

Determinants of Earnings Inequalities: Panel data evidence from South Africa

Andrew Kerr*

September 26, 2011

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. Human capital theory posits that workers' productivities explain their earnings, and sector choices, and implies that sectoral 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 can generate wage differentials between the sectors for otherwise identical workers, which is described as segmentation.

In this chapter I test these alternative explanations using a panel data set that allows me to explore the role of observed and unobserved heterogeneity in determining the outcomes observed in the South African labour market. The data set also allows me to analyse the relative contributions of observed and unobserved human capital, as well as 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 analysis 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.

*DataFirst, University of Cape Town: andrew.kerr@uct.ac.za

But given the large differences in labour market outcomes in South Africa across all racial groups, I have chosen to further limit my analysis to black workers.

The structure of the chapter is as follows: Section 2 reviews evidence on possible explanations of earnings differentials in South Africa. The KwaZulu-Natal Income Dynamics Study is described in Section 3. Section 4 tests the ability of both human capital theory and segmentation hypotheses to explain earnings, as well as analysing the effects of allowing for unobserved heterogeneity to influence earnings. The robustness of the results to corrections for attrition is considered in Section 5. Section 6 concludes.

2 Identifying Sources of Wage 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 the wage paid. It is argued that this results in wage differentials across sectors that are inconsistent with the competitive model.

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 any unobserved individual productivity-enhancing abilities. In this model any wage differentials across 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.

Allowing for 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 rapidly became a prominent feature of labour economics (Card 1999). Wages are generally observed to increase with education, training and experi-

ence (Mincer 1973), evidence in favour of the basic tenants of human capital theory. I explore the 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 types of employment. What has been called ‘occupational choice’ has been incorporated into modern human capital theory (cf Boskin (1974) and Fleisher (1970)). In this extensions individuals choose their education and training levels, as well as their sector of employment, to maximise their lifetime utility, based on their (possibly unobserved) productive capabilities in each sector and the benefits and costs of education and training required to enter each sector. One should thus observe workers in occupations where their expected future lifetime utility is highest, although Boskin (1974) does consider the role of imperfect capital markets. 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 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, were 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 develop 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 Segmented Labour Market Theories

Despite the success of the competitive model with human capital in explaining wages, other economists have emphasised that the competitive model was not sufficient to explain wage differentials between individuals (Leontaridi 1998). This has been argued to begin with the American Institutional school of thought (Dunlop (1957), Kerr (1954)) and was further developed in the 1960s as social scientists grappled with the new micro data available to them, who argued that similar workers were observed to earn substantially different amounts depending on which sector they worked in (Leontaridi 1998). Instead of explaining these within the neoclassical paradigm, as occupational choice did, some economists emphasised institutional constraints in labour markets that prevented the bargaining down of wages in high wage sectors by workers in low wage sectors and uncoupled the productivity-wage relationship central to human capital theory. Institutional constraints were argued to lead to different

wage setting mechanisms in different sectors in the market and to similar workers being paid different wages, depending on which part of the market they were able to access.

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, because of 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 conditional on finding formal sector employment 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 with the 1972 International Labor Organization (ILO) study in Kenya (ILO 1972) and anthropological work by Hart (1973) in urban Ghana. This highlighted that many urban residents were making a living beyond the control of the state and 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, or casual employment in informal enterprises, and was considered a free entry sector. This meant that anyone was 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 it 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 meant 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 African labour markets.

In the wake of the early work on the informal sector there has been a large volume of research on it in economics, anthropology and other disciplines. 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” (ILO 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, as well as 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 this hypothesis in my empirical analysis below, in addition to exploring the role of human capital.

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, with tertiary education associated with marginal returns of over 50 percent a year when using OLS and age earnings profiles suggesting 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 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, that show 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 times¹.

