_u(y^{**}) = z(y^{**}) + \alpha(V_s(y^{**}) - V_u(y^{**})). \tag{3.12}$$

Equation (3.1), the self-employment asset equation, then implies that

$$V_s(y^{**}) = \frac{y^{**} + q_s V_u(y^{**})}{r + q_s}. \tag{3.13}$$

I can then use equations (3.12) and (3.13) to substitute out $V_s(y^{**})$ and solve for $V_u(y^{**})$:

$$rV_u(y^{**}) = \frac{z(y^{**})(r + q_s) + \alpha y^{**}}{r + q_s + \alpha}. \tag{3.14}$$

y^{**} is the value of y such that $V_u(y^{**}) = V_e(y^{**})$ and this combined with equation (3.3), the wage employment asset equation, implies that $rV_e(y^{**}) =$

$w(y^{**})$. This and equation (3.14) imply that

$$w(y^{**}) = \frac{z(y^{**})(r + q_s) + \alpha y^{**}}{r + q_s + \alpha}. \tag{3.15}$$

Now the bargained wage equation (3.7), with $V_u(y^{**}) = V_e(y^{**})$ and $rV_e(y^{**}) = w(y^{**})$, implies that $w(y^{**}) = y_w$. This is intuitive because the highest wage firms can pay must be paid to workers with the highest ability in self-employment who still take wage employment, and this is the wage which leaves the firm indifferent between agreeing to the match and returning to search for another worker, where the expected return is zero in equilibrium as a result of free entry.

I am then able to form an equation for y^{**} in terms of the exogenous parameters of the model:

$$\alpha y^{**} + z(y^{**})(r + q_s) = y_w(\alpha + q_s + r). \tag{3.16}$$

With a functional form for $z(y)$ I can solve directly for y^{**}. I also have one equation with two unknowns y^* and θ, and require a second to be able to solve for these variables. To do this it is necessary to first solve for the steady state employment flows, which I undertake in the next section.

3.5.7 Equilibrium Employment Flows

In the steady state equilibrium of this model the unemployment rate cannot be changing, implying that the flows into and out of unemployment must be equal. These flows differ by the workers' ability type, and by the sector of employment they choose. There are thus different steady state conditions for individuals with $y < y^*$, who take only wage employment, for those with $y^* < y < y^{**}$, who take both wage and self-employment opportunities, and for those with $y > y^{**}$, who take only self-employment. These steady state conditions give the unemployment rates for the different groups of workers and the overall unemployment rate for the population.

Following Albrecht et al. (2009), I let $u(y)$ be the fraction of time an individual of productivity y spends unemployed, $n_e(y)$ be the fraction of time spent in wage employment and $n_s(y)$ be the fraction of time spent in self-employment, meaning that $u(y) + n_e(y) + n_s(y) = 1$[8].

Individuals with productivity $y < y^*$ only move between wage employment and unemployment, implying that $n_s(y) = 0$. This means there is only one steady state condition for this group, which is that the flow of workers out of wage employment must equal the flow into wage employment. q_e is the rate at which wage jobs are destroyed, and $m(\theta)$ is the rate at which unemployed workers match with firms. The steady state condition is thus $q_e(1 - u(y)) = m(\theta)u(y)$, with

$$u(y) = \frac{q_e}{m(\theta) + q_e} \tag{3.17}$$

and

$$n_e(y) = 1 - u(y) = \frac{m(\theta)}{m(\theta) + q_e}. \tag{3.18}$$

Individuals with $y^* < y < y^{**}$ take both wage and self-employment opportunities, meaning there are two steady state conditions for these individuals. Firstly, the flows into wage employment equal the flows into unemployment from wage employment, and secondly, the flows out of self-employment equal the flows into unemployment from self-employment. Hence $m(\theta)u(y) =$

[8]I do not allow for non-participation in this model.

$q_e n_e(y)$ and $\alpha u(y) = q_s n_s(y)$. Combining these two implies

$$u(y) = \frac{q_e}{m(\theta) + q_e + \frac{\alpha q_e}{qs}}, \tag{3.19}$$

$$n_e(y) = \frac{m(\theta)}{m(\theta) + q_e + \frac{\alpha q_e}{qs}} \tag{3.20}$$

and

$$n_s(y) = \frac{q_e \alpha}{q_s(m(\theta) + q_e + \frac{\alpha q_e}{qs})}. \tag{3.21}$$

Individuals with $y > y^{**}$ take only self-employment so there is again only one steady state condition, that the flow from unemployment to self-employment equals the reverse flow from self-employment into unemployment. This implies $\alpha u(y) = q_s n_s(y)$, so that $n_s(y) = 1 - u(y)$, $n_e(y) = 0$,

$$u(y) = \frac{q_s}{\alpha + q_s} \tag{3.22}$$

and

$$n_s = \frac{\alpha}{\alpha + q_s}. \tag{3.23}$$

Total unemployment can then be calculated as

$$u = \int_0^{y^*} u(y)f(y)dy + \int_{y^*}^{y^{**}} u(y)f(y)dy + \int_{y^{**}}^1 u(y)f(y)dy, \tag{3.24}$$

where $u(y)$ varies for each of the three categories of worker, and $f(y)$ is the ability distribution in the population.

3.5.8 Finding the Equilibrium

The model can be solved by finding a second equation containing the two unknowns y^* and θ, which can be obtained from the free entry condition

$\prod_v(y) = 0$. This, together with equation (3.5), the vacant job asset equation, implies

$$c = \frac{m(\theta)}{\theta} E[\max[(\prod_e(y), 0]]. \tag{3.25}$$

Now in the equilibrium $\prod_e(y) \geqslant 0$ because firms only match with workers that they do not make a loss from hiring and hence

$$c = \frac{m(\theta)}{\theta} E[\prod_e(y)]. \tag{3.26}$$

The value of a filled job to a firm is uncertain because the wage paid depends on the ability of the worker employed. In equilibrium only workers with productivity $y < y^{**}$ take wage employment offers. To calculate the expected value of a filled job, $E[\prod_e(y)]$, I also need to know the distribution of workers with which firms can potentially match. I know the distribution of worker ability in the population but this is different to the distribution of unemployed workers who would take wage employment offers from firms. Following Albrecht et al. (2009), I denote the density of ability for those in unemployment as $f_u(y)$ and then use Bayes' Law to write

$$f_u(y) = \frac{u(y)f(y)}{u}. \tag{3.27}$$

Hence I can rewrite the free entry condition (3.25) as

$$c = \frac{m(\theta)}{\theta} \int_0^{y^{**}} \prod_e(y) \frac{u(y)f(y)}{u} dy. \tag{3.28}$$

The equation for a filled job, equation (3.4), with $\prod_v = 0$, implies that $\prod_e(y) = \frac{y_w - w(y)}{r+q_e}$. I also know the value of $w(y)$ from equation (3.7), the outcome of the Nash Bargain with free entry. Hence I can write the free entry condition as

$$c = \frac{m(\theta)(1-\gamma)}{\theta(r+q_e)} \int_0^{y^{**}} (y_w - V_u(y)) \frac{u(y)f(y)}{u} dy. \tag{3.29}$$

I now have the second of the two equations in y^* and θ, which can be solved to find the equilibrium values of these variables, once I have values for $V_u(y)$ for those workers with productivity below y^{**}, who take wage employment and who can be hired by firms.

The values of $V_u(y)$ differ for individuals with productivity above and below y^*. Workers with $y^* < y < y^{**}$ take both wage and self-employment, meaning that the self-employment and wage employment asset equations (equations (3.1) and (3.3)) still apply and the unemployment asset equation (equation (3.2)) becomes

$$rV_u(y) = z(y) + \alpha(V_s(y) - V_u(y)) + m(\theta)(V_e(y) - V_u(y)). \tag{3.30}$$

Solving these three equations for rV_u implies that

$$rV_u(y) = \frac{z(y)(r+q_s)(r+q_e) + \alpha y(r+q_e) + \gamma m(\theta) y_w (r+q_s)}{(r+q_s)(r+q_e) + \alpha(r+q_e) + \gamma m(\theta)(r+q_s)}. \tag{3.31}$$

For workers with $y < y^*$ only the wage job asset equation (3.3) still applies and the unemployment asset equation (3.2) becomes

$$rV_u(y) = z(y) + m(\theta)(V_u(y) - V_e(y)), \tag{3.32}$$

since unemployment is preferred to self-employment for $y < y^*$ individuals. Solving equations (3.3) and (3.32) implies that

$$rV_u(y) = \frac{z(y)(r+q_e) + \gamma m(\theta) y_w}{r + q_e + \gamma m(\theta)}. \tag{3.33}$$

I now have the values for $rV_u(y)$ in terms of the exogenous parameters. I can then also substitute the values of y^* from equation (3.11) and y^{**} from equation (3.16) into the free entry condition (3.29). This is a non-linear equation in theta, meaning I am required to solve it numerically, which I explain in more detail below, after proving that a unique equilibrium to the model exists.

3.5.9 The Existence of a Unique Equilibrium

A steady state equilibrium is one in which the expected value of maintaining a vacancy is zero, matches are consummated only if it is in the interests of the firm and individual to do so, the steady state employment flows conditions hold, wage employment is not worthwhile for individuals with $y > y^{**}$ and self-employment is not worthwhile for individuals with $y < y^*$ (Albrecht et al. 2009). A value of θ together with cutoffs y^* and y^{**} and an unemployment rate $u(y)$ at which the above conditions are satisfied is an equilibrium of the model.

Since I have shown above that $u(y)$ and y^* are uniquely determined by θ, and that y^{**} is determined by the exogenous parameters of the model, a unique equilibrium exists if there is a unique value of θ that solves the free entry condition (3.28). If I assume decreasing returns to scale in the matching function m, a standard assumption in the literature (Cahuc and Zylberberg 2004), then $\frac{m(\theta)}{\theta}$ is strictly decreasing in θ. $u(y)$ has different values for $y < y^*$ and $y^* < y < y^{**}$ but, as can be seen from equations (3.17) and (3.22), both are strictly decreasing in θ. Since u is the sum of these two terms for $u(y)$ that are both strictly decreasing in θ, it is also strictly decreasing in θ (u also includes a term for $u(y)$ above y^{**} which does not depend on θ).

To show that the free entry condition is decreasing in θ I now only need to show that $rV_u(y)$ is increasing in θ. Firstly, for $y < y^*$, equation (3.33) can be written as

$$rV_u(y) = \frac{z(y)A + By_w}{A + B}, \tag{3.34}$$

where only B depends on θ. I can differentiate $rV_u(y)$ and find that

$$\frac{\partial rV_u(y)}{\partial \theta} = \frac{(y_w - z(y))AB'}{(A + B)^2}. \tag{3.35}$$

Since $y < y^*$ this means $y < y_w$ and since $z(y) < y$ this implies $y_w - z(y)$ is always positive. The other terms are clearly also all positive so $rV_u(y)$ is increasing in θ for $y < y^*$.

For $y^* < y < y^{**}$ the value of $rV_u(y)$ is found in equation (3.31), which can be written as

$$rV_u(y) = \frac{z(y)D + Ey + Fy_w}{D + E + F}, \tag{3.36}$$

where only F depends on θ. To determine when $rV_u(y)$ is increasing in θ I calculate

$$\frac{\partial rV_u(y)}{\partial \theta} = \frac{F'[D(y_w - z(y)) + E(y_w - y))]}{(D + E + F)^2}. \tag{3.37}$$

Since F' and $(D + E + F)^2$ are always positive this implies that $rV_u(y)$ is increasing in θ if $D(y_w - z(y)) + E(y_w - y)) > 0$. Since $y^* < y < y^{**}$ this means $y_w - z(y) \geq 0$ and $y_w - y \geq 0$. And since D and E are both positive this implies $rV_u(y)$ is increasing in θ for $y^* < y < y^{**}$.

Having shown that the right hand side of the free entry condition is strictly decreasing in θ, to prove that the equilibrium is unique I must show that as $\theta \to \infty$ the right hand side of the free entry condition tends to zero and that it tends to ∞ as $\theta \to 0$. $\frac{m(\theta)(1-\gamma)}{\theta(r+q_e)}$ clearly tends to ∞ as $\theta \to 0$ and tends to zero as $\theta \to \infty$. Thus I only need to show that the definite integral is bounded to show that the equilibrium is unique. u and $u(y)$ are unemployment rates that are bounded between 0 and 1 for any value of θ. $f(y)$ is the distribution of abilities in the population and so does not depend on θ, and neither does y_w. $V_u(y)$ is the value of unemployment. As $\theta \to 0$ the labour market becomes very tight and matching with a firm increasingly unlikely. Thus, in the limit, individuals would not find wage employment. But as the model has been set up they can still find self-employment, where the arrival rate does not vary with θ. Thus the value of unemployment is still bounded in this case.

A similar argument holds when $\theta \to \infty$ and wage jobs are immediately available for those who become unemployed, except that some individuals still take self-employment, even though wage jobs are available instantly, because their employers would not pay a wage greater than the common level of productivity. Thus the definite integral is bounded and there is a unique value of θ that solves the free entry condition.

I have outlined a unique equilibrium above in which there are two cutoffs y^* and y^{**}. Though the equilibrium will always be unique there may actually be two, one or zero cutoffs, depending primarily on the exogenous value of y_w but also on the other exogenous parameters. At values of y_w close to 1 there are no individuals who only take self-employment and some who take both wage and self-employment. If y_w takes a value above 1, however, taking self-employment becomes relatively less lucrative and at a high enough value of y_w eventually all individuals choose only wage employment. Due to the high levels of self-employment in Tanzania I focus on the two cutoff type of equilibrium in this chapter, though I do discuss the implications of varying y_w and some of the other exogenous parameters below.

3.5.10 Equilibrium Earnings distributions in Wage and Self-Employment

In the model I have outlined wages depend on worker's outside options, which are themselves determined by individuals' self-employment productivities, and the self-employed earn their productivity. This gives rise to a distribution of wages and also of self-employment earnings. I now calculate these distributions in equilibrium, which then enables me to calculate the average wage paid in each sector.

$f(y)$ is the density of abilities in the population. If $h(y)$ is the density of ability in wage employment and $P(E)$ is the probability of being in wage employment, Bayes Law then implies that

$$h(y) = \frac{P(E|y)f(y)}{P(E)} = \frac{n_e(y)f(y)}{\int_0^1 n_e(y)f(y)dy}. \tag{3.38}$$

For simplicity I assume a uniform distribution of ability between 0 and 1. It is clear from equations (3.18) and (3.20) above that the probability of wage employment, n_e, differs for different values of y. This means

$$\begin{aligned}
h(y) &= \frac{\frac{m(\theta)}{m(\theta)+q_e}}{\frac{y^* m(\theta)}{m(\theta)+q_e} + (y^{**} - y^*)\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}} && 0 \leq y < y^* \\
&= \frac{\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}}{\frac{y^* m(\theta)}{m(\theta)+q_e} + (y^{**} - y^*)\frac{m(\theta)}{m(\theta)+q_e+\frac{\alpha q_e}{q_s}}} && y^* < y < y^{**} \\
&= 0 && y^{**} < y \leq 1
\end{aligned}$$

Since $h(y)$ takes a constant value for $0 \leq y < y^*$, another constant value on $y^* < y < y^{**}$ and is zero beyond y^{**} calculating the distribution of wages $m(w)$ is relatively simple. The distribution will also have two constant parts on the domain of w, but with a much narrower base than the ability distribution. If I let the height of the first block of the $h(y)$ distribution be a then the probability mass within this block is ay^*. The area of the first block of the distribution of w will be the same and thus the height of this block will be $\frac{ay^*}{w(y^*)-w(0)}$. Using similar reasoning, the height of the second block will be $\frac{b(y^{**}-y^*)}{w(y^{**})-w(y^*)}$. Solving the model numerically I will be able to calculate the equilibrium wages at $y = 0$, $y = y^*$ and $y = y^{**}$, obtain the density and then be able to solve for the average wage paid in the equilibrium. The average wage is calculated as $\int_0^{y^{**}} wm(w)dw$.

Calculating the density of earnings in self-employment is also simple. Again, the density will differ for those with ability y above and below y^{**} since the probability of self-employment varies with ability. Denoting the probability of self-employment as $P(S)$ I can write the distribution of ability in self-employment, again using Bayes' rule, as

$$g(y) = \frac{P(S|y)f(y)}{P(S)} = \frac{n_s(y)f(y)}{\int_0^1 n_s(y)f(y)dy}. \tag{3.39}$$

From equations (3.21) and (3.23) I have the probability of being in self-employment for individuals with $y > y^*$ and I know that $P(S) = 0$ if $y < y^*$. Thus it is possible to write down the density of ability in self-employment as

$$
\begin{aligned}
g(y) &= 0 && 0 < y < y^* \\
&= \frac{\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})}}{((y^{**}-y^*)\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})})+(1-y^{**})\frac{\alpha}{\alpha+q_s}} && y^* < y < y^{**} \\
&= \frac{\frac{\alpha}{\alpha+q_s}}{((y^{**}-y^*)\frac{q_e\alpha}{q_s(m(\theta)+q_e+\frac{\alpha q_e}{q_s})})+(1-y^{**})\frac{\alpha}{\alpha+q_s}} && y^{**} < y \leq 1
\end{aligned}
$$

Since earnings in self-employment are simply the value of individual productivity in self-employment, the ability distribution is also the earnings distribution in self-employment.

3.6 Numerical Solutions to the Model

Solving for two unknowns θ and y^* in two non-linear equations requires numerical methods. I use Broyden's method of solving nonlinear equations numerically, a quasi-Newton method first outlined by Broyden (1965). This was implemented using the Compecon Matlab package, a companion package to the book *Applied Computational Economics and Finance* (Miranda and Fackler 2002).

The two non-linear equations to be solved are the free entry condition (3.29) and the solution for y^* (3.11). A solution requires assuming values for the exogenous parameters and functional forms of the model, or estimating the model and parameters using data. As far as I am aware, there have been no matching models estimated on African, or possibly even any developing country data, and as a result there are no empirical estimates of most of the parameters for which I require values. I do not attempt to prove the parameters are identified or estimate my model on the Tanzanian Urban Panel Survey data here either; rather, I simply follow other papers in using what seem to be reasonable parameter values and then explore what the model predicts about the earnings distributions across wage and self-employment, transitions between the two sectors, and average wages in each sector.

3.6.1 Exogenous Parameter Values

The values of the exogenous parameters which I need in order to solve the model are those of the interest rate, the job destruction rate in self-employment and wage employment, the arrival rate of self-employment opportunities, firm productivity, the cost to a firm of holding a job open and the bargaining power parameter. I also require a functional form for $z(y)$, which shows how unemployment income depends on ability as well as a functional form for the matching function $m(\theta)$. I assume that the unit of time is one year.

Providing empirical estimates of the rest of the exogenous parameters and functional forms is challenging and I do not attempt to explore the conditions under which the parameters would be identified. I thus follow some of the assumptions of Albrecht et al. (2009), the only other two sector developing country microeconomic matching model I am aware of, in assuming values for these parameters and the functional forms, but modify these slightly for the Tanzanian context in some cases. I assume the matching function to be of the form $m = 4\theta^{.5}$, the cost of holding a job open $c = .3$, the arrival rate of self-employment offers $\alpha = 3$, the real interest rate $r = 0.05$ and a uniform distribution of self-employment productivity between 0 and 1. I assume that $y_w = .5$ which, given $y \sim \text{uniform}(0, 1)$, is the mean and median of this productivity. I discuss the implications of other potential underlying distributions below. I finally assume that $z = .3y$, so that individuals have unemployment income that is 30 percent of their self-employment productivity.

3.6.2 Equilibrium Solution

Having assumed the value of parameter values, as well as of functional forms, I now explore the solution to the baseline model, before undertaking some comparative statics.

The equilibrium outcome with the chosen parameter values and functional forms is shown in the third column of Table 3.6. In this baseline equilibrium $y^* = .3936$, $y^{**} = .5608$ and $\theta = .51$. This means that 39 percent of the population choose wage employment only, around 17 percent choose either wage or self-employment and 44 percent choose only self-employment. Total unemployment is 13.5 percent, with unemployment being 15 percent for individuals with $y < y^*$, 7.8 percent for individuals with $y^* < y < y^{**}$ and 14 percent for high productivity individuals with $y > y^{**}$. Individuals with productivity $y^* < y < y^{**}$ take both wage and self-employment opportunities so their unemployment rate is much lower than for those with $y < y^*$ or $y > y^{**}$ who take only wage and self-employment offers respectively. The vacancy rate $v = u\theta$ so v =6.75 percent. The average wage paid is .44 and the average earnings in self-employment is .73, meaning that the model generates earnings in self-employment that are much higher than those in wage employment. I discuss this further in section 3.6.2 below.