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 pertaining to Bargaining Councils, which can be created by agreement of trade unions representing 50% of workers in any industry and firms employing 50% 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 % higher than industries in neighbouring areas without these agreements, and argues that the Bargaining Council system results in lower employment and smaller firm size of firms in regions and industries covered by Bargaining Council agreements.

Moll (1996) explores the impact of Industrial Councils, the forerunners of Bargaining Councils, and provides some descriptive evidence that these created wage premia even for non-union workers covered by a 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 % 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, through their effect on firm level bargaining, and in addition to their role in facilitating Bargaining Council agreements. Union premiums of 60 percent in the 1993 PSLSD

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

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, that this reduces the premium to around 20%, which is comparable with premiums in OECD countries, and hence 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 actually be explained by individual differences in unobserved ability or unobserved school quality, however. The only study that takes into account unobserved heterogeneity using South African data is Badaoui et al. (2008). The authors use the Labour Force Survey panel 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 this study to my undertakings in this chapter. Firstly, Badaoui et al. (2008) 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 this paper tells one nothing about *how* this result is obtained or the institutions or regulations that 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.

Second, 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* 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.

3 Description of the KIDS Data and Survey Methodology

This section provides a brief description of the KwaZulu-Natal Income Dynamics Study (KIDS), a 3 wave panel conducted in 1993, 1998 and 2004 in KwaZulu-Natal province, which was the largest province by population when the survey was undertaken² 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 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 settlement 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 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.

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 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).

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

³The 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

The questions asked in KIDS do not allow for the categorisation of jobs according to the ILO or Statistics South Africa definitions ⁴. Following Chen (2006), however, the KIDS data do allow me to explore the effects of public sector wage setting and unionisation on earnings, which can be argued to be two important dimensions of formality. 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 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 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 1 also shows the large hourly earnings differences between regular, casual and self-employment that I highlighted at the start of this chapter and that 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 paying 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 how much 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 three-fold since 1993. This increase in the earnings gap over the 11 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% for unionised public sector employees.

⁴Further details of the relevant labour market questions asked in the KIDS are provided in the Appendix to this chapter

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

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 raising of the wages of black individuals in the public sector to the level of their white counterparts as a result of the new post-1994 government, or the effects of the formation of the Public Service Co-ordinating Bargaining Council, which occurred in 1997 (Bhorat et al. 2007). Table 1 also shows that between 1998 and 2004 real wages remained constant or decreased slightly. Interestingly, Table 3 shows that, on average, those working in casual or self-employment have similar weekly working hours to those in regular employment. This means that despite similar hours worked, 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 governments major challenges. Significant educational expansion between 1993 and 2004 is evident in the KIDS data, as shown in Table 3. Public sector employees have the highest levels of education across the 3 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, 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 following children of core members⁶. The large attrition rate is a concern and this may be the reason some of the changes in employment and earnings over the 3 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 5. I now describe the attrition in KIDS in greater detail.

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 for four main reasons. The first is that 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 re-

⁶Core members are defined in the Appendix to this chapter

leased survey data. The second reason would be because 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 is that the survey team could find no-one in the community the household was living in 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 that the survey team found no trace of would probably also fall into this fourth category.

In addition to the exit of households from the sample, it was also possible for individuals to exit 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 and 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 1998. 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 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 a household member owned the household's dwelling 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 5 shows some correlates of individual attrition (all individuals, including those who exit the sample due to household attrition, older than 15). 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: this is because the survey methodology meant core household members were tracked if they moved, whilst other household members were not. Table 5 shows similar 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 5, when I explore the effects of attrition on the results I obtain in the following sections.

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.

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 7 shows basic OLS earnings functions for the 3 waves of the panel, with the log of real hourly wage being the dependent variable⁷. Missing earnings or hours data, for example 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. I experimented with linear, squared and cubic continuous education variables, but these restrictions were not supported by the data.

⁷The 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 6 in the Appendix to this chapter shows basic summary statistics for the variables included in the regressions.

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 years of labour market experience would result in a 40 year old gaining a 55% 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 3, are substantially reduced in the OLS earnings functions for each of the 3 waves, conditioning on a range of covariates. The six-fold raw difference in average earnings over the 3 waves of KIDS 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 explains earnings differentials within the private sector, but not between the public and private sectors. The union premium in the private sector is estimated to be around 60%, but is roughly 30% when I control for industry and occupation, shown in the last column of Table 7. This is consistent with the analysis of Butcher and Rouse (2001)⁸.

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 the wage as an individual moves between 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 8 and 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

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}. \quad (1)$$

The α_i terms are the individual specific effects. x_{it} and β are $K \times 1$ vectors, and I assume the ε_{it} are iid $[0, \sigma^2]$. The fixed effects estimator is obtained by

⁸With 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 that I discussed in Section 2

subtracting the time averaged model from the original model, giving

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i). \quad (2)$$

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} , ie that $E[\varepsilon_{it}|x_{it}, \alpha_i] = 0$. Alternatively one can estimate the fixed effects model using the first difference estimator, subtracting the model lagged one period, to obtain

$$y_{it} - y_{i,t-1} = (x_{it} - x_{i,t-1})' \beta + (\varepsilon_{it} - \varepsilon_{i,t-1}). \quad (3)$$

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.

4.3 The effects of unobserved heterogeneity

Table 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 3 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 91 percent more than those in private, 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 significant⁹, two and a half fold difference between the highest and lowest paying sectors, suggesting that controlling for unobserved heterogeneity further reduces the differentials within the private sector, such that the differences are not significant in the fixed effects regression. 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, but that the private sector union premium is much lower than in the OLS regressions and insignificant.

The small union premium in the private sector is an important finding, as it suggests that the larger union premium in the cross section results may

⁹Table 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

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 having to pay higher wages, by increasing their capital stock and lowering their labour usage but hiring 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 for 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 section results also hold for the sub-sample of individuals observed more than once. The last column in Table 7 shows the key OLS results are indeed similar in the reduced sample, with earnings increasing as the degree of formality of employment increases. I now explore some other sources of bias in my results.

4.4 Measurement Error using 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:

$$y_{it} = \alpha_i + x_{it}^* \beta + \varepsilon_{it}. \quad (4)$$

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

$$\Delta y_{it} = \beta \Delta x_{it}^* + \Delta \mu_{it} \quad (5)$$

$$= \beta \Delta x_{it} + \Delta \mu_{it} - \beta \Delta v_{it}. \quad (6)$$

If I define $\rho = \text{Cor}(x_{it}^*, x_{i,t-1}^*)$, then it can be shown that

$$plim \hat{\beta} = \beta + (plim \frac{1}{N} \sum_{i=1}^N \Delta x_{it}^2)^{-1} plim \frac{1}{N} \sum_{i=1}^N (\Delta x_{it} \Delta \mu_{it} - \beta \Delta x_{it} v_{it}) \quad (7)$$

$$= \beta - \frac{\beta \sigma_v^2}{(1 - \rho) \sigma_{x^*}^2 + \sigma_{v_2}^2}. \quad (8)$$

This inconsistency is larger than in the cross sectional case when $\rho > 0$. It is also clear from equation (8) that as $\rho \rightarrow 1$ the inconsistency becomes large.

KIDS has longer periods between waves than many other panel surveys, with 5 years between the first 2 waves and 6 years between the second and third waves. This would help to alleviate attenuation bias if the dependent 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 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 10 shows the sector coefficients in the second difference results are generally not significant, probably 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 error¹⁰.

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 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 second column in Table 12 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 1. Column 3 shows the results excluding those who enter the public sector, which are radically different, with no significant public sector premium¹¹. Columns 4-6 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 workers, 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

¹⁰Butcher 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

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

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.