Earnings variation in wage employment

Those who take only wage employment in the equilibrium ($0 < y < y^*$) have an extremely low variance of earnings since $w(0) = .43$ whilst $w(y^*) =$

$w(.39) = .447$. This narrow distribution is a result of identical firm productivity and the small variation in the outside options of those who choose only wage employment in the equilibrium. For these individuals, their self-employment productivity does not enter directly into their outside option in the wage bargain with firms, because they never actually exercise their self-employment option in the equilibrium of the model. Thus only the difference in workers' unemployment income, which is assumed to be a function of self-employment productivity, generates wage differentials amongst these workers.

There is more wage variation over a narrower range of abilities for those who also take self-employment opportunities ($y^* < y < y^{**}$) since $w(y^*) = w(.39) = .447$ and $w(y^{**}) = w(.56) = .5$. This is because these individuals do exercise their self-employment options, which enter the wage bargain and generate greater variation in individuals' outside options.

Having shown how it is possible to generate the distribution of earnings conditional on wage employment in Section 3.5.10 above, I now undertake this and describe the equilibrium wage density. It consists of two rectangular blocks, one between $w(0) = .43$ and $w(y^*) = .447$ containing mass .82 (in the row of Table 3.6 labeled "wmass 1") and the other between $w(y^*) = .447$ and $w(y^{**}) = .5$ containing mass .18 (in the row of Table 3.6 labeled "wmass 2"). The corresponding cdf is shown in Figure 3.4 (labeled baseline), which emphasises the lack of variation in wages. Allowing unemployment income to vary by self-employment ability thus buys some degree of realism at the lower end of the earnings distribution, in that there is some variation in the bargained wage for individuals taking only wage employment, which would not be the case if unemployment income was fixed (there would be a mass point). Figure 3.4 suggests, however, that the amount of variation is limited.

Earnings variation in self-employment

Those taking only self-employment simply earn their self-employment productivity, which means that there is a much larger amount of variation in self-employment earnings compared to wages earned working for firms with identical productivity. The minimum amount earned in self-employment is y^*, whilst the maximum is the highest productivity in the distribution of y, which are .39 and 1 in the baseline model. Thus the maximum and minimum earnings for the population are respectively higher and lower than the maximum and minimum earnings in wage employment.

The probability of self-employment differs above and below y^{**}, because unemployed individuals with productivity $y < y^{**}$ also take wage employment offers if these arrive before self-employment opportunities. This means the density of earnings conditional on being self-employed also consists of two rectangular blocks, the first between y^* and y^{**} and the second between y^{**} and 1. The first contains mass .17 (the row labeled "smass 1" in Table 3.6) and the second .83 in the baseline equilibrium equilibrium (in the row of Table 3.6 labeled "smass 2").

Transitions

The comparative advantage of individuals in one or other sector in the model means that most individuals prefer to take opportunities in only one of the sectors. Only those individuals with self-employment productivity near the constant level of firm productivity take both wage and self-employment opportunities. This is consistent with the evidence from the TUPS data, shown in Table 3.4, which suggests that there is little movement between wage and self-employment.

Analysis of the Model Solution

The ILFS data suggest that earnings in self-employment exhibit a higher variance than those in wage employment but that the medians of the two distributions are roughly equal. The TUPS OLS regressions suggested earnings in wage employment in small firms was lower than in self-employment after controlling for basic observable human capital. This goes against the view that self-employment is a worse outcome than wage employment but accords with the view of Maloney (2004) that in Latin America, informal salaried workers are generally considered to be the worst off of the employed, below even the informally self-employed.

In the model I have built, average self-employment earnings are much higher than those in wage employment as a result of high productivity individuals choosing self-employment. This is partly as a result of the assumed underlying productivity distribution, which was chosen for its simplicity rather than for its realistic properties. The differential would be much lower with an ability distribution that had more mass near the centre. This would result in a higher proportion of individuals taking both wage and self opportunities and give less weight to those individuals with higher self-employment ability in calculating the average. My model will always have average self-employment earnings higher than the average wage paid, however, without more fundamental changes, some of which I discuss below.

The model I have built above can be thought of as an assignment model in the tradition of Roy (1951), where workers have a choice of sector in which to seek employment. In a review of assignment models, Sattinger (1993) notes that different models generate earnings distributions that can amplify or compress the underlying ability distribution assumed to be present in the population. In my model the identical productivity across all firms compresses the underlying ability distribution very substantially, so that there is much less inequality in wages than in ability. There is no such compression in self-employment, where earnings are equivalent to productivity. The compressed wage distribution is unrealistic but qualitatively, the distribution results are in line with the data I have presented from Tanzania, in that self-employment earnings exhibit a higher variance than wage employment earnings and that the lowest and highest earnings observed are in self-employment. This difference in variance is much higher in my model than in the data, however. One possible way of generating more variation in the wages of those only taking wage opportunities is to increase frictions in the model, which I show below. By changing the functional form of the matching function in wage employment and the exogenous arrival rate of self-employment opportunities, it is possible to increase frictions, make differences in unemployment income more important and thus increase the variation in wages paid.

Although I have shown some weaknesses in my model, it does generate realistic predictions for the transitions between wage and self-employment and the relative size of the variances of earnings in the two sectors without imposing the cost of extra complexity. It also compares well with the Albrecht et al. (2009) model which has constant earnings in the informal sector.

3.6.3 Comparative Statics

Having explored the equilibrium outcomes in the baseline model, I now explore how the equilibrium changes as I vary some of the exogenous parameters. To start with, I explore the effect of raising firm productivity, shown in Table 3.6, where y_w varies across the forth row. This could be thought of as exogenous development progress, where firms become more productive. This could possibly as a result of increased product market competition as trade barriers are lowered and imports increase, or as local production by more productive foreign-owned firms increases. As firm productivity rises, the percentage of workers willing to take only wage opportunities increases, as does the self-employment ability of the worker with productivity y^*, who is indifferent between unemployment and self-employment. The percentage of individuals willing to take both wage and self-employment opportunities first rises and then falls as productivity increases whilst the percentage of the population in self-employment declines as productivity increases. The variance of wages in wage employment is constant and then decreases at high levels of productivity. Both wage and self-employment average earnings rise as productivity increases: in wage employment this is because firm productivity rises, and in self-employment because rising productivity pushes lower productivity individuals who were self-employed into wage employment and thus raises the average productivity of those still in self-employment. Unemployment decreases as firm productivity rises, mainly as a result of decreased unemployment amongst the low ability types who take only wage employment.

It is also possible to explore how changing the matching function and the exogenous arrival rate of self-employment opportunities affects the wage and self-employment earnings distributions. I noted above that increased frictions would raise the importance of unemployment income and hence possibly generate more variation in the wage employment distribution. Table 3.7 shows the results of varying both the matching function for wage employment and the exogenous arrival rate of self-employment opportunities α. The matching function is of the form $a\theta^{.5}$ and I allow a to take on the values 1,2 and 3. The variation in wages for those who always take only wage employment is proportionately much higher when the matching rate in wage employment is lower. Figure 3.4 shows the cdf for $a = 1$ and $\alpha = .5$. The difference between $w(0)$ and $w(y^*)$ doubles when a is only 1 (compared to $a = 4$ in the baseline equilibrium). In absolute terms it is still low, however, only around a maximum of 5 percent of the variation in abilities of the individuals who take only wage employment. Unfortunately, increased frictions come at a high cost as the unemployment rate that results from large frictions is very high, much higher than what is seen in the data. I discuss other ways of increasing variation in wages in the following section.

3.6.4 Potential extensions

The model I have built is relatively simple since it assumes identical firms and only allows individual heterogeneity in one sector. Despite this, it generates some realistic outcomes in predicting the pattern of transitions, and also the relative size of the variation in wage and self-employment earnings. There are several possible extensions which would enhance the predictions of the model but which I have not pursued in this chapter. Varying firm productivity would be one possibility for improving the model since this would increase the variation in wages across individuals in wage employment. Another way to increase wage variation without increasing the amount of heterogeneity, would be to follow the original Mortensen and Pissarides (1994) model and include random productivity shocks that lower the productivity of the match and hence the wage paid. This is also the way match break up is modeled in the formal sector in Albrecht et al. (2009). This assumption would mean that the matches that exist in the equilibrium have existed for different lengths of time, have experienced different numbers of productivity shocks and will thus pay many different wages, even amongst those having the same initial productivity. This requires introducing an endogenous reservation wage that depends on the tightness of the labour market so that when a productivity shock lowers the wage below some level, the match dissolves and the firm and the worker look for other opportunities. This would be more realistic since it endogenises the separation rate rather than having an exogenous separation rate as in my model.

Productivity shocks or heterogeneous firm productivity would generate more realistic variation in wages but this would not get around high ability individuals still taking only self-employment. One potential way of addressing this issue is to allow for individual heterogeneity in both sectors, as in the Roy (1951) model. There would then be a question of how different assumptions about the correlation of individual ability in the two sectors affects the model's predictions. Positively correlated ability would likely result in many more transitions between wage and self-employment than are shown in the data.

My model has different assumptions about arrivals of opportunities in each sector. The rate of matching for wage employment is assumed to vary with labour market tightness but the arrival rate of self-employment opportunities is exogenously determined. Although I have not presented any empirical evidence of this, it would seem preferable to have the self-employment opportunity arrival rate also depend on labour market tightness. It seems intuitive that the self-employment arrival rate is higher the tighter the labour market is, although it is not clear what effect this would have on the uniqueness of the equilibrium in the current model.

The question of the origins of the firms that workers match with has not been addressed in this chapter. This is an important issue in a labour market where self-employment is so prevalent. Future work should incorporate the possibility of the self-employed owning firms and employing other workers, as in Banerjee and Newman's (1993) model of occupational choice and economic development.

3.7 Conclusion

In this chapter I have presented some stylised facts from the Tanzanian labour market and built a model to explain these. I have moved beyond a focus on average wage differentials in the choice between wage and self-employment in Tanzania and presented a model which posits a central role for individual heterogeneity in both the employment choice and earnings determination. I have shown that there is a large amount of heterogeneity in earnings both within and between what have traditionally been called the formal and informal sectors. I have also shown that there are also very few transitions between small firm wage employment and self-employment, despite these having traditionally been lumped together as the informal sector. I have also presented some evidence that small firm employees earn less than the self-employed in Tanzania, challenging the notion that self-employment is a residual sector for those unable to find wage employment.

The model of the urban Tanzanian labour market that I have built can explain some of these outcomes, using a matching framework, and allowing for variation in self-employment productivity that implies an opportunity cost of taking wage employment which is increasing in self-employment productivity. Workers' outside options include this opportunity cost and so bargained wages also vary as a result, despite individuals having constant productivity in wage employment.

The model generates distributions of earnings in wage and self-employment that exhibit characteristics that are broadly similar to the empirical distribution, including higher variance in self-employment earnings than wage employment earnings. The transitions predicted by the model are also similar to those observed empirically, in that there is little movement between wage and self-employment, as a result of most individuals in the population working only in the sector where they have a comparative advantage. As a result of selection by high ability individuals into self-employment, the model I have built generates average earnings in self-employment that are much higher than those in wage employment. This is not too dissimilar from the evidence I presented that earnings in small firms are roughly 15 percent lower than in self-employment.

I have suggested several ways of increasing the complexity of the model and generating patterns more similar to those observed empirically. In particular, allowing heterogeneous firms, a more complicated ability distribution in self-employment and possibly adding ability in wage employment would generate outcomes closer to those observed in the data. I believe the strength of the model is its relative simplicity in capturing the distinction between wage and self-employment in the matching framework and its prediction of some of the stylised facts from the urban Tanzanian labour market.

Table 3.1: Urban Employment by Sector and education

	2001 N	2006 N	Pooled < O levels	O levels or Higher
Employment Categories				
Self Emp no Employees	2168	3294	4985	477
Self Emp with Employees	458	522	729	251
Small Firm Wage Employment	470	660	1012	118
Large Firm Wage Employment	238	718	681	275
Public Employment	1080	1055	868	1267
Agriculture	1853	2785	4530	108
Total	6815	9736	13732	2819

Source: ILFS 2001 and 2006. O levels represents 11 years of education under the new education system and 12 under the old system

Table 3.2: Median Monthly Earnings

	2001	2006
Employment Categories		
Self Emp no Employees	30000	29464
	(98007)	(172078)
Self Emp with Employees	78571	56250
	(492279)	(259446)
Small Firm Wage Employment	30000	28125
	(42826)	(69443)
Large Firm Wage Employment	45000	45313
	(49195)	(136469)
Public Employment	78000	75000
	(192792)	(258443)
Frequency	4962	6951

Standard Deviations in Parenthesis. Earnings expressed in monthly constant 2001 Tanzanian Shillings. The average exchange rate was Tsh 916 per US dollar in 2001 and 1261 in 2006 (International Monetary Fund 2010)

Table 3.3: OLS earnings regressions from the TUPS

	2004	2005	Pooled
	(1)	(2)	(3)
Total Exp	0.024**	0.014	0.027***
	(0.012)	(0.01)	(0.008)
$\exp^2/100$	-.030	-.015	-.036***
	(0.021)	(0.019)	(0.013)
Ln(Hours)	0.025	0.298**	0.028
	(0.117)	(0.136)	(0.073)
Ln(firsiz)	0.18***	0.168***	0.149***
	(0.033)	(0.048)	(0.022)
Ln(emps)	0.615***	0.492***	0.555***
	(0.123)	(0.135)	(0.076)
Tenure	-.006	0.017	0.006
	(0.013)	(0.012)	(0.007)
$\text{Tenu}^2/100$	0.011	-.040	-.012
	(0.033)	(0.037)	(0.02)
Male	0.217***	0.299***	0.313***
	(0.065)	(0.065)	(0.044)
Educ	0.0004	-.008	0.01
	(0.024)	(0.026)	(0.019)
$\text{Educ}^2/100$	0.634***	0.496***	0.436***
	(0.183)	(0.172)	(0.131)
Pub Ent	-.085	0.193	0.112
	(0.215)	(0.236)	(0.135)
Civ Ser	0.985***	0.908***	0.826***
	(0.17)	(0.125)	(0.073)
Priv Wag	-.241*	-.177	-.121
	(0.124)	(0.116)	(0.074)
Const.	9.626***	8.735***	9.634***
	(0.483)	(0.569)	(0.316)
Obs.	568	579	1643
R^2	0.354	0.35	0.327

Source: Tanzanian UPS 2004-2005. A year dummy for 2004 was included but not reported.

Table 3.4: Occupation Transitions

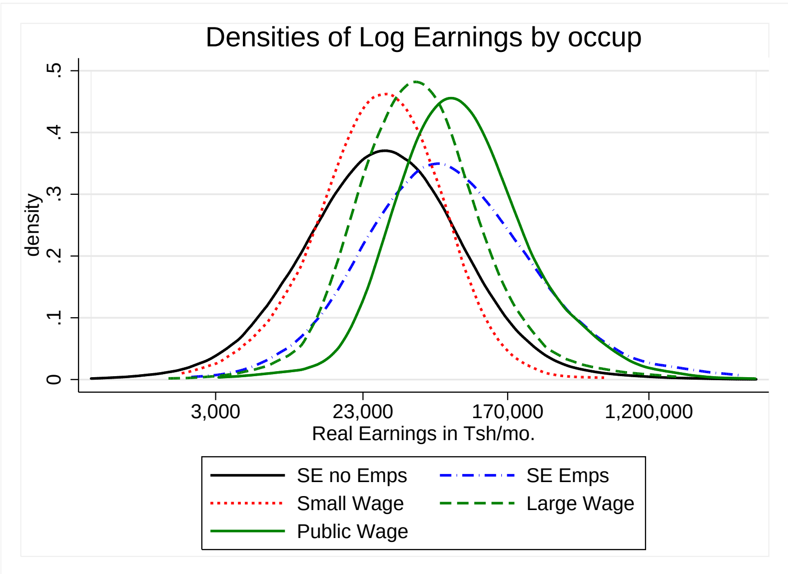
	Occupation						
Occupation	S/E no emps	S/E with emps	Small firm	Large firm	Public	none	Total
	%	%	%	%	%	%	%
S/E no emps	87	4	1	0	1	7	100
S/E with emps	43	48	0	2	0	7	100
Small firm	2	2	49	22	10	15	100
Large firm	2	0	6	34	52	6	100
Public	4	0	4	11	70	11	100
none	55	16	2	2	4	20	100
Total	55	9	6	8	13	9	100

Source: 2004 and 2005 rounds of the Tanzanian Urban Panel Survey. Total number of observations observed in both 2004 and 2005 in this table is 457

Table 3.5: Model Parameters

Parameter		Description
Exogenous	r	Interest rate
	y	Self-employment productivity
	ky	Unemployment income
	α	Matching rate in self employment
	m	Matching function in wage employment
	y_w	Firm productivity
	q_e	Wage job destruction rate
	q_s	Self-employment job destruction rate
	c	Cost of holding a vacant job open
	γ	Bargaining power parameter
Endogenous	V_u	Value of unemployment asset
	V_s	Value of self-employment asset
	V_e	the value of wage employment asset
	θ	Tightness of the labour market
	w	Wage in wage employment
	u	Unemployment rate
	v	Vacancy rate
	$\prod_e$	Value of a filled job asset
	$\prod_v$	Value of a vacant job asset

Figure 3.1: EARNINGS DISTRIBUTIONS, ILFS



Source: pooled earnings data from Tanzanian ILFS 2001 and 2006. The kernel density was estimated using Stata.

Table 3.6: VARYING FIRM PRODUCTIVITY

θ	0.18534	0.32621	0.50762	0.73923	1.039	1.4413	1.9975	2.3829
y^*	0.20731	0.29917	0.39364	0.49024	0.5888	0.68945	0.79211	0.88912
y^{**}	0.33649	0.44866	0.56082	0.67299	0.78515	0.89731	1	1
y_w	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$w(0)$	0.24153	0.335	0.43038	0.5273	0.62564	0.72545	0.8267	0.92439
$w(y^*)$	0.25365	0.34958	0.44682	0.54512	0.6444	0.74473	0.84606	0.94456
$w(y^{**})$	0.3	0.4	0.5	0.6	0.7	0.8	0.89765	0.971
avg wage	0.25372	0.34859	0.44497	0.54256	0.64125	0.74111	0.8417	0.93632
wmass 1	0.79041	0.8061	0.81696	0.82535	0.83233	0.83854	0.85003	0.92078
wmass2	0.20959	0.1939	0.18304	0.17465	0.16767	0.16146	0.14997	0.079219
avg self earnings	0.62249	0.67494	0.72783	0.78045	0.83155	0.87716	0.89606	0.94456
smass 1	0.11543	0.14093	0.17343	0.21989	0.29683	0.46044	1	1
smass2	0.88457	0.85907	0.82657	0.78011	0.70317	0.53956	0	0
u	0.1538	0.1454	0.13466	0.12211	0.10797	0.092203	0.07572	0.072335
u1	0.22502	0.17956	0.14926	0.12693	0.10924	0.0943	0.081256	0.07491
u2	0.095748	0.086436	0.078741	0.072055	0.065987	0.060225	0.054625	0.051682
u3	0.14286	0.14286	0.14286	0.14286	0.14286	0.14286	0	0

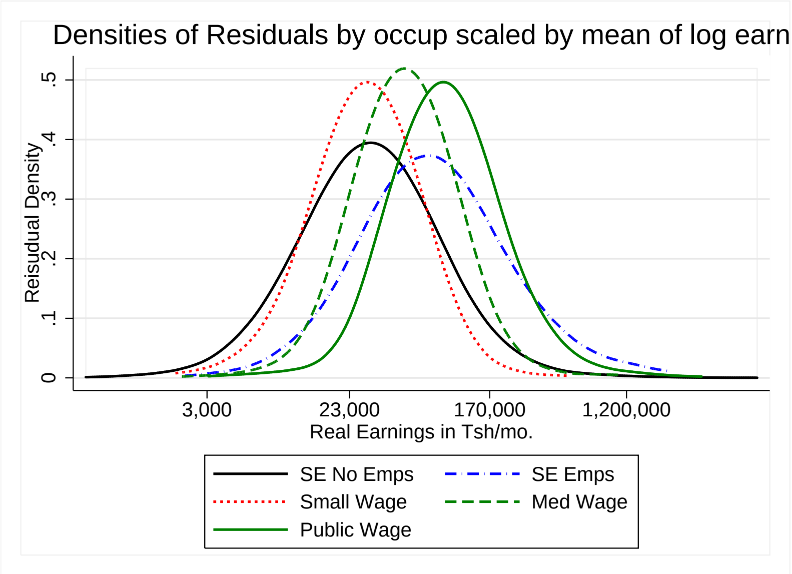
Note: each column represents a model solution, which changes as firm productivity varies, shown in the forth row. wmass1 is the percentage of wage employees who only take wage employment in the equilibrium and wmass2 is the percentage of wage employees who also take self-employment in the equilibrium. smass1 is the percentage of the self-employed who also take wage employment in the equilibrium and smass2 is the percentage of the self-employed who only take self-employment in the equilibrium. u is total unemployment, $u1$ is the unemployment rate for $y < y^*$ types, $u2$ is the unemployment rate for $y^* < y < y^{**}$ types and $u3$ is the unemployment rate for $y > y^{**}$ types.