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 13 and 14 show OLS and fixed effects and differenced estimates using gross pay for those who were asked about it and reported it. The results are similar to those using net pay in Tables 7 and 10 though there are now larger differences between self and casual employment and the other categories where gross and net pay differ. However the union premium in the private sector is much larger and significant 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.

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 15 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.

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 are corrected using a Heckman selection correction, following Wooldridge (2002).

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} :

$$s_{it} = 1[w_{it}\delta + v_{it} > 0]. \quad (9)$$

Tables 17 and 18 in the Appendix to this chapter¹² shows the results of 3 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 members 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 3 also showed, which is indicative of the tracking methodology that households that split were tracked if these contained “core” members, as defined in the Appendix to this chapter. 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.

Results from the third column, modelling individual attrition in 2004 using 1998 individual, household and community characteristics are similar. 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 actually seem not to, for example high levels of reported HIV/AIDS or violence do not affect the probability of an individual exiting the sample.

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

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 17 and 18 models 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.

5.2 Correcting for Possible Attrition Bias

In Section 4 individual earnings were modeled as

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}. \quad (10)$$

In correcting for possible attrition bias both Inverse Probability Weighting for the pooled cross section and the Heckman selection correction of the differenced regressions rely on slightly different versions of the so-called “ignorability” assumption (Wooldridge 2002):

$$P(s_{it} = 1|y_{it}, x_{it}, w_{i1}) = P(s_{it} = 1|w_{i1}). \quad (11)$$

This means that the y_{it} and x_{it} must not influence attrition once the impact of the w_{i1} has been accounted for.

The Heckman selection correction allow for an unobserved effect in the earnings equation but not in the attrition equation, and the IPW method does not allow for an unobserved effect in either equation. If attrition is a function of unobservables then these methods are not valid. The Heckman correction also requires the strict exogeneity of the x_{it} . In addition, both methods assume the normality of the errors in the selection equation, a strong distributional assumption (Wooldridge 2002). Methods that do not require such strong assumptions do exist, but I do not attempt to implement these here.

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. The coefficients can then be estimated consistently, if the selection on observables assumption holds. These results are shown in the

second to last column of Table 7, and are not significantly different from the uncorrected results. An optimistic interpretation of this result is that attrition is not a serious concern. It could also indicate, however, that the selection on observables assumption has been violated, given the fairly low predictive power of the variables in the attrition model.

In allowing for attrition in the differenced estimation, I follow Wooldridge (2002) and assume joint normality of $\Delta\varepsilon_{it}$ 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. These are fairly strong assumptions but are commonly used in the literature. It can then be shown that

$$E(\Delta y_{it} | \Delta x_{it}, w_{it}, s_{it} = 1) = \beta \Delta x_{it} + \rho \lambda(w_{it} \delta). \quad (12)$$

I can then use the probit regressions to model the selection process, and the inverse Mills ratio is used as the correction term which generates consistent estimates using the differenced estimator.

Table 16 shows the first and second differenced equations corrected for selection into earnings in future periods, with those who attrite not having an earnings observation in future periods. The results again do not differ dramatically compared 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 16. The results are similar to the uncorrected and attrition corrected results.

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 the panel dimension of KIDS enables one to answer 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 2 and 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 was driven by those entering the public sector, rather than those leaving it, and I have 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 either that 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.

Appendix

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. when they are together they share food from a common source (i.e. they cook and eat together)
3. contribute 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. However one important difference 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.

Table 1: MEDIAN HOURLY EARNINGS

	1993		1998		2004	
	<u>Median N</u>		<u>Median N</u>		<u>Median N</u>	
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: 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 3: MEAN WEEKLY HOURS WORKED AND EDUCATIONAL ATTAINMENT (YEARS)

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

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

Table 4: HH ATTRITION DESCRIPTIVE STATISTICS

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

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

Table 5: INDIV ATTRITION DESCRIPTIVE STATISTICS

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

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

Table 6: EARNER SAMPLE SUMMARY STATISTICS

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

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

Table 7: OLS EARNINGS FUNCTIONS

	93	98	04	pooled	IO
	(1)	(2)	(3)	(4)	(5)
Age	0.064 (0.015)***	0.039 (0.019)**	0.065 (0.02)***	0.062 (0.01)***	0.056 (0.01)***
Age ²	-0.0006 (0.0002)***	-0.0003 (0.0002)	-0.0007 (0.0002)***	-0.0006 (0.0001)***	-0.0006 (0.0001)***
Incomplete prim educ	0.26 (0.094)***	0.527 (0.127)***	0.222 (0.142)	0.302 (0.068)***	0.227 (0.066)***
Complete prim educ	0.442 (0.124)***	0.791 (0.154)***	0.382 (0.176)**	0.504 (0.085)***	0.4 (0.084)***
Incomplete secondary educ	0.691 (0.097)***	1.073 (0.128)***	0.558 (0.137)***	0.738 (0.068)***	0.547 (0.069)***
Complete secondary educ	0.884 (0.139)***	1.422 (0.145)***	0.822 (0.151)***	1.040 (0.081)***	0.73 (0.084)***
Tertiary educ	1.515 (0.187)***	1.731 (0.208)***	1.429 (0.177)***	1.537 (0.107)***	1.063 (0.119)***
Married	0.178 (0.07)**	-0.131 (0.076)*	0.113 (0.084)	0.061 (0.044)	0.053 (0.043)
Male	0.439 (0.063)***	0.316 (0.068)***	0.124 (0.074)*	0.299 (0.039)***	0.252 (0.04)***
Public Union	0.615 (0.152)***	0.977 (0.127)***	0.947 (0.132)***	0.896 (0.076)***	0.69 (0.08)***
Public non-Union	0.444 (0.122)***	0.808 (0.125)***	0.869 (0.177)***	0.664 (0.077)***	0.486 (0.08)***
Private Union	0.517 (0.09)***	0.504 (0.103)***	0.456 (0.15)***	0.506 (0.062)***	0.278 (0.066)***
Casual Employment	-0.452 (0.11)***	-0.147 (0.092)	-0.325 (0.102)***	-0.283 (0.055)***	-0.246 (0.067)***
Self employment	-0.449 (0.084)***	-0.144 (0.114)	-0.849 (0.133)***	-0.463 (0.061)***	-0.375 (0.074)***
Const.	-0.605 (0.302)**	-0.080 (0.403)	-0.141 (0.432)	-0.537 (0.208)***	-0.353 (0.215)
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.

Table 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

Source: own calculations from KIDS.

Table 9: EARNER MOVEMENT 98-04

	regprivatenonunion	regpubunion	regpubnonunion	regprivatunion	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
regprivatunion	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

Source: Own calculations from KIDS.

Table 10: FIXED EFFECTS AND DIFFERENCED EARNINGS FUNCTIONS

	Fixed Effects	First Difference	Second Difference
	(1)	(2)	(3)
Age	0.083 (0.042)**	0.064 (0.043)	-0.369 (0.32)
Age ²	-0.001 (0.0005)**	-0.0009 (0.0005)*	0.003 (0.003)
Public Union	0.654 (0.167)***	0.488 (0.157)***	0.391 (0.359)
Public non-Union	0.484 (0.16)***	0.352 (0.166)**	0.753 (0.401)*
Private Union	0.169 (0.095)*	0.125 (0.099)	-0.033 (0.233)
Casual Employment	-0.207 (0.116)*	-0.271 (0.122)**	-0.407 (0.268)
Self employment	-0.108 (0.166)	-0.129 (0.179)	-0.274 (0.269)
1998 Year Dummy	0.411 (0.104)***	0.547 (0.109)***	
2004 Year Dummy	0.289 (0.194)		
Const.	0.184 (0.9)	-0.098 (0.137)	-0.793 (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 11: F TESTS OF EQUALITY OF SECTORAL COEFFICIENTS FROM FIXED EFFECTS REGRESSION