Table 3.7: VARYING THE LEVEL OF FRICTIONS IN WAGE AND SELF EMPLOYMENT

θ	0.2253	0.36047	0.39689	0.28277	0.43905	0.476	0.35455	0.53396	0.57886
y^*	0.19068	0.30465	0.35526	0.20425	0.31625	0.36443	0.21804	0.32747	0.37387
y^{**}	0.78947	0.78947	0.78947	0.57223	0.57223	0.57223	0.54126	0.54126	0.54126
α	0.5	0.5	0.5	2.5	2.5	2.5	4.5	4.5	4.5
m	1	2	3	1	2	3	1	2	3
y_w	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$w(0)$	0.32536	0.38048	0.40803	0.33147	0.38661	0.41324	0.3378	0.39264	0.4187
$w(y^*)$	0.34534	0.40232	0.42763	0.35212	0.40812	0.43221	0.35902	0.41373	0.43694
$w(y^{**})$	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
avg wage	0.39428	0.42436	0.44095	0.37085	0.4118	0.43245	0.36665	0.41207	0.43376
wmass 1	0.32514	0.44845	0.49732	0.65516	0.74539	0.77577	0.77506	0.83459	0.85392
wmass2	0.67486	0.55155	0.50268	0.34484	0.25461	0.22423	0.22494	0.16541	0.14608
avg self earnings	0.62824	0.71696	0.76048	0.61813	0.68584	0.71516	0.61957	0.6815	0.70791
smass 1	0.65856	0.51134	0.41646	0.42221	0.29332	0.22328	0.38636	0.26505	0.20034
smass2	0.34144	0.48866	0.58354	0.57779	0.70668	0.77672	0.61364	0.73495	0.79966
u	0.40611	0.30497	0.25471	0.22237	0.18752	0.1627	0.17428	0.14589	0.12455
u1	0.513	0.29398	0.20921	0.48461	0.27394	0.19457	0.45644	0.25491	0.1797
u2	0.33906	0.22719	0.17301	0.14157	0.1156	0.098623	0.089358	0.077382	0.068658
u3	0.5	0.5	0.5	0.16667	0.16667	0.16667	0.1	0.1	0.1

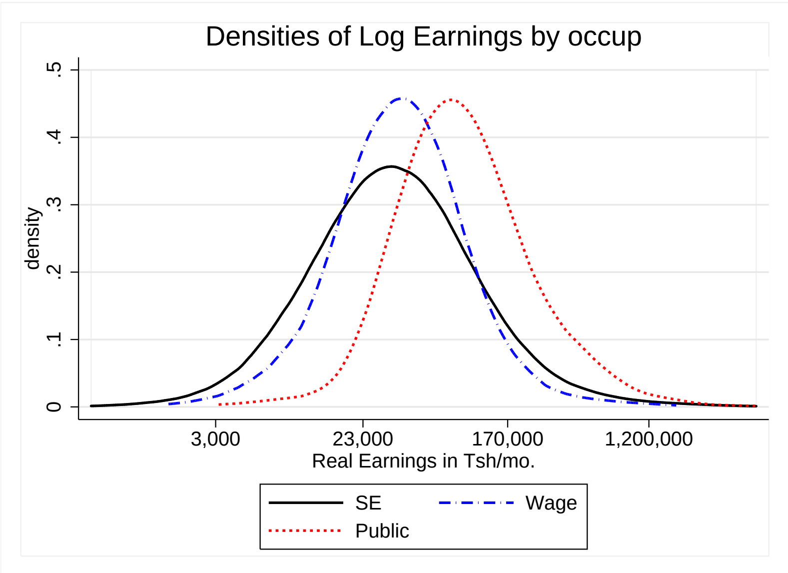
Note: each column represents a model solution, which changes as frictions vary, shown in the forth and fifth rows. wmass1 is the percentage of wage employees who only take wage employment in the equilibrium and wmass2 is the percentage of wage employees who also take self-employment in the equilibrium. smass1 is the percentage of the self-employed who also take wage employment in the equilibrium and smass2 is the percentage of the self-employed who only take self-employment in the equilibrium. u is total unemployment, $u1$ is the unemployment rate for $y < y^*$ types, $u2$ is the unemployment rate for $y^* < y < y^{**}$ types and $u3$ is the unemployment rate for $y > y^{**}$ types.

Figure 3.2: EARNINGS DISTRIBUTIONS, WITH CONTROLS FOR HUMAN CAPITAL



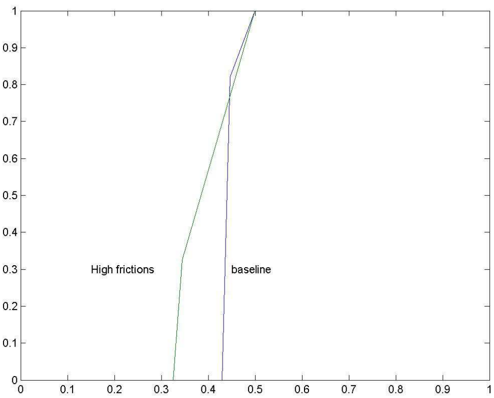
Source: pooled earnings data from Tanzanian ILFS 2001 and 2006. The kernel density was estimated using Stata.

Figure 3.3: EARNINGS DISTRIBUTIONS, ILFS



Source: pooled earnings data from Tanzanian ILFS 2001 and 2006. The kernel density was estimated using Stata.

Figure 3.4: Wage CDFs for baseline and high frictions model



3.A Appendix

3.A.1 Measuring Inflation

Price increases over the period makes comparing incomes at different times more difficult. A consumer price index (CPI) is collected monthly by the Tanzanian National Bureau of Statistics and is shown in Table 3.8. It indicates there was inflation of around 31 percent over the 5 year period between the two surveys. However some concerns have been raised about its accuracy (National Bureau of Statistics, Tanzania 2007)[9]. It is a Laspeyres index based on the 2000 Household Budget Survey, which means it does not take into account changes in consumption patterns that occurred post-2000. If relative prices or patterns of expenditure have changed over the period then this will not be an accurate measure of how the prices households face have changed (Deaton and Tarozzi 2000).

It is possible to calculate alternative price indices based on the Household Budget Surveys of 2000 and 2007, nationally representative samples of households in these years, which collected prices faced by households and household expenditure. Using these data a Fisher ideal index (which measures price changes over an average consumption basket across the two survey years (Deaton and Tarozzi 2000)) is constructed for both food and non food inflation in the 2007 HBS report (National Bureau of Statistics, Tanzania 2007). This suggests that food inflation was actually nearly 100 percent over this 7 year period, whilst the CPI puts food inflation at 52 percent over the same period. A similar trend is found in non-food inflation: the HBSs indicate inflation was over 100% but the CPI implies it was

[9]NBS collects its own CPI data but outsourced the 2007 HBS analytical report to a group of UK researchers, which explains how two NBS sources give different numbers for inflation.

29 percent (National Bureau of Statistics, Tanzania 2007). A Fisher ideal index is then calculated using weights for food and non-food inflation of 0.72 and 0.28 respectively, based on the share of food and non-food expenditure for the poorest 25 percent. The resultant index of price increases between 2000 and 2007 is 1.93, compared to 1.47 using the CPI.

Unfortunately the HBS years do not correspond to the years the ILFS was conducted, meaning some interpolation will be required to obtain the estimated price inflation over the period. To do this I assume a constant rate of inflation over the period and I then wish to find a π that satisfies $(1+\pi)^7 = 1.93$. This generates a value of π of 9.8 percent compared to an average yearly inflation figure of roughly 5.5 percent from the CPI. This interpolation means I estimate that the increase in prices between 2001 and 2006 was 60 percent, nearly double the 31 percent calculated from the CPI. Despite the assumption of constant inflation and the weighting of the index using the poorest 25 percent of the population, given the problems with the CPI, I believe this is more reliable than the CPI data from the National Bureau of Statistics and this is what I use to deflate earnings in the surveys I use.

3.A.2 The Evolution of the value of the asset

Following Cahuc and Zylberberg (2004) I provide an intuitive explanation for the value of an asset which gives a flow income $\omega(x)$ over a small interval of time dt and which can evolve to a new state with net present value $\bar{\Pi}(t+dt)$ over the small time interval dt with a probability of $\lambda(t)dt$. It is possible to write the value of the asset as

$$\Pi(t) = \frac{1}{1+rdt}(\omega(t) + \lambda(t)dt\bar{\Pi}(t+dt) + (1-\lambda(t)dt)\Pi(t+dt)) \qquad (3.40)$$

Rearranging it is possible to obtain

$$r\Pi(t) = \omega(t) + \lambda(t)(\bar{\Pi}(t + dt) - \Pi(t + dt)) + \frac{\Pi(t + dt) - \Pi(t)}{dt} \tag{3.41}$$

Letting $t \to 0$ we then have

$$r\Pi(t) = \omega(t) + \lambda(t)(\bar{\Pi}(t) - \Pi(t)) + \dot{\Pi}(t) \tag{3.42}$$

The time derivative of the asset $\dot{\Pi}(t)$ is assumed to be zero in the asset equation for a job in my model and equation (3.42) is the basis of the asset equations I use in this chapter.

Table 3.8: CPI as measured by the National Bureau of Statistics

2000	2001	2002	2003	2004	2005	2006	2007
77.9523	81.9648	86.3236	90.9018	95.2067	100	107.251	114.786

Source: International Monetary Fund (2010)

3.A.3 Human Capital Regressions

Table 3.9: Basic Human Capital Regressions from the 2006 ILFS

	occup1	occup2	occup3	occup4	occup5
	(1)	(2)	(3)	(4)	(5)
male	0.614	0.478	0.435	0.308	0.09
	(0.029)***	(0.08)***	(0.053)***	(0.054)***	(0.038)**
age	0.025	0.021	0.049	0.065	0.069
	(0.012)**	(0.029)	(0.019)**	(0.019)***	(0.017)***
agesq	-.029	-.009	-.042	-.062	-.052
	(0.014)**	(0.034)	(0.024)*	(0.023)***	(0.02)***
years of education	0.029	0.027	0.092	-.002	0.045
	(0.014)**	(0.04)	(0.027)***	(0.025)	(0.025)*
educsq	0.251	0.4	0.006	0.698	0.311
	(0.112)**	(0.24)*	(0.194)	(0.144)***	(0.122)**
D2001	0.022	0.228	0.16	-.089	0.119
	(0.029)	(0.075)***	(0.048)***	(0.05)*	(0.036)***
Obs.	4429	831	875	886	1646
R^2	0.131	0.16	0.193	0.31	0.259
F statistic	110.921	26.227	34.523	65.908	95.572

Notes: dependent variable is log(income) in all columns. *, **, *** denote significance at 10%, 5% and 1% levels.

Chapter 4

Estimating the Returns to Education in Tanzania

4.1 Introduction

The magnitude and shape of the returns to education is an important research area in labour and education economics and has stimulated a large literature in applied econometrics. Education is considered a crucial tool to raise incomes and thus quantifying returns at different levels of education is highly relevant for policy makers' decisions about how to invest in education. In the last forty years economists have become increasingly aware of the difficulties of estimating the causal effect of education on individual earnings. A major insight from this research is that unobserved individual heterogeneity can confound attempts to estimate the return to education, if unobservably more able individuals take more education and benefit more from it. In this paper I attempt to explore this insight and explore its implications for the problem of estimating the returns to education in Tanzania.

Linear estimates of the returns to education have traditionally been considered a good approximation of the shape of the earnings function, with the returns to education generally estimated to be around 6-7 percent for developed countries and slightly higher in developing countries. One of the major advances made in applied labour economics in recent years has been to recognise and identify the biases present in the estimates of the return to education, due to the endogeneity of schooling choices. This endogeneity results from heterogeneity in individual ability, both in schooling and in earning income, since higher ability individuals are likely to obtain more education and could also benefit more from it. This heterogeneity makes it much more difficult to obtain the causal effect of education on earnings. A discussion of the implications of this heterogeneity for estimation techniques and for the parameter(s) one wishes to estimate is an important contribution of this chapter.

I consider a model of schooling choice in which this choice is determined by its costs and benefits, which vary by individual. I discuss the applicability of the model to developing countries and its implications for estimating the returns to education. I then use two large sample cross-sectional data sets from Tanzania to estimate the returns to education. I consider only the private, pecuniary benefits of education and disregard other possible non-pecuniary or social benefits such as improved health (Söderbom et al. 2006) or more effective civic institutions (Walter 2007). In my estimation of the returns to educational investment I also disregard the costs of education that are necessary to calculate a true return on the educational investment. In doing so I follow almost all of the recent literature on estimating the returns to education.

I attempt to resolve concerns about the effects of unobserved heterogeneity by using exogenous variation in the number of years of education individuals obtained before, and after, a significant reform in the Tanzanian education system in the mid 1960s. I begin with Ordinary Least Squares (OLS) estimates of the returns to education, which have been shown to have a number of biases and which I discuss. I then use the control function approach and two-stage least squares and discuss what it is possible to estimate with each, as well as the assumptions required to do so, in the case of the Tanzanian education reform.

In section 4.2 I discuss the literature that has estimated the returns to education and outline the problems involved. In section 4.3 I set out a model of education and discuss its relevance for the Tanzanian economy. Section 4.4 explores the implications of individual heterogeneity for attempts to estimate the returns to education. In section 4.5 I outline the reform in the Tanzanian education system that I exploit as a source of exogenous variation in education. In section 4.6 I describe the data source I use to estimate the returns to education in Tanzania, section 4.7 describes my estimation and results and section 4.8 concludes.

4.2 Challenges in estimating the returns to education

A framework for thinking about investment in, the production of, and the economic returns to, what was first called 'human capital' by Pigou (1928), as noted in Polachek (2007), was developed in the the work of Schultz (1958), Mincer (1958), Schultz (1959), Becker (1962) and Mincer (1973). The explosion of subsequent research has led to the phrase "human capital" being widely adopted outside economics and has dramatically changed the way policy makers think about the provision and subsidisation of education in the last 50 years (Becker 1993). Mincer's (1973) work was a landmark study. Mincer showed that earnings are correlated with education levels at an individual level, and helped to launch a vast amount of research in applied microeconometrics and labour economics, the objective of which has been to estimate the returns to education. One of the results of this explosion of research has been the incorporation of human capital into the growth theory literature (Aghion and Howitt 1998), and returns to education have been estimated at the macro level (Krueger and Lindahl 2001). In this paper, however, I estimate the returns to education at the individual level.

Economists have long realised that the observed correlation between education and earnings, at the individual level, may not be a causal effect of education. Educational choices may be determined by an individual's taste for, and ability in schooling, as well as the individual's ability in the labour market. This endogeneity of education is the key problem faced by a researcher wishing to explore the causal effect of education on earnings.

One of the early concerns was the possibility of correlation between education and an unobserved and individual-specific constant in the earnings equation, that has subsequently been labeled "ability bias". Belzil (2006) notes that earlier studies attempted to control for ability directly by using measures of IQ. Another way of solving this problem was to use data on twins, who were argued to have the same innate ability, as pioneered by Taubman (1976). Ashenfelter and Krueger (1994) note that some of these studies found estimates of returns to education that were higher than those estimated using OLS, leading Griliches (1977) to argue that ability bias was actually small and that the true effect of education on earnings might even be larger than that estimated by OLS, because of the attenuation bias introduced by measurement error in education.

In the 1990s there was a move towards using aspects of the educational system that generated exogenous variation in education as instrumental variables. The most famous of these is probably the Angrist and Krueger (1991) study, using quarter of birth as an instrument for education, which provided exogenous variation in schooling levels due to compulsory school leaving laws in the US. In another important paper, Card (1995) uses proximity of an individual's home to a college as an instrument for educational attainment. In these and many other studies, estimates of the returns to education using IV were also found to be higher than OLS estimates. Card (2001) notes this was generally taken to imply that ability bias was small, as Griliches (1977) had argued, that the lower OLS estimates reflected measurement error in education, and that IV corrected the attenuation bias introduced by this measurement error.

The possibility that there may also be individual heterogeneity in the returns to education has been recognised more recently by economists working on the returns to education. The concern about the impact of heterogeneity in the returns to education on estimates of this return has shifted the methodology used in estimating the returns to education towards that used in the treatment effects literature, which places a strong emphasis on the heterogeneity of responses to treatment. In economics this literature originally focused on evaluating job training programs (Ashenfelter 1978), but has rapidly expanded its reach to the fields of development economics, immigration, public finance and the returns to schooling, amongst others (Angrist and Pischke 2010).

A key implication of heterogeneity in treatment effects, when estimating the returns to education, is that there are several different possible average effects of education that can be estimated. Card (2001) thus attempts to explain IV estimates that are above OLS estimates, not as the result of resolving measurement error-induced attenuation bias, but as a result of estimating the average effect of education for a particular sub-group with higher returns than the population average, those who were induced to change their level of schooling as a result of the particular instrument used in the study.

Card's (2001) explanation for why IV estimates are higher than OLS requires him to assume that returns are concave. Card (2001) argues that the supply side variation in education that is used as an instrument affects those who would have lower education levels. He assumes that these low education individuals have higher returns to education, which implies a concave schooling-earnings profile, and that IV estimates capture their returns to education rather than some average return from the whole population.

In contrast to Card's assumption of concavity, much of the literature estimating returns for the US assumed that returns were approximately linear, following Mincer (1973). In the following section I outline the debate on the shape of the returns to education, particularly within the African context.

4.2.1 The Shape of the Returns to Education

Despite the assumption of linear or concave returns to education in the US in the majority of the literature, this is still an area of active debate. Belzil and Hansen (2002) argue that because unobserved heterogeneity cannot be identified in reduced form models estimated using IV, strong assumptions about the effects of this heterogeneity that are required to estimate reduced form models are not appropriate. Instead, they estimate a structural model of educational choices using maximum likelihood estimation on the National Longitudinal Survey of Youth data from the US and allow highly flexible effects of heterogeneity and schooling. Unlike Card (2001), these authors find highly convex returns to education, with returns below 1 percent per year until grade eleven and exceeding 10 percent only for the final three years of a four year college degree.

In addition to being a question of interest for applied microeconometricians, the shape of the earnings function is also an important question for policy makers in developing countries. The arguments of Psacharopoulos (1994), that returns to education were concave and generated by schooling-induced increases in productivity, were used to justify policies where the largest investments went towards primary education in developing countries, as in the World Bank's (1995) objective of "Education For All". Bennell (1996) argues, however, that many of the studies used by Psacharopoulos (1994) to justify the claim that marginal returns to primary school are higher than those for secondary and tertiary education, are actually of very low quality.

Despite the arguments of Psacharopoulos (1994), there seems to be a shift towards a consensus that returns are convex in developing countries. Convexity in returns to education has also been found in studies for a number of African economies. These include Knight and Sabot (1990) for the 1980s and Söderbom et al. (2006) for the 1990s, in both Tanzania and Kenya, for Tanzania and Ghana in the 2000s (Rankin et al. 2010), for Nigeria in the 1990s and 2000s (Oyelere 2008) and for South Africa in the 1990s and 2000s (Keswell and Poswell 2004).