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

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

Table 12: FIRST DIFFERENCED EARNINGS FUNCTIONS EXCLUDING MOVERS INTO AND OUT OF PUBLIC SECTOR EMPLOYMENT

	Orig	Exclout	Exclinto	Pub	PubExclout	PubExclinto
	(1)	(2)	(3)	(4)	(5)	(6)
Δ age	0.049 (0.042)	0.045 (0.042)	0.044 (0.044)	0.048 (0.042)	0.046 (0.042)	0.043 (0.044)
Δ agesq	-.0007 (0.0004)	-.0007 (0.0004)*	-.0005 (0.0004)	-.0007 (0.0004)*	-.0007 (0.0004)*	-.0005 (0.0004)
Δ pubunionD	0.461 (0.153)***	0.423 (0.183)**	0.155 (0.209)			
Δ pubnonunionD	0.344 (0.153)**	0.462 (0.195)**	0.039 (0.196)			
Δ publicD				0.32 (0.122)***	0.335 (0.157)**	0.031 (0.174)
Δ privunionD	0.125 (0.114)	0.107 (0.116)	0.144 (0.12)			
Δ casearner1	-.255 (0.116)**	-.296 (0.119)**	-.170 (0.119)	-.304 (0.11)***	-.342 (0.113)***	-.215 (0.113)*
Δ selfearner	-.147 (0.158)	-.155 (0.159)	-.181 (0.171)	-.177 (0.156)	-.182 (0.158)	-.204 (0.169)
cons	-.107 (0.14)	-.019 (0.147)	-.222 (0.148)	-.084 (0.139)	-.002 (0.146)	-.202 (0.148)
Obs	628	585	570	628	585	570
R^2	0.084	0.076	0.062	0.079	0.072	0.059

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

Table 13: OLS GROSS HOURLY EARNINGS

	93	98	04
	(1)	(2)	(3)
Incomplete prim educ	0.534 (0.134)***	0.256 (0.159)	0.381 (0.076)***
Complete prim educ	0.789 (0.164)***	0.361 (0.201)*	0.535 (0.097)***
Incomplete sec educ	1.088 (0.134)***	0.62 (0.152)***	0.824 (0.076)***
Complete sec educ	1.526 (0.153)***	0.819 (0.17)***	1.149 (0.092)***
Tertiary educ	1.840 (0.216)***	1.472 (0.203)***	1.659 (0.122)***
Married	-.067 (0.079)	0.113 (0.094)	0.113 (0.049)**
Male	0.32 (0.072)***	0.221 (0.084)***	0.323 (0.044)***
Age	0.073 (0.024)***	0.042 (0.029)	0.071 (0.014)***
Age ²	-.0008 (0.0003)***	-.0004 (0.0004)	-.0008 (0.0002)***
Public Union	0.982 (0.132)***	1.280 (0.15)***	1.084 (0.086)***
Public non-Union	0.92 (0.13)***	0.818 (0.196)***	0.799 (0.086)***
Private Union	0.56 (0.106)***	0.462 (0.169)***	0.579 (0.068)***
Casual Employment	-.317 (0.1)***	-.512 (0.117)***	-.389 (0.063)***
Self employment	-.318 (0.118)***	-1.083 (0.151)***	-.541 (0.068)***
Const.	-.586 (0.469)	0.362 (0.575)	-.779 (0.268)***
Obs.	821	632	2369
R^2	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 14: FIXED EFFECTS AND DIFFERENCED EARNINGS FUNCTIONS USING GROSS EARNINGS