As well as suggesting that investing only in primary schooling may be bad policy, these findings of convexity raise concerns about the possible influence of some of the biases discussed above. Many of the studies using African data do not account for the endogeneity of education and estimate the returns using OLS, including all of those cited in the review by Keswell and Poswell (2004) for South Africa, as well as Knight and Sabot (1990). It may be that what seems to be convex returns to education could actually just be a form of ability bias, if individuals with higher ability in the labour market also invest in more education, or returns bias, if individuals who benefit more from education obtain more of it.

There are some studies using African data that acknowledge the difficulties individual heterogeneity creates for estimating the returns to education and attempt to solve these. Oyelere (2008) uses a binary instrument, indicating whether or not an individual was exposed to a free education program rolled out across different states, and at different times in Nigeria, in attempting to solve endogeneity problems, but still finds convex returns. The author, however, does not allow for individual heterogeneity in individual returns.

Söderbom et al. (2006) use a control function approach in attempting to solve the endogenous education problem. The authors, following the work of Card (1995), who used proximity to college as an instrument for college attendance, claim to have solved the endogeneity problem, using instruments for education that include distance from primary and secondary schools. Theirs is possibly the only study that allows for heterogeneous returns but they do not find this to be important, and their preferred specification eventually excludes terms that capture this form of heterogeneity. In general, however, a lack of good quality data, as well as the fact that addressing heterogeneous returns is a relatively new area, in which advances are being made rapidly, has meant that there is a paucity of studies using African micro data that acknowledge the difficulties individual heterogeneity causes when estimating the returns to education, or that attempt to solve these.

There is also a lack of clarity in what is being estimated when there is heterogeneity in returns, since the returns may differ for each individual. In the following section I outline a model of educational choice that clarifies the endogeneity concerns I have discussed in this section.

4.3 A Model of Educational Choice

In this section, following Card (2001), I specify a general model of educational choice, modeling the returns to education and its costs, and discuss the model's relevance for the developing country context. Card (2001) assumes risk averse individuals, who have utility function $u()$ that is increasing and concave in consumption, and who have an infinite planning horizon. This planning horizon is assumed to begin at the minimum school leaving age, so Card is assuming that individuals are old enough to process information about their own ability, and the state of the labour market, and to make the sorts of decisions the model requires of them at this time. With high enough poverty levels that a significant number of Tanzanians acquire no schooling at all, one could question the relevance of this model for Tanzania.

Orazem and King (2008) is one of the few models that explains educational attainment in low income countries and this model has similar assumptions about educational choice to Card's (2001). The authors build a model of educational choice in which individuals maximise their utility over a lifetime. Households are included and are assumed to have differential costs of borrowing but the model does not specify the relationship between the household and its members, nor how borrowing by a younger member to pay for schooling investments might be repaid. I thus use Card's (2001) model of individual decision making and highlight issues of relevance to the Tanzanian context.

In this model individuals are assumed to experience a disutility of schooling compared to work $\phi(t)$ that is convex in the schooling level and make a once and for all decision to leave school and begin working. Card specifies life cycle utility as

$$V(S, c(t)) = \int_0^S (u(c(t) - \phi(t)))e^{-\rho t}dt + \int_S^\infty u(c(t))e^{-\rho t}dt. \tag{4.1}$$

An individual is assumed to be able to borrow or lend at a fixed interest rate R, pay tuition costs $T(t)$, earn $p(t)$ in part time earnings and $y(S,t)$ in full time earnings where S is the level of schooling and t is the age of the individual. As a result of the assumption of no credit constraints there is simply a inter-temporal budget constraint, which Card (2001) writes as

$$\int_0^\infty c(t)e^{-Rt}dt = \int_0^S (p(t) - T(t))e^{-Rt}dt + \int_S^\infty y(S,t)e^{-Rt}dt. \tag{4.2}$$

It is then possible to form a Langrangian for the choice of the optimal schooling level and optimal consumption path:

$$\begin{aligned}\Omega(S, c(t), \lambda) =& V(S, c(t)) - \lambda(\int_0^\infty c(t)e^{-Rt} \\ & - \int_0^S (p(t) - T(t))e^{-Rt}dt - \int_S^\infty y(S,t)e^{-Rt}dt).\end{aligned} \tag{4.3}$$

The optimal schooling level is found from the first order condition for S:

$$\Omega_S(S, c(t), \lambda) = \lambda e^{-RS}(MB(S) - MC(S)), \tag{4.4}$$

where

$$MB(S) = \int_0^\infty \partial y(S, S+\tau)/\partial S e^{-R\tau}d\tau \tag{4.5}$$

and

$$MC(S) = y(S,S) - p(S) + T(S) + 1/\lambda e^{-(\rho - R)S}\phi(S). \qquad (4.6)$$

The marginal benefit of schooling is the extra earnings it brings at every date in the future, discounted to the present as in equation (4.5). The marginal cost of schooling is foregone earnings (ameliorated by part-time earnings while in school $p(S)$), tuition costs $T(t)$, and its inherent disutility relative to work, the so-called "psychic costs" of schooling which Cunha et al. (2005) argue are actually an important part of schooling decisions.

Card (2001) assumes that log earnings are additively separable in education and labour market experience, implying that $y(S,t) = f(S)h(t-S)$, where $t-S$ is a rough measure of labour market experience[1]. Now marginal benefits are expressed in equation (4.5) as a function of $y(S, S+\tau)$, so this and the separability assumption imply that $y(S, S+\tau) = f(S)h(S+\tau+S) = f(S)h(\tau)$. Hence it is possible to write the marginal benefit as

$$MB(S) = f'(S)\int_0^\infty h(\tau)e^{-R\tau}d\tau = f'(S)H(R), \qquad (4.7)$$

where $H(R)$ is defined by Card as a decreasing function of the interest rate.

Card (2001, pg. 1130) notes that if "$MC(S)$ rises faster than $MB(S)$, a necessary and sufficient condition for an optimal schooling choice is that $MC(S) = MB(S)$." The assumption that $MC(S)$ rises faster than $MB(S)$ guarantees an interior solution and hence that the equality of marginal costs and benefits is a necessary and sufficient condition for an optimal schooling

[1]This assumes no prolonged periods of unemployment.

choice. This condition implies

$$f'(S)/f(S) = 1/H(R)\{1 + (T(S) - p(S))/f(S) + 1/\lambda e^{-(\rho - R)S}\phi(S)/f(S)\}. \tag{4.8}$$

Card (2001) notes that if the earnings experience profiles are concave then $H(R) = 1/(R - g)$, where g is the constant growth rate of earnings that gives the same value of discounted lifetime earnings as $h(t - S)$. Card (2001) then assumes part time earnings while at school are roughly equal to tuition costs which then implies the optimal schooling decision is given by

$$f'(S)/f(S) = R - g + \rho e^{-\rho S}\phi(S) \equiv d(S). \tag{4.9}$$

The left hand side is the proportional increase in earnings per year for the S^{th} year of schooling. The right hand side is the marginal cost of schooling, which in simpler models is the interest rate, but in this model is modified due to life-cycle earnings growth g and disutility of the S^{th} year of schooling relative to work.

Card (2001) notes that individual heterogeneity can be due to differences in benefits or costs, which he specifies as

$$f'(S)/f(S) = b_i - k_1 S \tag{4.10}$$

for benefits, with $k_1 > 0$, and as

$$d(S) = r_i + k_2 S \tag{4.11}$$

for costs, with $k_2 > 0$. Equation (4.10) implies that returns to education are concave, with the marginal benefit declining with education level. Equation

(4.11) gives a simple linear expression of the disutility of school relative to work that is increasing in the schooling level. r_i represents heterogeneity in R the interest rate or g the growth rate in lifetime earnings as a result of accumulated labour market experience. Card (2001) does not explicitly mention which of these it represents but it certainly does allow for interest rates to differ across individuals. It is assumed to be a random variable. The model could actually thus be interpreted as allowing credit constraints to affect some individuals, as is likely to be the case in Tanzania.

Equating these costs and benefits, Card (2001) obtains the optimal schooling level:

$$S_i = (b_i - r_i)/k, \tag{4.12}$$

where $k = (k_1 + k_2)$ and $k > 0$, because of the assumptions made about k_1 and k_2. Card (2001) notes that his assumptions generate an optimal level of schooling that is linear in the individual specific heterogeneity terms, which plays an important part in his later analysis. An obvious question is why there is no heterogeneity in the $k_1 S$ part of the marginal benefits. Card (2001) does not mention this but it would imply an optimal level of schooling that was non-linear in (what would become) the three individual heterogeneity terms. It would seem this is the reason Card (2001) does not include such heterogeneity in his model. He does allow the heterogeneity in returns to be correlated with schooling, however, so the model does actually allow for a fairly general effect of individual heterogeneity on marginal returns.

Integrating equation (4.10) implies the following model of earnings

$$log y_i = \alpha_i + b_i S_i - \frac{1}{2} k_1 S_i^2. \tag{4.13}$$

The model specified in equation (4.13) ignores determinants of earnings other than experience and education, but I do include these other determinants in my estimation of the returns to education in section 4.7. In the following section I discuss the Card (2001) model, after which I explore the ways researchers have used equation 4.13 to estimate the returns to education, in section 4.4.

4.3.1 Some Remarks on Card's model

Card's (2001) model assumes returns to be a non-linear concave function of education. The assumption of concave returns is not necessary, however, to generate an interior solution in this model. Costs would still rise faster than benefits in the case of increasing convex returns, if $k_2 > k_1$, which could still generate an interior solution[2]. This is relevant because in the estimation section below I find some evidence that there are convex returns to education in Tanzania and I do not want this to invalidate my use of Card's (2001) model, and the estimation techniques it suggests, or those of other authors his model has inspired.

Although Card (2001) makes certain assumptions about the earnings function in his model, including the separability of education and experience and the concavity of the relationship between education and earnings, he is agnostic about how schooling is translated into earnings. Card notes that the returns to schooling may be generated by schooling enhancing productivity, as in Becker's (1962) human capital model, but also mentions that the returns to education may result from its value as a signalling device as in Spence (1973); where individuals obtain costly education to signal their abilities to potential employees, but in which education generates no pro-

[2]This point is also made by Florens et al. (2008)

ductivity enhancing human capital. In this case the returns to education are not necessarily reflective of the skills or productive capacity individuals have acquired through education and Card's (2001) model allows for this possibility. This is relevant in the Tanzanian case, where much of the return to tertiary education seems to come from being able to access public sector employment, implying that tertiary education is not necessarily generating higher productivity levels.

The model I developed in the previous chapter suggests another avenue for education to influence earnings, if education increases self-employment productivity. Although I did not model this explicitly, the model predicts earnings would increase for any individual if their education level was raised. This would occur either through a direct effect of increased productivity and earnings in self-employment, or by raising the outside option for those in wage employment. In the following section I explore some of the different strategies that have been adopted to estimate the returns to education in the presence of heterogeneity in returns to education.

4.4 Strategies for Estimating the Returns to Education

4.4.1 The Treatment Effects Framework

Attempts to estimate equations similar to equation (4.13), which has both an individual-specific intercept term and an individual-specific slope coefficient, have spawned an extremely large body of work, which has become known as the treatment effects literature. As I discussed in section 4.2, this was first developed in models similar to the one presented above but with a focus on an individual specific coefficient on a dummy variable that represented the effect of some "treatment", like participation in a labour market training programme. Much of recent microeconometrics has been concerned with understanding what it is possible to estimate with models of this nature (Heckman 2001). In this section I introduce the treatment effects framework, starting with the case of a simple binary treatment, and then extending this to multiple treatments, as is the case for education.

In a simple binary treatment effects framework, with a binary treatment w, individuals have a potential outcome y_1 under treatment and an outcome y_0 without treatment. The researcher wishes to estimate the difference between these two. Individuals, however, are not observed in both states, so it is impossible to obtain a counterfactual to assess how an individual who benefitted from a particular policy would have fared without the introduction of the policy, and vice versa. The other important issue is that $y_1 - y_0$ is individual specific, which implies that there is self-selection into treatment. Individuals who benefit more from treatment are more likely to obtain treatment and this creates an endogeneity problem when attempting to estimate the effects of treatment.

A key implication of individual differences in outcomes that has been emphasised in the treatment effects literature, is that there are several possible effects one could potentially estimate (Wooldridge 2002). Candidate effects include the average treatment effect (ATE), which is the expected value of the difference between the outcome under treatment and without treatment over the entire population. In the case of a binary treatment this is $ATE \equiv E(y_1 - y_0)$. This has been criticised by Heckman (1997), who argued that this is unhelpful for policy analysis, since, to use Heckman's example, the responses of millionaires would then be included in the effects of a training programme. Wooldridge (2002), however, argues that this is not really a valid objection because one can easily limit the population one is studying.

An alternative is the average treatment effect on the treated (ATET), which is the expected value of the treatment for those who actually took it: $ATET \equiv E(y_1 - y_0|w = 1)$. Another alternative is the local average treatment effect (LATE), first outlined by Imbens and Angrist (1994), which is the average effect of treatment for the members of the population who were induced to change their treatment status as a result of a change in some instrumental variable that affects treatment status.

When moving from a binary to a multi-valued treatment defining both the ATE and LATE can be more complicated. It is not obvious what the counterfactual is because there are many possibilities (Wooldridge 2002). In the case of a so-called a random coefficients model, like equation (4.13) above, the ATE is the average partial effect (Wooldridge 2002). In the model outlined in section 4.3 this would be defined as $\partial E(y|S; a, b)/\partial S = E(b_i)+k_1$, taking the expectation over the entire population (Florens et al. 2008). This implies that there is not one ATE of education, as in the binary treatment case. Instead, the ATE varies with the level of treatment.

4.4.2 The Implications of Heterogeneous Returns for Estimation of the ATE of Education

Having introduced the treatment effects framework and having noted the implications of heterogeneity in treatment, particularly self-selection into treatment and the resulting endogeneity, as well as that there are several possible treatment effects to estimate, in this section I explore the implications of Card's model of heterogeneous returns to education for attempts to estimate the ATE of education.

In using Card's model to estimate the ATE of education, equation (4.13) can be written as

$$logy_i = a_0 + \bar{b}S_i - \frac{1}{2}k_1 S_i^2 + a_i + (b_i - \bar{b})S_i. \qquad (4.14)$$

It is clear from equation (4.14) that self-selection would also be a concern if one used this to estimate the returns to education. The presence of individual heterogeneity in both the intercept and the slope of the earnings function introduces correlation between the error term and years of schooling and means that estimating $E(b_i) + k_1$, the ATE, using OLS is not possible. Finding an estimation procedure that can consistently estimate the ATE would thus be an important step forward.

Instrumental variables is the most commonly used estimation strategy to tackle endogeneity and ability bias. Card (2001) explores this approach in the context of his model of endogenous schooling with heterogeneous returns and assumes the existence of a set of instrumental variables related to the marginal cost of schooling:

$$r_i = Z_i \pi_1 + \eta_i, \qquad (4.15)$$

with η_i containing other unobserved cost and taste factors. The optimal schooling decision (4.9) can then be rewritten as

$$S_i = Z_i \pi_+ \xi_i, \qquad (4.16)$$

where ξ_i is a function of $b_i - \bar{b}$, the parameter governing heterogeneity in returns to education.

Many studies estimating the returns to education incorporate heterogeneity only in the intercept term of the returns to education equation. If this is true and $b_i = \bar{b}$ then IV can be used on the simultaneous equations system (4.14) and (4.16), if Z_i and a_i are uncorrelated. This was how the earlier literature from the 1990s attempted to deal with endogeneity concerns, as was discussed in section 4.2.

IV cannot be used to estimate the ATE of education in the heterogeneous returns case (when b is not constant) without further assumptions, however, since Z_i could be correlated with $(b_i - \bar{b})S_i$ in the income equation, equation (4.14). This understanding, along with a belief that the homogeneity of returns assumption ($b_i = b \; \forall \; i$) is too strong, has meant that the literature estimating the returns to education has followed several different paths. I discuss each of these below and then outline the estimation strategy I follow.

4.4.3 The LATE interpretation of IV

The first possible option is to abandon attempts to estimate an ATE and instead to ask what IV *can* estimate if Z_i is correlated with the unobserved ability terms (Card 2001). This was first investigated by Imbens and Angrist (1994) in the context of a binary treatment and extended for the case of a multi-valued treatment (such as education) by Angrist and Imbens (1995). Imbens and Angrist (1994) suggest an non-parametric alternative to the homogeneity assumption, the monotonicity assumption, which is sufficient to identify a local average treatment effect (LATE), in the case of a binary treatment, and what they term an average causal response (ACR), in the case of a multi-valued treatment (Angrist and Imbens 1995). I follow the subsequent literature in also referring to the ACR as a LATE.

I noted above that the LATE is more complicated in the case of a multi-valued treatment and thus defining it requires slightly different notation.

Assume Z is a binary instrument measuring whether an individual was exposed to the pre- or post-reform educational system. Following Angrist and Imbens (1995), let $S_Z \in \{0, 1, 2, ..., J\}$ be completed years of schooling for $Z \in \{0, 1\}$. Angrist and Imbens (1995) assume that S_Z exists for each person both pre- and post-reform, even though only one S_Z is observed. Thus S_0 is number of years of education pre-reform and S_1 is years of education post-reform. They further assume that the random variables S_0, S_1, Y_0, Y_1, $...Y_J$ are jointly independent of Z, at least after conditioning on covariates, a non-parametric version of the usual uncorrelatedness assumption when using IV. The other assumption is monotonicity: with probability 1 either $S_1 - S_0 \geq 0$ or $S_1 - S_0 \leq 0$ for each person. This means that the instrument does not generate an increase in treatment for some individuals and a decrease for others.

Under these assumptions Angrist and Imbens (1995) then go on to prove that the instrumental variable estimator first suggested by Wald (1940) is a weighted average of the treatment effects in the population:

$$\frac{E[Y|Z=1] - E[Y|Z=0]}{E[S|Z=1] - E[S|Z=0]} = \sum_{j=1}^{J} \omega_j E[Y_j - Y_{j-1}|S_1 \geq j > S_0], \qquad (4.17)$$

where

$$\omega_j = \frac{P(S_1 \geq j > S_0)}{\sum_{i=1}^{J} P(S_1 \geq i > S_0)}. \qquad (4.18)$$

This result is obtained from a very general model. The model allows for heterogeneous effects of education, as well as the possibility of non-linearities in education, and the estimated coefficient on the S term is a weighted average of all these effects. It, however, does not provide direct evidence on the type of heterogeneity in returns or the shape of the earnings function, because the differing treatment effects for all individuals whose treatment status was affected by the reform are aggregated together. This simple result is also obtained without covariates. This is partly addressed by Angrist and Imbens (1995), who allow for discrete covariates in a fully saturated model. The result becomes more complicated with continuous regressors (Abadie 2003), though two-stage least squares provides an good approximation in the binary treatment case (Angrist and Pischke 2009). Card (2001), amongst others, simply assumes this is also true in the multiple treatment case and I follow him in assuming my two-stage least squares estimates approximate the LATE in a model with heterogeneous, multi-valued treatment effects and covariates.

Monotonicity is a crucial assumption in obtaining these results. Angrist and Imbens (1995) show that the estimated treatment effect can be negative or zero if this assumption does not hold, even if the treatment effect for all individuals is positive, and illustrate this in the case of a binary treatment variable and a binary instrument. To do this Angrist and Imbens (1995) compare outcomes for different values of the instrument:

$$\begin{aligned} E[Y|Z=1] - E[Y|Z=0] =& E[Y_0 + (Y_1 - Y_0)S_1|Z=1] \\ & - E[Y_0 + (Y_1 - Y_0)S_0|Z=0]. \end{aligned} \tag{4.19}$$

Since the other random variables are independent of Z by assumption, the mean conditional on Z is the same as the unconditional mean, implying that the right-hand side of equation (4.19) can be written as

$$E[(Y_1 - Y_0)(S_1 - S_0)], \tag{4.20}$$

which, by the law of iterated expectations, means that

$$\begin{aligned} E[(Y_1 - Y_0)(S_1 - S_0)] =& E[Y_1 - Y_0 | S_1 - S_0 = 1] P[S_1 - S_0 = 1] \\ & - E[Y_1 - Y_0 | S_1 - S_0 = -1] P[S_1 - S_0 = -1]. \end{aligned} \tag{4.21}$$

Those who do not change their schooling due to the instrument, those for whom $S_1 - S_0 = 0$, "do not contribute anything to comparisons of average outcomes by instrument status" (Angrist and Imbens 1995, pg 434). Those who contribute include those who are induced by the reform either into treatment, if they were previously not treated, or out of treatment, if they were previously treated. Equation (4.21) shows this average value could be negative or zero, even if treatment effects are positive for each individual. This is why the requirement of monotonicity is so important. In this example monotonicity would imply either $P[S_1 - S_0 = 1]$ or $P[S_1 - S_0 = -1]$ is zero and thus the remaining non-zero term would be the only contributor to the (positive) average value of treatment. A similar argument holds for multi-valued treatments (Angrist and Pischke 2009). I discuss this further in the context of the Tanzanian reform in section 4.6.2.