	Fixed Effects	First Difference	Second Difference
	(1)	(2)	(3)
Age	0.107 (0.049)**	0.112 (0.051)**	-.391 (0.366)
Age ²	-.001 (0.0005)**	-.001 (0.0006)**	0.003 (0.004)
Public Union	0.769 (0.164)***	0.613 (0.176)***	0.659 (0.391)*
Public non-Union	0.513 (0.171)***	0.334 (0.212)	0.969 (0.491)**
Private Union	0.307 (0.114)***	0.319 (0.133)**	0.56 (0.328)*
Casual Employment	-.212 (0.14)	-.236 (0.149)	-.173 (0.314)
Self employment	-.158 (0.182)	-.128 (0.203)	0.214 (0.492)
1998 Year Dummy	0.479 (0.122)***	0.853 (0.13)***	
2004 Year Dummy	0.238 (0.225)		
Const.	-.337 (1.040)	-.238 (0.155)	-1.200 (0.306)***
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 15: MALE AND FEMALE REGRESSIONS

	OLS Female	OLS Male	Fixed Effects Female	Fixed Effects Male
	(1)	(2)	(3)	(4)
Age	0.053 (0.015)***	0.069 (0.015)***	0.073 (0.043)*	0.138 (0.042)***
Age ²	-0.0005 (0.0002)***	-0.0007 (0.0002)***	-0.0005 (0.0005)	-0.002 (0.0005)***
Incomplete prim educ	0.142 (0.095)	0.465 (0.096)***		
Complete prim educ	0.482 (0.12)***	0.525 (0.121)***		
Incomplete sec educ	0.708 (0.094)***	0.775 (0.097)***		
Complete sec educ	0.994 (0.114)***	1.078 (0.115)***		
Tertiary educ	1.567 (0.142)***	1.399 (0.165)***		
Married	0.007 (0.058)	0.069 (0.068)		
Public Union	0.971 (0.118)***	0.831 (0.1)***	0.631 (0.258)**	0.676 (0.214)***
Public non-Union	0.75 (0.112)***	0.557 (0.106)***	0.523 (0.227)**	0.419 (0.225)*
Private Union	0.411 (0.098)***	0.536 (0.08)***	0.104 (0.183)	0.16 (0.152)
Casual Employment	-0.244 (0.077)***	-0.353 (0.08)***	-0.328 (0.155)**	-0.083 (0.168)
Self employment	-0.528 (0.08)***	-0.339 (0.094)***	-0.047 (0.242)	-0.206 (0.194)
Const.	-0.301 (0.291)	-0.431 (0.292)	-0.372 (0.879)	-0.792 (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 included but not reported. Reference employment category is private non-union regular employment.

Table 16: ATTRITION CORRECTED DIFFERENCED EARNINGS FUNCTIONS

	First Diff	Second Diff	FirstDiff93 98	FirstDiff98 04
	(1)	(2)	(3)	(4)
Age	0.055 (0.046)	-.434 (0.358)	0.05 (0.064)	-.010 (0.068)
Age ²	-.0007 (0.0004)	0.004 (0.003)	-.0009 (0.0005)*	0.00005 (0.0007)
Public Union	0.438 (0.156)***	0.473 (0.338)	0.392 (0.189)**	0.641 (0.261)**
Public non-Union	0.294 (0.158)*	1.067 (0.375)***	0.133 (0.186)	0.778 (0.271)***
Private Union	0.097 (0.119)	0.028 (0.297)	0.088 (0.129)	0.251 (0.233)
Casual	-.265 (0.121)**	-.281 (0.291)	-.138 (0.158)	-.345 (0.175)**
Self employed	-.176 (0.158)	-.421 (0.428)	-.147 (0.182)	-.121 (0.318)
1998 Year Dummy	0.619 (0.113)***			
lambda9898	-.205 (0.133)			
lambda0404	0.328 (0.113)***			
lambda93930404		-.390 (0.493)		
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 17: INDIVIDUAL ATTRITION PROBIT REGRESSIONS

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

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Marginal Affects Reported. Regression continued in Table 18