The LATE has been criticised because it estimates the effect of treatment for those who were induced into changing their education level as a result of the effect of the instrument which was used, as I explained in the previous paragraph, which is generally an unidentifiable sub-population (Heckman 1997). This means that the estimated LATE will vary with the instrument used and that there are potentially many different LATEs that could be estimated, depending on which instruments are available and used. Heckman (1997) further argues that defining a parameter of interest by an instrument not part of the outcome equation is unusual and bad practice, whilst Heckman and Vytlacil (2005) argue that the policy question being investigated when a LATE is estimated is often not clear. Despite this, the LATE is very simple to estimate, allows for heterogeneous treatment effects and requires no parametric assumptions, all of which have made it extremely popular.

As I discussed in section 4.2, Card (2001) uses the LATE to provide an explanation for why many of the IV estimates in the returns to education literature are actually above the OLS estimates. He argues that the educational reforms used as instruments in many studies, such as changes to the costs of education, predominantly affect those with low education and higher marginal costs of education. I have shown above that the LATE is weighted by the returns to the levels of schooling of those induced by the instrument to take more schooling. If these individuals have low schooling and higher returns, as is the case in Card's (2001) model of concave returns to education, then this explains why estimates of the LATE may be higher than the ATE in the population.

4.4.4 Stronger Assumptions for IV

If one does not want to abandon efforts to estimate an ATE, then one option is to find assumptions under which IV still produces a consistent estimate of the ATE of education (Wooldridge (1997), Heckman and Vytlacil (1998), Card (2001), Wooldridge (2003)). It is not impossible for IV to recover a consistent ATE under heterogeneous returns to education, but one needs assumptions that are stronger than the usual uncorrelatedness assumption. Wooldridge (1997) assumes homoscedasticity of returns in a linear returns to education model and shows that the ATE can be consistently estimated if, in addition, the individual heterogeneity terms are assumed mean independent of Z_i. In general for a random variable u to be mean independent of X it must be the case that $E(u|X) = E(u)$. It is a stronger assumption than uncorrelatedness ($E(uX) = 0$) and implies that not only is u uncorrelated with X but also that u is uncorrelated with any other non-linear function of X. This effectively implies that the conditional mean of S_i is correctly specified in the first stage regression (Wooldridge 2002), a fairly strong assumption.

In terms of the model I set out above the assumptions required are that $E[\xi_i^2|Z_i] = \sigma_\xi^2$ (homoscedasticity) and that the conditional expectation of $b_i - \bar{b}$ is linear in the schooling residual:

$$E[(b_i - \bar{b})|S_i, Z_i] = \psi_0 \xi_i. \tag{4.22}$$

This means that the conditional mean of the ability that affects the returns to education is not correlated with either education or the instruments once the schooling residual is controlled for. Under these assumptions the conditional expectation of the final term in equation (4.14) can be shown to equal $\psi_0 \sigma_\xi^2$, so that

$$E[logy_i|Z_i] = a_0 + \bar{b} Z_i \pi - \frac{1}{2} k_1 (Z_i \pi)^2 + (\psi_0 - \frac{1}{2} k_1) \sigma_\xi^2 \tag{4.23}$$

The last term in the conditional expectation is a constant and does not depend on Z_i which implies that the ATE can be consistently estimated.

Unfortunately the assumptions put forward by Wooldridge (1997) are unlikely to be satisfied in the case of an instrument that affects the supply side of schooling. Card (2001) argues that the effect of these reforms is to alter the mapping between ability and schooling, implying that there will be a correlation between $(b_i - \bar{b})S_i$ and Z_i, and violating assumption (4.22) above. In addition, he argues that the effect of reforms that lower the marginal cost of schooling will be to lower the effect of cost differences r_i (in equation (4.15) above), and that this will lower the variance in ξ_i post-reform, which violates the homoscedasticity assumption Wooldridge (1997) makes.

Wooldridge (2003) provides some slightly weaker assumptions than his earlier paper, under which two-stage least squares consistently estimates the ATE, but also acknowledges that one of his key assumptions may not hold for the same reasons pointed out by Card (2001) and explained above. As a result of this research I do not interpret my two-stage least squares estimates of the returns to education in section 4.7 as average treatment effects. Instead, following Angrist and Imbens (1995) and Card (2001), I interpret these as LATEs.

4.4.5 The Control Function Approach

The final option I discuss is the control function approach to estimating an ATE, Card (2001), Blundell et al. (2005), which has an advantage over IV in being able to consistently estimate an ATE, even when the assumptions required for IV to consistently estimate an ATE do not hold. This approach does, however, require parametric assumptions about the correlation between unobservables in the earnings (4.14) and schooling (4.16) equations and then the inclusion of functions in the earnings equation that capture these assumed relationships (Card 2001). This solution was first introduced by Garen (1984), who showed the exact form the control function should take in the binary treatment case.

Garen (1984) also showed that there was a similar approximate solution when the treatment variable was assumed continuous and when there was no heterogeneity in treatment effects. This was expanded by Card (2001), whose solution can give a consistent estimate of the ATE even in the presence of heterogeneity in treatment effects and a change in the mapping between schooling and ability, which Card (2001) argues will make the estimates of the IV method proposed in Wooldridge (1997) inconsistent. This does, however, come at the cost of extra parametric assumptions compared to the IV approach discussed above.

The assumptions required for consistent estimation of the average treatment effect of schooling in the continuous education model, following Card (2001), are that a_i, b_i and ξ_i are mean independent of Z_i, the reform used as the instrument, and also that the conditional expectations of the unobserved ability components are linear in the residual (ξ_i) from the first stage schooling regression. These can be written as

$$E[b_i - \bar{b}|S_i, Z_i] = \psi_0 \xi_i \tag{4.24}$$

and

$$E[a_i|S_i, Z_i] = \lambda_0 \xi_i. \tag{4.25}$$

The first of these is identical to the assumption required for IV to estimate the ATE, equation (4.22) above. These equations imply that the earnings ability components are not correlated with schooling or the instrument but are linear functions of the schooling residual, which can be thought of as a combination of schooling ability and unobserved schooling costs.

These assumptions imply

$$E[logy_i|S_i, Z_i] = a_0 + \bar{b}S_i + \frac{1}{2}k_1 S_i^2 + \lambda_0 \xi_i + \psi_0 \xi_i S_i, \tag{4.26}$$

which can then be estimated using the predicted residual from the first stage regression of S_i on Z_i, see equation (4.16) above. Since one is estimating the residual and then using this estimate in the second stage, the standard errors in the second stage regression are not correct, as they do not take into account the uncertainty in the estimation of the residual. Bootstrapping can, however, generate the correct standard errors (Wooldridge 2002). The restriction that the conditional expectation is linear in the schooling residual can be relaxed by adding non-linear terms in the schooling residual to equation (4.26), as in Söderbom et al. (2006).

The control function requires an extra linear conditional expectation assumption compared to the IV method discussed above, equation (4.25). I noted above that when there are changes in the mapping between abilities and schooling, that a major reform on the supply side is likely to bring, equation (4.24) is unlikely to hold. A simple extension when there is a binary reform is to allow the conditional expectation to be linear before and after the reform but to have a different coefficient for each period (Card 2001). Thus equations (4.24) and (4.25) become

$$E(b_i - \bar{b}|\xi_i) = \psi_{00}(1 - Z_i))\xi_i + \psi_{01}(1 - Z_i))\xi_i \tag{4.27}$$

and

$$E[a_i|S_i, Z_i] = \lambda_{00}(1 - Z_i)\xi_i + \lambda_{01}(1 - Z_i)\xi_i. \tag{4.28}$$

These imply that

$$\begin{aligned} E(log(y_i|S_i, Z_i)) = & a_0 + \bar{b}S_i - 1/2k_1S^2 + \lambda_{00}\xi_i - (\lambda_{01} - \\ & \lambda_{00})Z_i\xi_i + \psi_{00}S_i\xi_i + (\psi_{01} - \psi_{00})Z_iS_i\xi_i. \end{aligned} \tag{4.29}$$

In section 4.7 below I estimate an ATE using the control function approach and equation (4.29), as well as using two-stage least squares to attempt to estimate a LATE.

An important conclusion from the discussion above is that if one wishes to move beyond estimating the LATE with IV, and instead use the control function approach to find the ATE, then this comes at the cost of imposing fairly strong parametric assumptions. Some of these can be relaxed, but even then other strong assumptions are still required. In particular, mean independence of Z from ξ implies that one has correctly specified the conditional mean of the schooling equation (Wooldridge 2002). This is much stronger than the assumption of uncorrelatedness of Z and ξ, that is required for IV to consistently estimate the return to education when there is no heterogeneity in returns.

To see how mean independence might be violated, assume that schooling is determined by the instrument Z, as well as another variable X that is not correlated with schooling ability but which is unobserved to the econometrician:

$$S_i = Z_i\pi + X_i\gamma + \xi_i. \tag{4.30}$$

As before,

$$logy_i = a_0 + \bar{b}S_i - \frac{1}{2}k_1 S_i^2 + a_i + (b_i - \bar{b})S_i. \tag{4.31}$$

When X_i is an omitted variable in the first stage, the composite unobservable schooling error term from the first stage regression is

$$\hat{\xi}_i = X_i\gamma + \xi_i. \tag{4.32}$$

X_i being an omitted variable implies $E(\hat{\xi}_i|Z_i) \neq 0$, if X_i is correlated with Z_i, and thus mean independence is violated.

4.5 Tanzanian Education Reform

In this section I describe an education reform introduced in Tanzania in the 1960s that provides a source of exogenous variation in education which can potentially be used to consistently estimate the average treatment effect of education when there is heterogeneity in returns. Tanzania gained its independence from Britain in 1961 and inherited an education system consisting of four years of primary school and four years of middle school. Another four years of secondary school were required to obtain O levels and a further two were required to obtain A levels.

The "Five Year Plan" for 1964-1969 was the first attempt at development planning by the new ruling party, the Tanganyika African National Union (Morrison 1976). This plan had goals of self sufficiency in trained 'manpower', increasing per capita income from 20 to 45 pounds per year, and an increase in life expectancy from 35 to 50 years, all by 1980. A major reform in the mid 1960s, undertaken as part of the first "Five Year Plan", did away with middle school and made primary school seven years, replacing the previous system of four years of primary and then four years of middle school (Morrison 1976). The reform was implemented over a three year period between 1965 and 1967, first in cities, then in one group of rural regions and then in another group of rural regions (Ministry of Education, Tanzania 1960-1969). The entrance exam into Standard five was also "virtually abolished in towns and cities by 1966", though it took several more years for it to be abolished in rural areas (Morrison 1976).

The key result of these reforms was that the majority of children now obtained seven, rather than four years of education. Another consequence of the reforms was that all those who started secondary school post-reform obtained one year less of education than they would have had pre-reform.

In my estimation below I use this reform as a source of exogenous variation in the number of years of schooling to estimate an ATE and LATE.

Subsequent to these reforms being tabled, and whilst they were being implemented, the Tanzanian president, Julius Nyerere, published further plans for reform, which he entitled "Education for Self-Reliance" (Morrison 1976). Nyerere criticised the education system for hampering Tanzania's development as a socialist country and linked the extension of primary school to seven years to dramatic changes in schooling curricula, to be designed to prepare students for the realities of agriculture that Nyerere argued they were most likely destined to find employment in. Morrison (1976) notes, however, that the changes implemented were actually much smaller than Nyerere envisaged, mainly because of a lack of funding for their implementation and teachers who were unable or unwilling to implement the changes, with Sabot (1979) noting that the Tanzanian educational system "has an inherent inertia."

Dramatic changes in the education curricula around the time of the reform would make my analysis problematic, since I could then be capturing the effects of these changes, rather than the increased years of schooling that many Tanzanians acquired post-reform. The evidence reviewed in the previous paragraph suggests that these changes were less important than Nyerere had hoped, however, and hence that my estimation strategy, described above in section 4.4, is still robust.

To explore the impact of these reforms on educational attainment I obtained data on enrollment during the time of the reform period from the Ministry of Education Annual Reports (Ministry of Education, Tanzania 1960-1969). From these data I can obtain a measure of educational attainment by taking the enrollment in a particular grade, in a particular year and subtracting the number of enrollments in the next school grade in the following year. This is a rough measure of educational attainment since the next years' enrollment in a particular grade includes those repeating that grade, as well as those who may have stopped school for a year and are now returning. Nevertheless, Figure 4.1 shows the expected outcomes obtain; the reform drastically reduced the number of individuals obtaining four years of primary school, increased the number finishing seven (rather than eight) years of primary school (primary plus middle in the case of those who obtained eight) and reduced the number obtaining eight years of primary plus middle school to zero.

4.6 Data source and Description

In this chapter I make use of the Tanzanian Integrated Labour Force Surveys (ILFS) from 2001 and 2006 to estimate the returns to education in Tanzania. These are cross-sectional, nationally representative surveys and were discussed in more detail in the previous chapter of this thesis. I do not make use of the 2006 population weights because the 2001 weights are not publicly available.

One important limitation of the ILFSs is that income data were not collected for those engaged in rural agriculture. In addition, most of those engaged in agriculture in urban areas engaged in home production rather than for market and imputed incomes were also not recorded, nor was there sufficient information asked to calculate these. I am thus forced to ignore agricultural earnings in this chapter and estimate the returns to education for the sub-population of urban workers in non-agricultural employment, which is much smaller than the total population. As explained in the previous chapter, I continue to use gross income minus expenses as my measure of self-employment earnings, adjusted for inflation using the HBS deflators.

4.6.1 A brief description of the data

I now provide a brief description of the data from the 2001 and 2006 ILFS surveys, limiting the sample to those aged between 15 and 65 inclusive at the time each survey was conducted. Table 4.1 shows labour force participation and employment levels in Tanzania over the two surveys, both for the full sample between the ages of 16 and 65 (columns 1 and 3), and the sample of earners used in the regression analysis (columns 2 and 4). The table shows that, conditional on participation, unemployment is extremely low, a phenomenon common to many countries in sub-Saharan Africa, although not to South Africa (Kingdon et al. 2006). The earner sample is predominantly male, reflecting Tanzanian women's roles in home, rather than market, production. The average level of education in both samples is slightly higher than completed primary. As a result of agricultural earnings not being captured in the ILFSs, urban dwellers are over-represented in the earner sample.

Table 4.1 also gives occupational breakdowns, showing that Tanzania is predominantly an agricultural economy, with subsistence agriculture being by far the most common form of employment. The share of government employment in total employment seems stable, but self-employment with no employees seems to have an substantially increased share over the period.

Table 4.1 shows that approximately 12 percent of the 2001 sample faced the pre-reform educational system and that this had decreased to 10 percent percent in 2006, both as a result of an ageing population and the cut-off I impose in limiting the analysis to those younger than 65 years of age. Figure 4.3 shows that education levels are higher in urban areas (for those aged between 15 and 65 in both surveys)[3]. Figure 4.4 further shows that education levels are higher amongst the population of earners, again because of the urban bias of the earner sample.

As I noted above, I am forced to limit my multivariate analysis to those either in wage employment or non-agricultural self employment, although my sample is still more representative than previous studies using Tanzanian data, which used only urban manufacturing employees (Söderbom et al. (2006), Knight and Sabot (1990)). Thus I am estimating the return to education for a sample of earners who have higher incomes and education levels than the average Tanzanian. Selection into earnings when using IV has generally been ignored in the literature for the US and other developed countries. It has been incorporated into models that estimate the returns to education in some other contexts (Chen and Hamori 2009), but this area of research does not yet allow for parameter heterogeneity as well as selection. As a result I do not undertake selection correction in this chapter and thus I am estimating the returns to education for a sub-sample of the working population, albeit one that is more representative than has been used in other work on Tanzania to date.

Table 4.2 shows how median earnings vary by education level and suggests a positive relationship between earnings and education[4]. Table 4.2 also shows median earnings levels by occupation, with government employees earning the highest on average, followed by the self-employed with employees, wage employees in private firms and then the self-employed with no employees. The data show stagnant real earnings over the period. This is suprising given that the World Bank records that Tanzania had real GDP

[3]The average difference is about two years in both surveys.

[4]The collapse in earnings for those with 15 or more years of education from 2001 to 2006 is likely the result of the small number of earners in the category.

growth of higher than 6 percent over the period (World Bank 2009) but is not inconsistent with the Household Budget Survey, which showed that poverty fell only very marginally between 2000 and 2007 and that this fall was not statistically significant (National Bureau of Statistics, Tanzania 2007).

4.6.2 Measuring Educational Attainment and the Effect of the Education Reform

In calculating individual years of education in the previous section I used specific questions on educational attainment from the 2001 and 2006 ILFSs. There were 17 codes for the level of education an individual had completed: pre-primary, standards 1-8, forms 1-6, tertiary non-university and tertiary university. Standard eight was the last year of middle school in the pre-reform education system. As a result of the education reform middle school was removed and seven years of primary school introduced. Accurately translating the survey information into actual years of completed schooling requires overcoming challenges posed by the survey data.

I noted above that the enrollment data from the Tanzanian Ministry of Education, shown in Figure 4.1, indicates that before the reform there was almost no one stopping school after completing standard seven, which was one year before the end of middle school. It also shows that post-reform there was no one stopping school after completing standard eight, which had been abolished. Pre-reform, the end of primary school (four years) was where the majority of children stopped school and the end of extended primary school (seven years) was where the majority of children stopped school post-reform.

The ILFS data, however, do not show the same situation. Figure 4.2

shows educational attainment data from the 2006 ILFS[5]. The "before" bars show that, for individuals whose age indicates they could have completed middle school before the reform (assuming that all individuals started school at age six, did not repeat any standards and correctly reported their age), those who report completing standard seven are nearly as numerous as those completing standard four. Figure 4.2 also shows that, for individuals who faced the pre-reform system, about half the number of those who report completing standard four completed standard eight. Thus the ratio of seven years to four years to eight years of education in the ILFS data is roughly 2:2:1. This is not consistent with the pre-reform educational attainment reported by the Ministry of Education in Figure 4.1, which shows that almost no one completed seven years of primary school pre-reform and that the ratio is roughly 0:5:1, depending on which year one uses (I only have data going back to 1961).

If individuals were much slower in reaching four, seven or eight years of education, then this could explain some of these individuals obtaining facing the post-reform system and obtaining seven years of education, despite being of an age that suggests that they should have faced the pre-reform system. I suspect, however, that most of the explanation for this difference is measurement error in the ILFS surveys generated by confusion of the survey enumerators about the length of primary school. I believe that some enumerators may have asked whether respondents had completed primary school and, if the answer was yes, they may have simply written down that the respondent had completed standard seven, without establishing whether completing primary meant four years under the old system or seven years under the new system[6]. If I re-assign four years of education to those who

[5]The 2001 survey data shows a similar pattern.

[6]I have undertaken a similar labour market survey for the Centre for the Study of

are reported to have seven years, but whose age makes them likely to have undergone the pre-reform system, based on the assumptions I am required to make and which I discussed above, I find that the pre-reform ILFS data match the Ministry of Education figures more closely, with a ratio of roughly 0:4:1, and so I continue to use this solution throughout the chapter.