Table 18: INDIVIDUAL ATTRITION PROBIT REGRESSIONS CONTINUED

	Attrition 93-98	Attrition 98-2004	Attrition 93-2004
	(1)	(2)	(3)
Core hh member	-0.478 (0.072)***	-0.563 (0.079)***	-0.406 (0.066)***
Log(residents)	-0.301 (0.077)***	-0.456 (0.059)***	-0.399 (0.034)***
Log(non-residents)	-0.019 (0.073)	-0.075 (0.059)	-0.044 (0.031)
Age of hh Head	0.003 (0.003)		0.0005 (0.001)
HH head is female	0.011 (0.079)	0.025 (0.07)	-0.018 (0.035)
Max educ level in hh	-0.002 (0.015)	-0.006 (0.018)	-0.031 (0.007)***
HH has flush toilet	0.323 (0.221)	-0.234 (0.238)	0.097 (0.118)
2nd income quintile	0.116 (0.108)	-0.097 (0.094)	0.142 (0.044)***
3rd income quintile	0.082 (0.116)	-0.067 (0.102)	0.108 (0.047)**
4th income quintile	0.095 (0.121)	-0.046 (0.105)	0.117 (0.052)**
5th income quintile	0.191 (0.136)	-0.090 (0.132)	0.193 (0.065)***
Community violence Dummy	0.102 (0.127)		0.199 (0.056)***
Rural area in 93	0.609 (0.266)**	0.356 (0.472)	0.003 (0.178)
Tarred road Dummy	0.189 (0.246)		-0.211 (0.176)
Old KwaZulu area	-0.635 (0.238)***	0.029 (0.608)	-0.547 (0.255)**
HH owns its house	0.002 (0.153)		-0.073 (0.064)
Obs.	8258	7654	8258

Notes: *, **, *** denote significance at 10%, 5% and 1% levels. Marginal Affects Reported. The omitted category is a single, female non-resident, non-labour force participant, 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.

References

- 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.
- Becker, Gary S.**, “Investment in Human Capital: A Theoretical Analysis,” *The Journal of Political Economy*, 1962, 70 (5), 9–49.
- 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.
- Boskin, Michael J.**, “A Conditional Logit Model of Occupational Choice,” *The Journal of Political Economy*, 1974.
- 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.
- Cameron, A. Colin and Pravin K. Trivedi**, *Microeconometrics: Methods and Applications*, Cambridge: Cambridge University Press, 2005.
- Card, David**, “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.
- 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, 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.

- Deaton, Angus**, *The Analysis of Household Surveys – A Microeconomic Approach to Development Policy*, Baltimore: Johns Hopkins University Press, 1997.
- 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.
- 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.
- Fleisher, Belton**, *Labor Economics: Theory and Evidence*, Prentice-Hall, 1970.
- Griliches, Zvi 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.
- 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.
- ILO**, *Employment, incomes and equality: A strategy for increasing productive employment in Kenya*, Geneva: International Labour Office, 1972.
- International Monetary Fund**, “International Financial Statistics September 2009, Series 199..AF.ZF,” 2009.
- 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.
- 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.
- 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**, *Schooling, Experience and Earnings*, Columbia University Press, 1973.
- Moll, Peter**, "Compulsory Centralization of Collective Bargaining in South Africa," *American Economic Review*, 1996, 86, 326–329.
- Muller, Colette**, "Measuring South Africa's Informal Sector: An Analysis of National Household Surveys," 2003. DPRU Working Paper.
- Posel, Dori and Daniela Casale**, "Internal migration and household poverty in post-apartheid South Africa," in Ravi Kanbur and Haroon Borat, eds., *Poverty and Policy in Post-Apartheid South Africa*, HSRC Press 2006.
- Psacharopoulos, George**, "Returns to investment in education: A global update," *World Development*, September 1994, 22 (9), 1325–1343.
- Roy, A.**, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, 3 (2), 135–146.
- 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.
- Wooldridge, Jeffrey M.**, *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA.: MIT Press, 2002.