4.6.3 The Tanzanian education reform and the Monotonicity assumption

In section 4.2 I noted that researchers were aware measurement error in education could bias the results of OLS, even without the endogeneity problems that result from the presence of unobserved individual heterogeneity. IV was thought to correct measurement error in education (Cameron and Trivedi (2005), Card (1999)) and findings that IV estimates were higher than OLS were often attributed to IV correcting for measurement error (Card 2001). Card's (2001) alternative explanation is that heterogeneity in the returns to education meant IV was estimating a LATE rather than an ATE and I noted that given the stringent conditions required to estimate an ATE with IV, I wish to interpret my 2SLS estimates below as LATEs.

It should be clear from the discussion above, however, that monotonicity, the key assumption required to be able to consistently estimate a LATE, does not hold in the case of the Tanzanian education reform. Most of the individuals for whom the reform created an exogenous shift in education, as a result of facing the post-reform system, obtained seven instead of four years of education. All those who obtained any level of high school education obtained one year less than if they had done the same in the pre-reform

African Economies in Tanzania, that asked a similar question about educational attainment. In training enumerators for the survey I found that the majority were unfamiliar with the pre-reform education system, which lends some support to my hypothesis system, however, because primary schooling was reduced by one year. This means that although most individuals who changed their educational status as a result of the reform obtained more education, some obtained one year less.

In section 4.4.3 I showed, for the binary treatment case, that the violation of monotonicity can lead to estimated treatment effects that are zero or even negative, despite treatment effects being positive for all individuals. This is also true for multi-valued treatments like education (Angrist and Pischke 2009). In practice, however, this violation of monotonicity may not be quite as serious in the case of the Tanzanian reform. Only those who obtained at least seven years of education had a year less as a result of the reform, and of these Figure 4.2 suggests a large majority would have 12 years (after completing O levels). One year less of primary education for these individuals would arguably have only a very small impact compared to the effect of education increasing from four to seven years, as would have occurred for the majority of those affected by the reform. I suggest that although LATE may still be biased downwards, as a result of the negative weighting for those who obtained less education and the violation of monotonicity, this bias is likely to be small.

4.7 Estimating the returns to education

Having explored the data, and having described some concerns about them, I begin the multivariate analysis by ignoring potential biases due to unobserved heterogeneity in ability and returns, and follow Mincer (1973) in using a semi-logarithmic specification of the earnings equation, which I estimate using OLS. In addition to years of completed education, I include a measure of age, age at reform year, and its square as a proxy for the relationship between earnings and labour market experience, and also interact this with a gender dummy, which I find to be important. I also include an urban dummy.

Table 4.3 shows results for 2001 and Table 4.4 shows 2006 results. The second column of each table follows much of the developed country literature and forces returns to education to be linear, showing a constant marginal return of 14 percent a year in 2001 and 12 percent a year in 2006, nearly double those found in the US. Including a squared term in column three allows marginal returns to vary and the OLS estimates show a convex pattern of returns to education in both cross sections, as in many of the studies on returns to education in Africa discussed in section 4.2 above, and contradicting the work of Psacharopoulos (1994). These imply marginal returns of 13.5 percent (13 percent) at standard seven and 16.5 percent (19 percent) for those who completed eleven years of education, to obtain O levels, in 2001 (2006).

These returns are much higher than those reported in Knight and Sabot (1990) and Söderbom et al. (2006) in Tanzania. Knight and Sabot (1990) use a survey of manufacturing wage employees in Dar es Salaam, Tanzania's largest urban centre, from 1980 and show that the returns to education are around eight percent a year without controls for occupation but that they are halved when these controls are included. They also find strongly non-linear effects, and those with secondary or post-secondary education have returns of around three percent and seven percent per year respectively when including occupation controls.

In their 1993 sample of urban wage employees Söderbom et al. (2006) find marginal returns to an extra year of education of five percent for years one to eleven with occupational controls, when using OLS. For those with twelve or more years they find returns of thirteen percent. Their 2001 sample provides dramatically different results, with marginal returns below 10 percent for years one to eleven and then jumping to 30 percent for those with twelve or more years of education. The nature of their sample is the most likely explanation for the difference between my results and theirs, since theirs is more selective, focusing only on urban manufacturing employees. Limiting my sample to urban wage employees does not change my results dramatically, however[7]. The evidence thus suggests that returns to education for urban wage employees rose between 1980 and the early 1990s and have risen again dramatically between the early 1990s and the 2000s. This may reflect the decreased enrollment in Tanzanian primary education that occurred from 1980 (Bommier and Lambert 2000), as well as the low secondary school enrollment that has occurred in much of the post-independence period (Knight and Sabot 1990).

I prefer a specification in which I do not include sector of employment or firm size, because these are themselves likely to be determined by the level of education attained and I wish to explore the total returns to education, including those that come from being able to access higher paying types of employment (Pereira and Martins 2004). I do include these in Columns 3 and 7 in Tables 4.3 for 2001 and 4.4 for 2006, however, for comparison with the earlier studies. They show that the linear returns are reduced to 10 percent in both 2001 and 2006, which is still higher than the numbers reported in Söderbom et al. (2006). These results also indicate that the convexity of the returns to education is reduced by controls for occupation in both surveys. They imply marginal returns of 10 percent (10 percent) at standard seven and 12 percent (15 percent) at eleven years of education,

[7] The last two columns of Table 4.6 show OLS estimates of the return to education for urban wage employees in 2001 and 2006

the end of O levels, in 2001 (2006) and are still higher than those reported in Söderbom et al. (2006) and Knight and Sabot (1990).

The key insight from the returns to education literature is that these results may be affected by ability bias and heterogeneity in returns, with findings of high returns possibly caused by ability bias or returns bias, if those who take more education benefit more from it. I thus explore the effect of allowing for individual heterogeneity, using the exogenous variation in years of education generated by the Tanzanian education reform, in the following sections.

4.7.1 Control function estimates of the return to education

The control function approach can generate consistent estimates of the ATE of education if the relationship between the unobservables in the earnings and education equations is correctly specified. This is possible even when the ability-schooling mapping is affected by the reform, unlike IV, but I noted that this requires a mean independence assumption that implies having correctly specified the educational attainment equation. Identification of the ATE occurs through fairly strong parametric assumptions about the relationship between individual heterogeneity in earnings and the education residual.

Column 1 and 3 of Table 4.3 and 4.4, for 2001 and 2006 respectively, show the estimation of the educational attainment equation, the first stage. In testing various specifications I find that allowing for age effects to differ by gender is important, and so I include these. The non-linear age terms control for increases in primary enrollment in the post-reform period (Bommier and Lambert 2000). Older age for both men and women is correlated with lower educational attainment. Men have a year's extra education compared to women on average. The reform dummy shows that the reform increased educational attainment by about half a year on average, after controlling for age effects which allow for secular increases in education levels over time.

In the simplest specification of linear returns the control function estimates suggest that OLS underestimates the ATE of education, with Table 4.4 suggesting it is around 15 percent in 2006, compared to 12 percent using OLS. In 2001 I find a larger change, column 8 of Table 4.3 shows the estimated return increases from 14 percent when using OLS to 21 percent.

Using the control function and allowing for non-linear returns by including a quadratic education term, shown in column 5 of Table 4.3 for 2001 and column 4 of Table 4.4, produces a dramatic effect. Returns are estimated to be concave rather than convex in both surveys, in contrast with my OLS results and most of the recent studies from developing countries. In 2001 the average treatment effect of education is estimated to be 22 percent in standard seven and 11 percent after the eleven years taken to achieve O levels. In 2006 the switch to concavity is more pronounced, with the ATE estimated to be 16 percent at standard seven, higher than the OLS estimate, but only 5 percent for completed O levels.

The differences in estimated returns between OLS and the control function when allowing for non-linear effects are illustrated in Figure 4.5 (for 2001) and Figure 4.6 (for 2006). In both cases the expected log earnings estimated with the control function are much lower at low levels of education but very similar around the median of seven years. The control function and OLS results differ for higher levels of education, however. The 2001 control function results are not very different to the OLS results but the 2006 results show lower expected log earnings in the control function compared to OLS.

Both the control function and OLS results suggest that returns to education are high in Tanzania. The control function estimates of the average treatment effect of education do indicate that there is some evidence of ability bias and returns bias, with marginal returns estimated to be much lower for secondary education than OLS in 2006. The high degree of significance of the coefficient on the residuals from the first stage regression implies that I can reject the exogeneity of education in the earnings function. I also find that the interaction of education and the residual from the first stage has a highly significant coefficient in all of the control function regressions, which indicates that there is heterogeneity in the returns to education, *contra* Söderbom et al. (2006), but consistent with the model I outlined in section 4.3.

4.7.2 Estimating the LATE of education

I noted in section 4.4 that IV can recover the ATE under some strong assumptions, including homoscedasticity of the schooling residual conditional on Z_i. Card (2001) points out that these are unlikely to hold if the entire mapping between schooling and ability varies for different values of the exogenous instrument, college proximity in his example, and provides evidence that this change does indeed occur in the US. This would then imply bias in the two-stage least squares IV estimates of the ATE of education.

If homoscedasticity does not hold, then one can use IV to estimate a LATE under much less restrictive assumptions. I noted above, however, that the Tanzanian reform does not satisfy the monotonicity assumption. Although many individuals who would have obtained only four years, obtained seven as a result of the reform, those who went on to high school obtained one year less. As I showed in section 4.4.3, this can result in estimated effects that are zero or negative, even if treatment effects are positive for all individuals.

Columns 2 and 4 of Table 4.5 report the two-stage least squares results, which I interpret as the LATE of education[8]. The results indicate that the 2001 and 2006 data show much smaller effects of education than those estimated with OLS or the control function and the coefficients are also not significantly different from zero. The violation of the monotonicity assumption is a plausible reason for the low two-stage least squares estimate of the LATE. In addition to these concerns there also seems to be an issue with the instrument strength. The last row of Table 4.5 reports the F-statistics for the Cragg and Donald (1993) test of weak instruments. A value of 10 or higher is generally regarded as evidence that the exogenous regressors are sufficiently relevant for finite sample bias not be a concern (Cameron and Trivedi 2005). Unfortunately neither the 2001 or 2006 instruments pass this test, suggesting another reason why the LATE estimates may be biased.

4.7.3 Robustness Checks

In this section I report some robustness checks on my main results. Combining the pre- and post-reform sub-samples to explore the effect of the reform may not be valid if the relationship between education and earnings differs dramatically across the two groups. Columns 5 and 6 of Tables 4.3 (for 2001) and 4.4 (for 2006) show OLS estimates of returns to education for those who faced the pre- and post-reform educational system respectively. The 2006 results show similar returns to education for both sub-samples, with some evidence of slightly stronger convexity in the post-reform sub-sample. The 2001 results suggest concavity in the pre-reform sub-sample, although the coefficient on the square of education is not significantly different from zero.

[8]Columns 1 and 3 of Table 4.5, for 2001 and 2006 respectively, show the the first stage regressions

This suggests there is not strong evidence that the different groups should not be pooled to estimate the returns to education in Tanzania.

Table 4.6 shows OLS estimates and the first and second stage results using the control function approach for men and for women separately, pooling the two surveys to increase the sample size. I again find that returns are high for both men and women and that the control function procedure raises returns at low education levels and lowers them at high levels of education compared to OLS for both genders.

At the time the reform was implemented the government of Tanzania was becoming an increasingly large employer through the expansion of state owned enterprises and through this employment it also aimed to compress the earnings distribution (James 1996). Individuals who obtained extra education as a result of the reform and also obtained employment in government jobs may thus be driving the results. One simple way to check this is to exclude all former public sector employees. Unfortunately I cannot do this I do not know the employment histories of the individuals in the sample. The best I can do is exclude all current government employees.

These results for 2001 and 2006 are shown in column 10 of Tables 4.3 and 4.4 respectively. They do not tell a consistent story. In 2006 the returns are still estimated to be high and concave but in 2001 returns are actually estimated to be negative, although the standard errors are very large. This may simply reflect the large number of government employees with high levels of education and the resultant reduction in the number of observations when government employees are excluded: in both the 2001 and 2006 samples government employees make up over half of those with 11 or more years of education.

4.8 Conclusion

In this chapter I have described a reform in the Tanzanian education system that generated exogenous shifts in the level of education for many Tanzanians and used this to estimate the returns to education and purge these of the effects of individual heterogeneity. The data I use enable me to estimate the returns for a more representative sample of the Tanzanian population than in previous work, although I am still forced to exclude those working in agriculture.

The estimates of the returns to education across several African economies indicate that returns are larger than in developed countries and that they are convex. If these estimates represent the true effects of education on productivity then this may have important policy implications for the way governments fund primary, secondary and tertiary education. Findings of convexity in returns may be due to contamination by unobserved ability, however. There are only a few studies which test this possibility in African labour markets and these show that returns are convex even after correcting for this source of endogeneity. This suggests ability or returns bias are not responsible for the convexity of returns estimated with OLS.

I estimated the returns to education using the ILFS data from 2001 and 2006. In the simplest linear specification estimated with OLS I find marginal returns to be around 12 percent in 2006 and 14 percent in 2001, higher than those reported for Tanzania in Söderbom et al. (2006) and Knight and Sabot (1990). Allowing for non-linear returns in OLS estimation suggests that returns are actually strongly convex. My estimates are still higher than those estimated previously when I confine my analysis to a sub-sample similar to that used in other studies.

The most general model I use allows for individual heterogeneity in ability and in the marginal returns for education, as well as non-linear returns to education. Heterogeneity in returns makes estimating an average treatment effect much harder. Three solutions exist. The IV and control function approaches both require fairly strong assumptions to estimate an ATE, while the LATE, though requiring fewer assumptions, is a weighted average of returns for a sub-population that is generally not identified. In allowing for potential unobserved heterogeneity in ability and returns I have attempted to estimate the average treatment effect using the control function approach and the LATE using two-stage least squares and compared these to the OLS estimates.

Estimating the average treatment effect in this more general model using the control function approach I generally still find high returns. I also find that returns are actually concave. These results suggest ability bias and returns bias may be important, with the true ATE lower at higher levels of education than indicated by the OLS estimates. In undertaking robustness checks I find similar results when I separate men and women and when I exclude government employees using the 2006 data, although not in the 2001 data.

The Tanzanian education reform I discussed does seem to offer a plausibly exogenous shift in education levels for those who faced the new educational system. It does not, however, satisfy the one crucial assumption required to estimate a LATE, monotonicity, since individuals at the lower end of the education distribution obtained more education, whilst those at the top end obtained less. The LATEs I estimate are lower than the linear OLS estimates, and in 2006 they are estimated to be negative. The simplest explanation for this is that the reform pushed different individuals in different directions, violating the monotonicity assumption and biasing the coefficients towards zero as a result. I also noted that weak instrument bias is affecting the LATE estimates.

The picture emerging from studies of returns to education in sub-Saharan Africa is one of larger returns than in developed countries. I find similarly large returns in OLS, roughly double those reported for the US labour market. The shape of the earnings function has been found to be convex in many studies using Africa micro data but this literature has not progressed much beyond estimating OLS returns in dealing with the endogeneity introduced by individual heterogeneity. I also find high and convex returns when using OLS. When I use the control function to estimate the average treatment effect I find generally high returns, although returns are estimated to be concave. I still find that returns are high for completed primary school, the median level of education but that they are lower than in OLS at secondary school level. This suggests that ability or returns bias, resulting from individual heterogeneity, may explain the convex OLS estimates of the returns to education in Tanzania.

Table 4.1: Descriptive Statistics

	2001		2006	
	All 16-65	**Earners**	**All 16-65**	**Earners**
Demographics				
Age	32.58	38.59	33.35	37.58
	(13.07)	(9.68)	(12.88)	(9.69)
Male	0.48	0.63	0.47	0.64
	(0.50)	(0.48)	(0.50)	(0.48)
Education (yrs.)	5.50	7.57	5.51	7.12
	(3.46)	(3.33)	(3.38)	(3.14)
Urban	0.33	0.78	0.54	0.82
	(0.47)	(0.41)	(0.50)	(0.38)
Pre-reform schooling	0.12	0.12	0.10	0.09
	(0.32)	(0.33)	(0.30)	(0.29)
Labour Market Outcomes				
Labour Force Participant	0.89	1.00	0.93	1.00
	(0.31)	(0.00)	(0.25)	(0.00)
Employed	0.83	1.00	0.88	1.00
	(0.37)	(0.00)	(0.32)	(0.00)
Self Emp with employees	0.02	0.09	0.02	0.08
	(0.13)	(0.29)	(0.13)	(0.26)
Self Emp no employees	0.09	0.44	0.12	0.48
	(0.29)	(0.50)	(0.33)	(0.50)
Private Emp	0.06	0.25	0.07	0.27
	(0.25)	(0.43)	(0.26)	(0.44)
Public	0.04	0.22	0.03	0.15
	(0.19)	(0.41)	(0.17)	(0.36)
Agriculture	0.61	0.00	0.54	0.00
	(0.49)	(0.00)	(0.50)	(0.00)
Frequency	30974	4962	36402	6915

Source: ILFS 2001 and 2006. Note: Standard deviations in parentheses.

Table 4.2: Median real earnings in Tanzanian Shillings by Education level and Occupation

	2001		2006	
	Median	N	Median	N
Occupation				
S/E with employees	65000	499	53571	583
	(473817)		(248207)	
S/E no employees	28821	2804	26786	4049
	(90958)		(172406)	
Private employee	30000	2043	31250	2624
	(446874)		(102067)	
Public Employee	76675	1120	75000	1108
	(189924)		(252890)	
Education Level				
None	15000	433	18750	727
	(47705)		(89059)	
1-4 years	28000	555	27054	649
	(113570)		(129346)	
5-8 years	30000	3902	30603	5569
	(86432)		(199493)	
9-14 years	70000	1445	62500	1615
	(596894)		(255311)	
15 or more years	245000	138	120313	112
	(436352)		(234615)	

Note: Standard deviations in parentheses. The average exchange rate was Tsh 916 per US dollar in 2001 and 1261 in 2006 (International Monetary Fund 2010)

Table 4.3: OLS and Control Function Estimates of Returns to Education, 2001

	first	ols0	ols0c	ols1	ols1a	ols1b	ols1c	cf-0	cf-1	cf-2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age at reform	-.039	0.033	0.02	0.031	0.007	0.032	0.019	0.032	0.031	-.106
	(0.014)***	(0.004)***	(0.004)***	(0.004)***	(0.294)	(0.004)***	(0.004)***	(0.004)***	(0.004)***	(0.065)
Age at reform ^{2}	-.003	-.0008	-.0005	-.0009	0.0001	-.001	-.0005	-.0004	-.00009	-.001
	(0.0009)***	(0.0002)***	(0.0002)**	(0.0002)***	(0.006)	(0.0005)**	(0.0002)**	(0.0003)	(0.0003)	(0.0008)*
Male*Age at reform	0.112	0.001	0.002	0.001	0.079	-.004	0.002	-.006	-.005	0.141
	(0.018)***	(0.005)	(0.005)	(0.005)	(0.324)	(0.006)	(0.005)	(0.007)	(0.007)	(0.072)*
Male*Age at reform^2	-.002	0.00007	-.00004	0.00007	-.002	0.0006	-.00004	0.0001	-.0001	-.003
	(0.001)	(0.0003)	(0.0003)	(0.0003)	(0.007)	(0.0006)	(0.0003)	(0.0003)	(0.0003)	(0.001)**
Log (hours)		0.106	0.177	0.121	0.451	0.078	0.189	0.127	0.119	0.255
		(0.042)**	(0.04)***	(0.042)***	(0.125)***	(0.044)*	(0.04)***	(0.046)***	(0.048)**	(0.052)***
Education		0.141	0.104	0.089	0.126	0.079	0.063	0.207	0.437	-.937
		(0.004)***	(0.004)***	(0.012)***	(0.028)***	(0.014)***	(0.012)***	(0.038)***	(0.076)***	(0.611)
Educ^2/(100)				0.343	-.061	0.453	0.273		-1.503	-.418
				(0.075)***	(0.173)	(0.084)***	(0.072)***		(0.433)***	(0.518)
Male d	0.628	0.4	0.377	0.402	0.046	0.382	0.379	0.367	0.381	1.404
	(0.111)***	(0.033)***	(0.032)***	(0.033)***	(3.873)	(0.039)***	(0.032)***	(0.042)***	(0.044)***	(0.499)***
S/E with employees			0.744				0.742			
			(0.047)***				(0.046)***			
Private Wage			0.109				0.103			
			(0.032)***				(0.032)***			
Public employment			0.714				0.705			
			(0.038)***				(0.038)***			
Pre-reform d	-.501									
	(0.266)*									
Urban d	1.195	0.513	0.568	0.514	0.779	0.474	0.568	0.423	0.405	2.379
	(0.116)***	(0.033)***	(0.031)***	(0.033)***	(0.099)***	(0.034)***	(0.031)***	(0.064)***	(0.066)***	(0.951)**
Residuals								-.097	-.216	1.032
								(0.039)**	(0.053)***	(0.604)*
Educ*Resid								0.005	0.021	0.012
								(0.0009)***	(0.005)***	(0.006)**
pre-reform*residual								-.008	-.028	0.029
								(0.02)	(0.021)	(0.026)
pre-reform*residual*S								-.003	-.003	-.008
								(0.002)*	(0.002)*	(0.003)***
Obs.	4962	4942	4938	4942	616	4326	4938	4942	4942	3862
R^2	0.1	0.314	0.378	0.317	0.362	0.314	0.38	0.32	0.322	0.271
F statistic	50.177	282.254	272.255	254.206	38.217	219.348	251.455			

Notes: dependent variable is years of education in column 1 and log(income) in all other columns. *, **, *** denote significance at 10%, 5% and 1% levels. d indicates a dummy variable. Second stage control function regressions have bootstrapped standard errors with 1000 replications to account for generated regressors.

Table 4.4: Estimates of returns to education, 2006

	FIRST	ols-0	ols-0c	ols-1	ols-1a	ols-1b	ols-1c	cf-0	cf-1	cf-2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age at reform	-.054	0.019	0.014	0.015	0.131	0.019	0.012	0.018	0.02	0.011
	(0.007)***	(0.002)***	(0.002)***	(0.002)***	(0.279)	(0.003)***	(0.002)***	(0.003)***	(0.003)***	(0.007)*
Age at reform ^{2}	-.003	-.0003	-.0001	-.0005	-.005	0.0003	-.0003	-.0002	8.95e-06	0.00004
	(0.0007)***	(0.0002)	(0.0002)	(0.0002)**	(0.007)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Male * Age at reform	0.074	-.003	-.003	-.002	-.161	-.005	-.003	-.005	-.007	-.004
	(0.009)***	(0.003)	(0.003)	(0.003)	(0.327)	(0.003)	(0.003)	(0.004)	(0.004)*	(0.006)
Male * Age at reform^2	0.0007	2.41e-06	-.00006	0.00008	0.005	-.0005	-6.79e-06	-.00008	-.0003	-.0006
	(0.0008)	(0.0002)	(0.0002)	(0.0002)	(0.008)	(0.0004)	(0.0002)	(0.0002)	(0.0003)	(0.0003)**
Log (hours)		0.115	0.178	0.151	-.377	0.21	0.201	0.155	0.151	0.247
		(0.037)***	(0.037)***	(0.037)***	(0.125)***	(0.038)***	(0.036)***	(0.039)***	(0.039)***	(0.043)***
Education		0.125	0.097	0.018	0.054	0.013	0.011	0.153	0.345	0.232
		(0.004)***	(0.004)***	(0.011)*	(0.03)*	(0.011)	(0.01)	(0.034)***	(0.075)***	(0.093)**
Educ^2/(100)				0.786	0.657	0.777	0.637		-1.323	-.839
				(0.072)***	(0.192)***	(0.079)***	(0.072)***		(0.459)***	(0.489)*
Male d	0.88	0.42	0.401	0.43	1.541	0.456	0.412	0.415	0.421	0.513
	(0.107)***	(0.035)***	(0.034)***	(0.034)***	(3.213)	(0.038)***	(0.034)***	(0.048)***	(0.046)***	(0.079)***
S/E with employees			0.488				0.486			
			(0.044)***				(0.044)***			
Private Wage			-.159				-.169			
			(0.032)***				(0.032)***			
Public employment			0.584				0.53			
			(0.037)***				(0.037)***			
Firm size > 10			0.365				0.345			
			(0.043)***				(0.043)***			
Pre-reform d	-.520									
	(0.244)**									
Urban d	1.052	0.27	0.239	0.283	0.334	0.277	0.248	0.247	0.23	0.274
	(0.097)***	(0.031)***	(0.031)***	(0.03)***	(0.103)***	(0.032)***	(0.031)***	(0.053)***	(0.054)***	(0.093)***
Residuals								-.088	-.189	-.130
								(0.034)***	(0.049)***	(0.074)*
Educ*Resid								0.008	0.022	0.017
								(0.0009)***	(0.005)***	(0.005)***
pre-reform*residual								0.015	-.002	-.049
								(0.019)	(0.021)	(0.026)*
pre-reform*residual*S								0.0005	0.0007	0.004
								(0.002)	(0.002)	(0.002)*
Obs.	6915	6915	6694	6915	631	6284	6694	6915	6915	5885
R^2	0.08	0.216	0.259	0.229	0.41	0.206	0.267	0.232	0.234	0.188
F statistic	54.67	237.96	194.433	228.264	48.004	180.421	187.558			

Notes: dependent variable is years of education in column 1 and log(income) in all other columns. *, **, *** denote significance at 10%, 5% and 1% levels. d indicates a dummy variable. Second stage control function regressions have bootstrapped standard errors with 1000 replications to account for generated regressors.

Table 4.5: LATE ESTIMATES OF RETURNS TO EDUCATION

	first2001	late2001	first2006	late2006
	(1)	(2)	(3)	(4)
Reform age	-.038	0.028	-.054	0.009
	(0.014)***	(0.007)***	(0.007)***	(0.011)
Reform age^2	-.003	-.001	-.003	-.0009
	(0.0009)***	(0.0007)	(0.0007)***	(0.0007)
Male*Reform age	0.112	0.007	0.074	0.009
	(0.018)***	(0.018)	(0.009)***	(0.013)
Male*Reform age^2	-.002	-6.99e-06	0.0007	0.0001
	(0.001)	(0.0004)	(0.0008)	(0.0003)
Education		0.079		-.032
		(0.159)		(0.17)
Male d	0.63	0.459	0.88	0.574
	(0.112)***	(0.105)***	(0.107)***	(0.157)***
Urban d	1.197	0.54	1.052	0.458
	(0.116)***	(0.193)***	(0.097)***	(0.209)**
Pre-reform d	-.500		-.520	
	(0.267)*		(0.244)**	
Obs.	4942	4942	6915	6915
R^2	0.101	0.286	0.08	0.025
F statistic	50.15	99.768	54.67	96.416
Cragg-Donald Wald F statistic		3.507		4.355

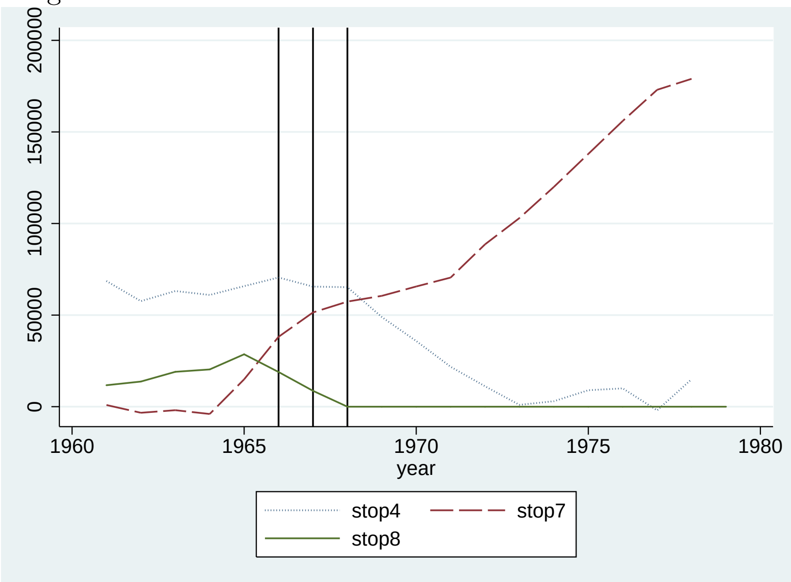
Notes: dependent variable is years of education in columns 1 and 5 and log(income) in all other columns. *, **, *** denote significance at 10%, 5% and 1% levels. d indicates a dummy variable.

Table 4.6: OLS and Control function estimates of returns to education for men and women separately, using pooled data, and for urban workers

	Ffirst	Fols	Fcf	Mfirst	Mols	Mcf	UW2001	UW2006
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform age	-.045	0.019	0.024	0.029	0.018	0.017	0.026	0.026
	(0.007)***	(0.002)***	(0.003)***	(0.005)***	(0.001)***	(0.002)***	(0.008)***	(0.005)***
Reform age^2	-.002	-.0004	0.00002	-.003	-.0004	-.0003	-.0008	-.0003
	(0.0006)***	(0.0001)***	(0.0002)	(0.0004)***	(0.00009)***	(0.0002)*	(0.0006)	(0.0004)
Male*Reform age							0.011	-.012
							(0.009)	(0.005)**
Male*Reform age^2							-.0004	-.00006
							(0.0007)	(0.0005)
Pre-reform d	-1.169			-.477				
	(0.319)***			(0.197)**				
Education		0.028	0.267		0.042	0.178	0.051	0.013
		(0.012)**	(0.064)***		(0.011)***	(0.08)**	(0.026)**	(0.02)
Educ^2/(100)		0.937	-.252		0.524	-.252	0.517	0.839
		(0.088)***	(0.401)		(0.067)***	(0.458)	(0.147)***	(0.128)***
2001 d	0.775	-.501	-.563	0.485	-.479	-.486		
	(0.104)***	(0.031)***	(0.041)***	(0.078)***	(0.024)***	(0.028)***		
Residuals			-.165			-.082		
			(0.044)***			(0.052)		
educehat			0.012			0.008		
			(0.004)***			(0.005)*		
Obs.	4359	4359	4359	7498	7498	7498	991	1588
R^2	0.102	0.253	0.256	0.068	0.229	0.229	0.327	0.291
F statistic	54.922	245.35		60.91	370.566		68.348	92.457

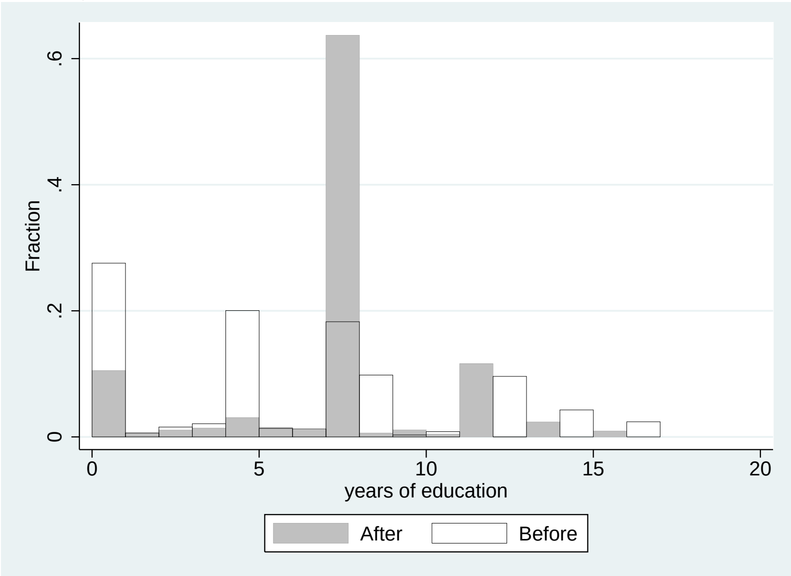
Notes: dependent variable is education in columns 1 and 4 and log(income) in all other columns. Region dummies included but not reported. *, **, *** denote significance at 10%, 5% and 1% levels. levels. d indicates a dummy variable. In column headings M indicates Male, F indicates female and UW indicates urban wage employees. Second stage control function regressions have bootstrapped standard errors with 1000 replications to account for generated regressors.

Figure 4.1: Number of pupils stopping at the end of each year



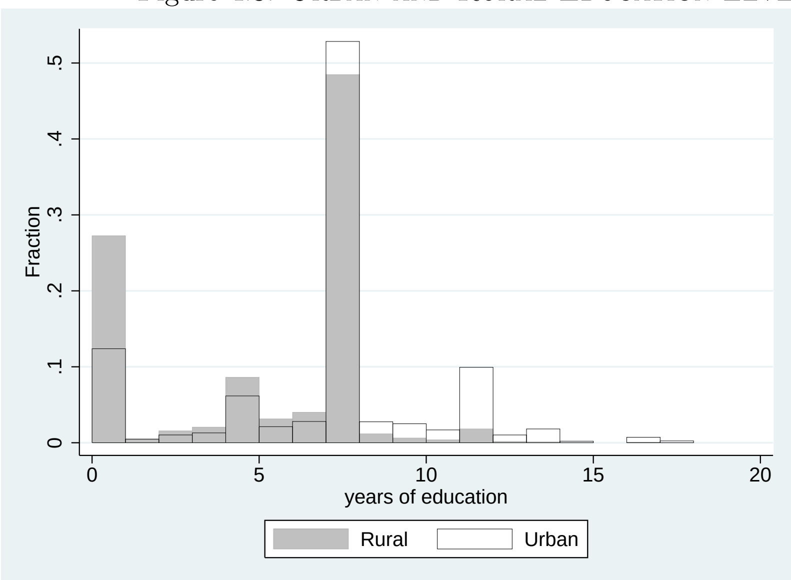
Source: Tanzanian Ministry of Education Annual Reports 1968-1979. The vertical lines represent the three years the reform was implemented in.

Figure 4.2: Education Levels as recorded in the 2006 ILFS



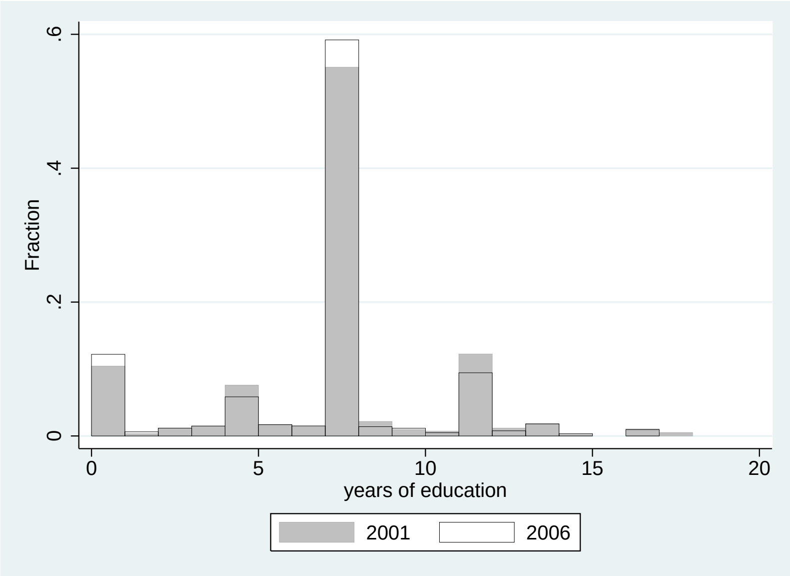
Source: Tanzanian ILFS 2006.

Figure 4.3: Urban and Rural Education Levels



Source: Tanzanian ILFS 2001 and 2006. Note: includes all individuals 15-65

Figure 4.4: Education Levels for the sample of earners in 2001 and 2006



Source: Tanzanian ILFS 2001 and 2006

Figure 4.5: Predicted log income by education level at the mean of all other regression variables in 2001

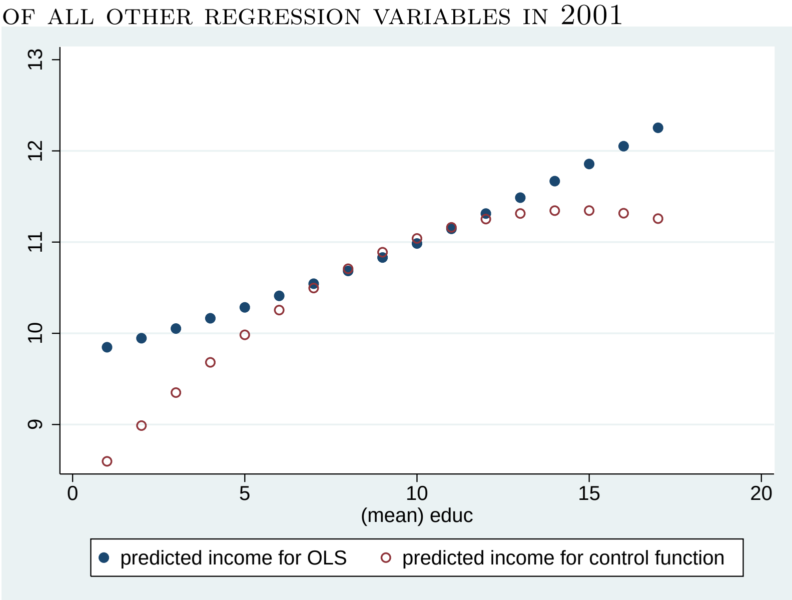
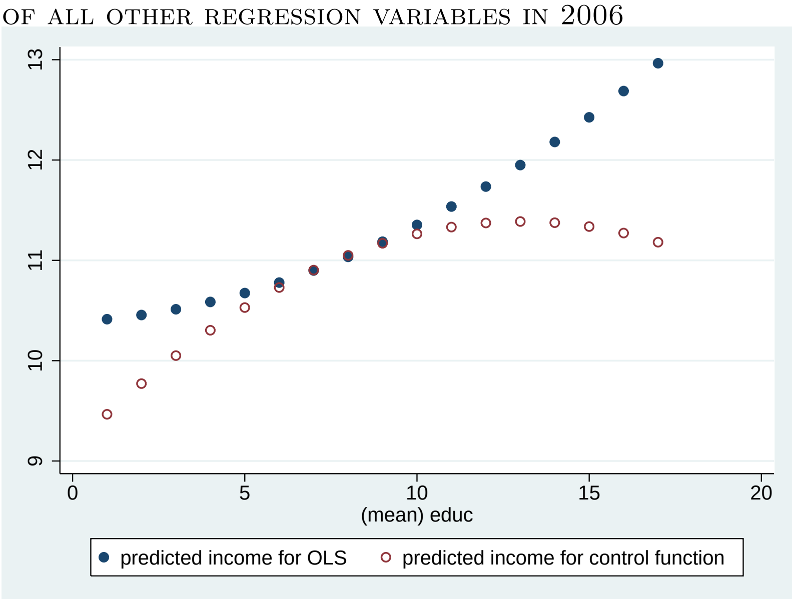


Figure 4.6: Predicted log income by education level at the mean of all other regression variables in 2006



Chapter 5

Conclusion

In this thesis I have explored the workings of the labour markets of two African countries, South Africa and Tanzania. In doing so I have shed light on the role of human capital, individual heterogeneity and institutions in determining labour market outcomes, particularly earnings and individuals' choices of where to work, in these two countries.

In chapter two I described the existence of extremely large average earnings differentials across different kinds of employment for black South Africans in the KwaZulu-Natal Income Dynamics Study data. I noted a 12-fold difference in the median earnings of government employees and the self-employed in the third wave of the survey conducted in 2004, a difference which had increased more than three times since the first wave in 1993. Using OLS to estimate the earnings differences across sectors I found that observable human capital can explain some of the differences within the private sector, but that there are still substantial public sector and union premia. The panel element of KIDS enables me to control for time invariant unobserved heterogeneity. Using panel estimators I found that the union premium is substantially diminished, whilst the public sector premium is hardly reduced at all. I argued that observable and unobservable human capital can explain most of the differences in earnings between sectors within the private sector but does not explain the public sector premium.

Further investigation revealed two interesting conclusions. Firstly, the public sector premium seems to be driven by the entry of low productivity private sector employees into the public sector. There is no decrease in earnings for those public sector employees who leave to enter the private sector, which suggests that the only workers who leave the public sector are higher productivity individuals. Secondly, I find that the low union premium estimated is not the result of measurement error that can sometimes confound estimation using panel data techniques. This suggests either that union members are unobservably more productive, or possibly that firms respond to unionisation by increasing their capital to labour ratios and hiring workers with higher skill levels. Exploring which of these is correct requires panel data on firms, which is not yet available in South Africa.

One of the most obvious differences between the Tanzanian and South African labour markets is the large number of individuals engaged in self-employment in Tanzania. The micro data from Tanzania I described in chapter three of this thesis suggest that there are not large differences in earnings between self-employment and wage employment, however, as one would expect if self-employment was a residual sector for those unable to find wage employment. Rather, the data suggest that self-employment is a viable alternative to wage employment in small firms for many individuals.

I developed a matching model to explain these outcomes, in which individual differences in self-employment productivity drive earnings determination in both sectors. The model highlights the fundamental difference in earnings determination between wage and self-employment, that in wage employment an individual bargains with a firm. This difference generates variation in earnings in both sectors, despite there being individual differences in productivity in only one of the sectors. This represents an improvement on other recent models of the labour market in developing countries, in which there is often no earnings variation within some, or even all, sectors.

In chapter four I used the Integrated Labour Force Survey data to estimate the returns to education in Tanzania. Studies using African micro data generally suggest that the returns to education are higher than in developed countries and that they are also strongly convex. I also found this pattern in Tanzania when using ordinary least squares. Previous research estimating the returns to education in Tanzania suggests that the returns have increased between 1980 and the early 1990s. My own OLS estimates suggest the returns increased again between the early 1990s and the two surveys I use from 2001 and 2006 and I showed that the increases in the 2000s are unlikely to be the result of differences in the sample composition between the data I use and that used in earlier studies.

Research using developed country data suggests that there are several reasons to mistrust the OLS estimates of the returns to education. If individuals differ in their earnings ability, ability in schooling and also have heterogeneity in their marginal returns then educational choices will be endogenous and OLS estimates are likely to reflect differences in individual heterogeneity rather than the true causal effect of education. I used an education reform that occurred in Tanzania in the 1960s as a source of exogenous variation in schooling which, under certain assumptions, allows one to estimate an average treatment effect, using the control function approach, as well as a local average treatment effect, using two-stage least squares.

Using the control function approach I found a reversal of the convexity result, with returns estimated to be concave. I still found generally high returns using the control function, though in 2006 the convexity is strong enough that the estimates of returns at secondary education are lower than previous estimates. I highlighted that for the control function approach to consistently estimate an ATE some strong parametric assumptions and the mean independence assumption are required, and noted that the parametric assumptions could be relaxed. I also found that the monotonicity assumption that is required to estimate a LATE was not satisfied in the case of the Tanzanian education reform I considered. The LATEs I estimated were generally lower than the control function and the OLS estimates, which I showed is what one would expect when the monotonicity assumption is not satisfied.

In this thesis I have highlighted the importance of accounting for individual heterogeneity in drawing conclusions from micro data about the functioning of labour markets. I have explored the role of labour market institutions, observable human capital and individual ability in earnings determination and the choice of where to work. Specific contributions have been to explore how the vast earnings inequalities in South Africa are created, instead of simply describing the presence or absence of a formal sector premium; developing what I believe to be the first matching model to highlight the implications of high levels of self-employment, inspired by descriptive evidence from the Tanzanian labour market and finally exploring the implications of individual heterogeneity for an analysis of the role of education in determining earnings and then estimating the returns to education in Tanzania.

Bibliography

Abadie, Alberto, "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, April 2003, *113* (2), 231–263.

Aghion, Philippe and Peter Howitt, "Market Structure and the Growth Process," *Review of Economic Dynamics*, January 1998, *1* (1), 276–305.

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, July 2009, *119* (539), 1105–1129.

Angrist, Joshua D. and Alan B Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," *The Quarterly Journal of Economics*, November 1991, *106* (4), 979–1014.

_ **and Guido W. Imbens**, "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 1995, *90* (430), 431–442.

_ **and Jörn-Steffan Pischke**, *Mostly Harmless Econometrics*, Princeton, NJ: Princeton University Press, 2009.

— **and** — , "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, Spring 2010, *24* (2), 3–30.

Ashenfelter, Orley C., "Estimating the Effect of Training Programs on Earnings," *The Review of Economics and Statistics*, February 1978, *60* (1), 47–57.

— **and Alan B. Krueger**, "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review*, December 1994, *84* (5), 1157–73.

Badaoui, Eliane, Eric Strobl, and Frank Walsh, "Is There An Informal Sector Wage Penalty? Evidence from South Africa," *Economic Development and Cultural Change*, 2008, *56*, 683–710.

Banerjee, Abhijit V. and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, April 1993, *101* (2), 274–98.

Becker, Gary S., "Investment in Human Capital: A Theoretical Analysis," *The Journal of Political Economy*, 1962, *70* (5), 9–49.

— , "Nobel Lecture: The Economic Way of Looking at Behavior," *Journal of Political Economy*, June 1993, *101* (3), 385–409.

Belzil, Christian, "The Return to Schooling in Structural Dynamic Models: A Survey," IZA Discussion Papers 2370, Institute for the Study of Labor (IZA) October 2006.

— **and Jorgen Hansen**, "Unobserved Ability and the Return to Schooling," *Econometrica*, September 2002, *70* (5), 2075–2091.

Bennell, Paul, "Rates of return to education: Does the conventional pattern prevail in sub-Saharan Africa?," *World Development*, January 1996, *24* (1), 183–199.

Bhorat, Haroon, Carlene van der Westhuizen, and Sumayya Goga, "Analysing Wage Formation in the South African Labour Market: The Role of Bargaining Councils," 2007. Paper presented at 2008 DPRU conference.

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky, "The Nash Bargaining Solution in Economic Modelling," *Rand Journal of Economics*, 1986, *17* (1), 176–188.

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi, "Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey," *Journal Of The Royal Statistical Society Series A*, 2005, *168* (3), 473–512.

Bommier, Antoine and Sylvie Lambert, "Education Demand and Age at School Enrollment in Tanzania," *Journal of Human Resources*, 2000, *35* (1), 177–203.

Boskin, Michael J., "A Conditional Logit Model of Occupational Choice," *The Journal of Political Economy*, 1974, *82* (2), 389–398.

Broyden, C G, "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Mathematics of Computation*, 1965, *19* (92), 577–593.

Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, May 1998, *39* (2), 257–73.

Butcher, Kristin F. and Cecilia E. Rouse, "Wage effects of unions and industrial councils in South Africa," *Industrial and Labor Relations Review*, January 2001, *54* (2), 349–374.

Cahuc, Pierre and Andre Zylberberg, *Labor Economics*, The MIT Press, 2004.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge: Cambridge University Press, 2005.

Card, David, "Using Geographic Variation in College Proximity to Estimate the Returns to Schooling," in L.N. Christofides, E.K. Grant, and R. Swidinsky, eds., *Aspects of Labour Market Behaviour: Essays in Honor of John Vanderkamp*, University of Toronto Press, 1995.

_ , "The causal effect of education on earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, April 1999, chapter 30, pp. 1801–1863.

_ , "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, September 2001, *69* (5), 1127–60.

Casale, Daniela and Dori Posel, "The Male Marital Earnings Premium in the Context of Bride Wealth Payments: Evidence from South Africa," *Economic Development and Cultural Change*, 2010, *58* (2), 211–230.

_ , Colette Muller, and Dori Posel, "Two million net new jobs? A reconsideration of the rise in employment in South Africa, 1995-2003," *The South African Journal of Economics*, 2004, *75* (2), 978–1002.

Chen, Guifu and Shigeyuki Hamori, "Economic returns to schooling in urban China: OLS and the instrumental variables approach," *China Economic Review*, June 2009, *20* (2), 143–152.

Chen, Marta, "Rethinking the Informal economy: linkages with the formal economy and the formal regulatory environment," in Basudeb Guha-Khasnobis, Ravi Kanbur, and Elinor Ostrom, eds., *Linking the Formal and Informal Economy: Concepts and Policies*, Oxford University Press 2006.

Cichello, Paul L., Gary S. Fields, and Murray Leibbrandt, "Earnings and Employment Dynamics for Africans in Post-apartheid South Africa: A Panel Study of KwaZulu-Natal," *Journal of African Economies*, June 2005, *14* (2), 143–190.

Cornwell, Katy and Brett Inder, "Language and Labour Markets in South Africa," *Journal of African Economies*, June 2008, *17* (3), 490–525.

Cragg, John G. and Stephen G. Donald, "Testing Identifiability and Specification in Instrumental Variable Models," *Econometric Theory*, April 1993, *9* (02), 222–240.

Cunha, Flavio, James Heckman, and Salvador Navarro, "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, April 2005, *57* (2), 191–261.

Deaton, Angus, *The Analysis of Household Surveys – A Microeconomic Approach to Development Policy*, Baltimore: Johns Hopkins University Press, 1997.

_ **and Alessandro Tarozzi**, "Prices and poverty in India," Working Paper 213, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies. July 2000.

Dunlop, John, "The Task of Contemporary Wage Theory," in John Dunlop, ed., *The theory of wage determination : Proceedings of a conference held by the International Economic Association*, MacMillan 1957.

Fafchamps, Marcel and Agnes R. Quisumbing, "Marriage, Bequest, and Assortative Matching in Rural Ethiopia," *Economic Development and Cultural Change*, January 2005, *53* (2), 347–80.

_ , Mans Sderbom, and Najy Benhassine, "Wage Gaps and Job Sorting in African Manufacturing," *Journal of African Economies*, November 2009, *18* (5), 824–868.

Falco, Paolo, Andrew Kerr, Neil Rankin, Justin Sandefur, and Francis Teal, "The returns to formality and informality in urban Africa.," *Journal of Labour Economics*, 2011, *18, Supplement 1*, S23 S31.

Fields, Gary S., "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs," *Journal of Development Economics*, 1975, *2* (2), 165–187.

_ , "Dualism In The Labor Market: A Perspective On The Lewis Model After Half A Century," *Manchester School*, December 2004, *72* (6), 724–735.

Fleisher, Belton, *Labor Economics: Theory and Evidence*, Prentice-Hall, 1970.

Florens, Jean-Pierre, James Heckman, Costas Meghir, and Edward Vytlacil, "Identification of Treatment Effects Using Control Functions in Models With Continuous, Endogenous Treatment and Heterogeneous Effects," *Econometrica*, September 2008, *76* (5), 1191–1206.

Garen, John, "The Returns to Schooling: A Selectivity Bias Approach With a Continuous Choice Variable," *Econometrica*, July 1984, *52*, 1199–1218.

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, January 1977, *45* (1), 1–22.

_ **and Jerry A. Hausman**, "Errors in variables in panel data," *Journal of Econometrics*, February 1986, *31* (1), 93–118.

Guha-Khasnobis, Basudeb, Ravi Kanbur, and Elinor Ostrom, "Beyond Formality and Informality," in Basudeb Guha-Khasnobis, Ravi Kanbur, and Elinor Ostrom, eds., *Linking the Formal and Informal Economy: Concepts and Policies*, Oxford University Press 2006.

Harris, John and Michael Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 1970, *60* (1), 126–142.

Hart, Keith, "Informal Income Opportunities and Urban Employment in Ghana," *The Journal of Modern African Studies*, 1973, *11*, 61–89.

Heckman, James J., "Effects of Child-Care Programs on Women's Work Effort," *Journal of Political Economy*, 1974, *82* (2), S136–S163.

_ , "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources*, 1997, *32* (3), 441–462.

_ , "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, August 2001, *109* (4), 673–748.

_ **and Edward Vytlacil**, "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling," *Journal of Human Resources*, 1998, *33* (4), 974–987.

_ **and** _ , "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, May 2005, *73* (3), 669–738.

Heintz, James and Dori Posel, "Revisiting informal employment and segmentation in the South African labour market," *South African Journal of Economics*, 2008, *76* (1), 26–44.

Hussmans, Ralph, "Measuring the informal economy: From employment in the informal sector to informal employment. ILO Working Paper 53," Technical Report, ILO, Washington, DC 2004.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, March 1994, *62* (2), 467–75.

International Labour Office, *Employment, incomes and equality: A strategy for increasing productive employment in Kenya*, Geneva: International Labour Organization, 1972.

International Monetary Fund, "Internation Financial Statistics September 2009, Series 199..AF.ZF," 2009.

_ , "Internation Financial Statistics April 2010, Series ..DE.ZF," 2010.

James, Jeffrey, "Public Choice, Technology and Industrialization in Tanzania: Some Paradoxes Resolved," *Public Choice*, December 1996, *89* (3-4), 375–92.

Kerr, Clarke, "The Balkanization of Labor Markets," in Bakke, E. Wight, Philip M. Hauser, Clark Kerr, Charles A. Myers, Gladys L. Palmer, Paul Webbink, Dale Yoder, ed., *Labor Mobility and Economic Opportunity*, Wiley 1954.

Keswell, Malcolm and Laura Poswell, "Returns To Education In South Africa: A Retrospective Sensitivity Analysis Of The Available Evidence," *South African Journal of Economics*, 2004, *72* (4), 834–860.

Kingdon, Geeta and John Knight, "Unemployment in South Africa: The Nature of the Beast," *World Development*, March 2004, *32* (3), 391–408.

_ , Justin Sandefur, and Francis Teal, "Labour Market Flexibility, Wages and Incomes in Sub-Saharan Africa in the 1990s," *African Development Review*, 2006, *18* (3), 392–427.

Knight, John and Richard Sabot, *Education, Productivity and Inequality, The East African Natural Experiment*, Oxford: Oxford University Press, 1990.

Krueger, Alan B. and Mikael Lindahl, "Education for Growth: Why and for Whom?," *Journal of Economic Literature*, December 2001, *39* (4), 1101–1136.

Leontaridi, Marianthi Rannia, "Segmented Labour Markets: Theory and Evidence," *Journal of Economic Surveys*, February 1998, *12* (1), 63–101.

Magruder, Jeremy, "High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa," 2009. University of California, Berkeley, Working Paper.

Maloney, William, "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico," *The World Bank Economic Review*, 1999, *13* (2), 275–302.

_ , "Informality Revisited," *World Development*, July 2004, *32* (7), 1159–1178.

Maluccio, John, "Attrition in the KwaZulu-Natal Income Dynamics Study, 1993-1998," 2000. FCND Discussion Paper No. 95, IFPRI.

May, Julian, Jorge Agüero, Michael R. Carter, and Ian M. Timaeus, "The KwaZulu-Natal Income Dynamics Study (KIDS) 3rd wave: methods, first findings and an agenda for future research.," *Development Southern Africa*, 2007, *24*, 629 – 648.

_ **, Michael R. Carter, Lawrence. Haddad, and John Maluccio**, "KwaZulu-Natal Income Dynamics Study (KIDS) 1993-98: a longitudinal household database for South African policy analysis," *Development Southern Africa*, 2000, *17*, 567–581.

Mincer, Jacob, "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 1958, *66*, 281–302.

_ , *Schooling, Experience and Earnings*, Columbia University Press, 1973.

Ministry of Education, Tanzania, "Annual Reports," 1960-1969.

Miranda, Mario and Paul Fackler, *Applied Computational Economics and Finance*, MIT Press, 2002.

Moll, Peter, "Compulsory Centralization of Collective Bargaining in South Africa," *American Economic Review*, 1996, *86*, 326–329.

— , "Primary Schooling, Cognitive Skills and Wages in South Africa," *Economica*, May 1998, *65* (258), 263–84.

Morrison, David, *Education and Politics in East Africa, The Tanzanian Case*, London: C. Hurst & Co Ltd., 1976.

Mortensen, Dale T., *Wage dispersion : why are similar workers paid differently*, Cambridge, MA. London: MIT Press, 2003.

— **and Christopher A. Pissarides**, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, July 1994, *61* (3), 397–415.

— **and** — , "New developments in models of search in the labor market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, April 1999, chapter 39, pp. 2567–2627.

Muller, Colette, "Measuring South Africas Informal Sector: An Analysis of National Household Surveys," 2003. DPRU Working Paper.

Nash, John, "Two-person cooperative games," *Econometrica*, 1953, *21* (1), 128–140.

National Bureau of Statistics, Tanzania, "Household Budget Survey 2007 Analytical Report," 2007.

Orazem, Peter F. and Elizabeth M. King, "Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy," in T. Paul Schultz and John A. Strauss, eds., *Handbook of Labor Economics*, Vol. 4, Elsevier, 2008, chapter 55, pp. 3475–3559.

Oyelere, Ruth, "Understanding Low Average Returns to Education in Africa: The Role of Heterogeneity across Education Levels and the Importance of Political and Economic Reforms," IZA Discussion Paper 3766, Institute for the Study of Labor (IZA) Oct 2008.

Pereira, Pedro Telhado and Pedro Silva Martins, "Returns to education and wage equations," *Applied Economics*, April 2004, *36* (6), 525–531.

Pigou, Arthur C., *A Study in Public Finance*, Macmillan, 1928.

Polachek, Solomon W., "Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications," IZA Discussion Papers 3181, Institute for the Study of Labor (IZA) November 2007.

Posel, Dori and Daniela Casale, "Internal migration and household poverty in post-apartheid South Africa," in Ravi Kanbur and Haroon Bhorat, eds., *Poverty and Policy in Post-Apartheid South Africa*, HSRC Press 2006.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, November 2002, *70* (6), 2295–2350.

Psacharopoulos, George, "Returns to investment in education: A global update," *World Development*, September 1994, *22* (9), 1325–1343.

Rankin, Neil, Justin Sandefur, and Francis Teal, "Learning and earning in Africa: Where are the returns to education high?," 2010. CSAE Working Paper WPS/2010-02.

Roy, A, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Sabot, Richard H., *Economic Development and Urban Migration: Tanzania 1900-1971*, Oxford: Oxford University Press, 1979.

Sandefur, Justin, Pieter Serneels, and Francis Teal, "African Poverty through the Lens of Labor Economics: Earnings and Mobility in Three Countries," in Pierella Paci and Pieter Serneels, eds., *Employment and Shared Growth*, World Bank, 2007.

Satchi, Mathan and Jonathan Temple, "Labor Markets and Productivity in Developing Countries," *Review of Economic Dynamics*, January 2009, *12* (1), 183–204.

Sattinger, Michael, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, June 1993, *31* (2), 831–80.

Schultz, T. Paul and Germano Mwabu, "Labor unions and the distribution of wages and employment in South Africa," *Industrial and Labor Relations Review*, July 1998, *51* (4), 680–703.

Schultz, Theodore W., "The Emerging Economic Scene and its Relation to High School Education," in F.S. Chase and H.A. Anderson, eds., *The High School in a New Era*, University of Chicago Press, 1958.

_ , "Investment in Man: An Economist's View," *Social Service Review*, Mar/Dec 1959, *33* (1), 109–117.

Söderbom, Måns, Francis Teal, Anthony Wambugu, and Godius Kahyarara, "The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing," *Oxford Bulletin of Economics and Statistics*, June 2006, *68* (3), 261–288.

Spence, A. Michael, "Job Market Signaling," *The Quarterly Journal of Economics*, August 1973, *87* (3), 355–74.

Taubman, Paul, "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins," *American Economic Review*, 1976, *66* (5), 848–870.

van der Berg, Servaas and Megan Louw, "Changing Patterns Of South African Income Distribution: Towards Time Series Estimates Of Distribution And Poverty," *South African Journal of Economics*, September 2004, *72* (3), 546–572.

Wald, A, "The Fitting of Straight Lines if Both Variables are Subject to Error," *Annals of Mathematical Statistics*, 1940, *11*, 284–300.

Walter, McMahon, "The Social and External Benefits of Education," in Geraint Johnes and Jill Johnes, eds., *International Handbook on the Economics of Education*, Edward Elgar 2007.

Wong, Linda Y., "Can the Mortensen-Pissarides Model with Productivity Changes Explain U.S. Wage Inequality?," *Journal of Labor Economics*, January 2003, *21* (1), 70–105.

Wooldridge, Jeffrey M., "On two stage least squares estimation of the average treatment effect in a random coefficient model," *Economics Letters*, October 1997, *56* (2), 129–133.

_ , *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA.: MIT Press, 2002.

_ , "Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model," *Economics Letters*, May 2003, *79* (2), 185–191.

World Bank, "Priorities and Strategies for Education: A World Bank Sector Review, Education and Social Policy Department," 1995. Washington DC.

_ , "Joint World Bank/IMF Debt Sustainability Analysis on United Republic of Tanzania," 2009.

Zenou, Yves, "Job search and mobility in developing countries. Theory and policy implications," *Journal of Development Economics*, June 2008, *86* (2), 336–355.