Do informal workers suffer an earnings penalty? A panel data analysis of earnings gaps in South Africa

KATRINA YU
Supervisors:
Dr. Jean-Pierre Cling (Université Paris 13)
Dr. Nicolas Pons-Vignon (University of Witwatersrand)

Erasmus Mundus Masters Program in Economic Policies in the Age of Globalisation (EPOG)

15 June 2015

This thesis is submitted in partial fulfilment of the requirements for the degrees of Masters of Commerce (Development Theory and Policy) at the University of Witwatersrand and Masters of Economics (Economics and International Finance) at Université Paris 13.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Jean-Pierre Cling and Nicolas Pons-Vignon, for their insightful guidance and advice over the past six months — this thesis would not be possible without your encouraging support. I would also like to express sincere gratitude to Mr. François Roubaud from Développement, institutions et mondialisation (DIAL) for taking the time to share his invaluable expertise and constructive comments with me.

I am also thankful to my sister, Serena, for her patient, technical guidance and a lifetime of unwavering support, and to Blake for his editing of my thesis and supply of endless laughter when it was needed most. Finally I am grateful to all my friends, both near and far, who have kept me sane throughout this whole process.

PLAGIARISM DECLARATION

I, Katrina Yu, declare that this research report is my own, unaided work.

This work is submitted in partial fulfilment of the requirements for the Master of Commerce in Development Theory and Policy degree at the University of the Witwatersrand, Johannesburg, and for the Master of Economics (Economics and International Finance) at Université Paris 13. It has not, either in whole or in part, been submitted before for any degree or examination in this, or any other University.

Katrina Yu
15 June 2015

ABSTRACT

Drawing on data from the National Income Dynamics Survey (NIDS) from 2008-2012, this paper investigates informal-formal earnings gaps in South Africa in order to assess the validity of labour market segmentation theories and to shed light on the motivations behind informal labour. In addition to controlling for observed individual characteristics using pooled OLS regression, fixed effects and quantile regressions are also estimated to control for time-invariant individual heterogeneity and to assess variations in earnings gap along the income distribution. Results are indicative of segmentation both within informal employment and in the labour market as a whole. Overall, informal employment in South Africa can be characterised as highly diverse and heterogeneous, with earnings differentials varying by employment type (whether salaried or self-employed), gender, and level of income. This earnings analysis is complemented by a brief exploration into the non-pecuniary characteristics of informal employment, with a consideration of the relationship between subjective wellbeing and informality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Given its prevalence in much of the developing world, informal activity lies at the heart of development policy and has drawn considerable research interest over the past four decades. However, since the term 'informal sector' was coined by Hart (1973), there has been no single definition of informality that has consistently underpinned the theoretical or empirical literature, such that even stylised facts about informality remain contentious (Badaoui et al., 2008). On top of this, a plurality of views on the origins and the causes of informal labour have fuelled the continuing theoretical and empirical debate. More recent discourse has shifted away from a dualistic representation of the labour market — in which informal and formal sectors are considered distinct entities and informal labour considered an involuntary outcome — towards a more heterogeneous and multi-segmented characterisation in which some segments of informal employment may reflect a voluntary choice given a worker's particular set of preferences, skills, and alternate earnings prospects.

While there are no standard tests of theoretical approaches to informality, this paper follows a major strand of research in empirically estimating informal-formal earnings gaps in order to shed light on the degree of segmentation both within informal employment in South Africa and in the labour market as a whole. As is further detailed in Chapter 3, the existence of an earnings penalty for an informal worker vis-à-vis an otherwise equally productive formal worker is evidence in support of theories of labour market segmentation[1] and offers a valuable insight into the motivations of informal workers. Specifically, the question this thesis seeks to answer is: "*Taking into account individual and job characteristics, do informal workers in South Africa suffer an earnings penalty?".* In answering this question, variations in the earnings gap will also be identified, according to employment segment (whether a wage employee or a self-employed worker), gender and level of income. These findings will considerably broaden the existing evidence base on the nature of the informal economy in South Africa, which in a country plagued by high levels of inequality, absolute poverty and unemployment, can play a significant role in poverty alleviation by providing potential avenues for income generation.

In this sense, an enriched understanding of informal labour is critical for researchers and South Africa's policymakers alike. This study aims to contribute to the stock of knowledge in several ways. Firstly, unlike previous empirical studies that have relied solely on Statistics South Africa's (SSA) Labour Force Survey (LFS), this paper draws on the National Income Dynamics Study (NIDS) — currently the only true national, panel data survey. The ability to exploit the panel nature of the NIDS will allow for the control of time-invariant individual heterogeneity, an element missing in much of the empirical literature. Secondly, the analysis

[1] Following Posel and Heintz (2008), this paper defines labour market segmentation as "the existence of barriers to mobility within labour markets that prevent employed individuals from maximising the returns to their labour by switching to a more highly remunerated type of employment".

[2] According to the latest QLFS in March quarter 2015, the official unemployment rate was 26.4%, while it was 50.3% for those aged from 15-24 (Stats SA, 2015). However, using the expanded definition that includes

covers a full sample of workers, including women and the self-employed, unlike many previous studies that focus only on male workers, or salaried workers. This not only provides a more complete analysis of informal employment, but also permits analysis from a gender perspective, which has not been considered previously in the South African context. Thirdly, quantile regression analysis is conducted in order to uncover variations of the earnings gap along the income distribution. The existing literature has shown this to be a valuable exercise and to the author's knowledge, Bargain and Kwenda (2011) is the only existing study to conduct this analysis in the South African context — yet their analysis is restricted to a sample of men employed full-time in the private sector. Lastly, the richness of the NIDS dataset also allows the characteristics of informal workers to be explored beyond monetary measures. In addition to earnings data, the NIDS also includes a number of health and subjective wellbeing indicators, such as level of life satisfaction. These indicators are briefly explored as a complement to the earnings analysis.

The paper will proceed as follows. Chapter II provides a brief outline of the key features of informal employment in the South African labour market. Chapter III explores the various theoretical approaches to informality that have emerged in the literature, before an in-depth literature review of previous empirical studies is conducted in Chapter IV. Chapter V follows with a discussion of the data, measurement and definition of informality and key descriptive statistics. Chapter VI outlines the empirical methodology, before the results are presented in Chapter VII. Further discussion of results is provided in Chapter VIII, before Chapter IX concludes.

II. CONTEXT: INFORMAL LABOUR IN SOUTH AFRICA

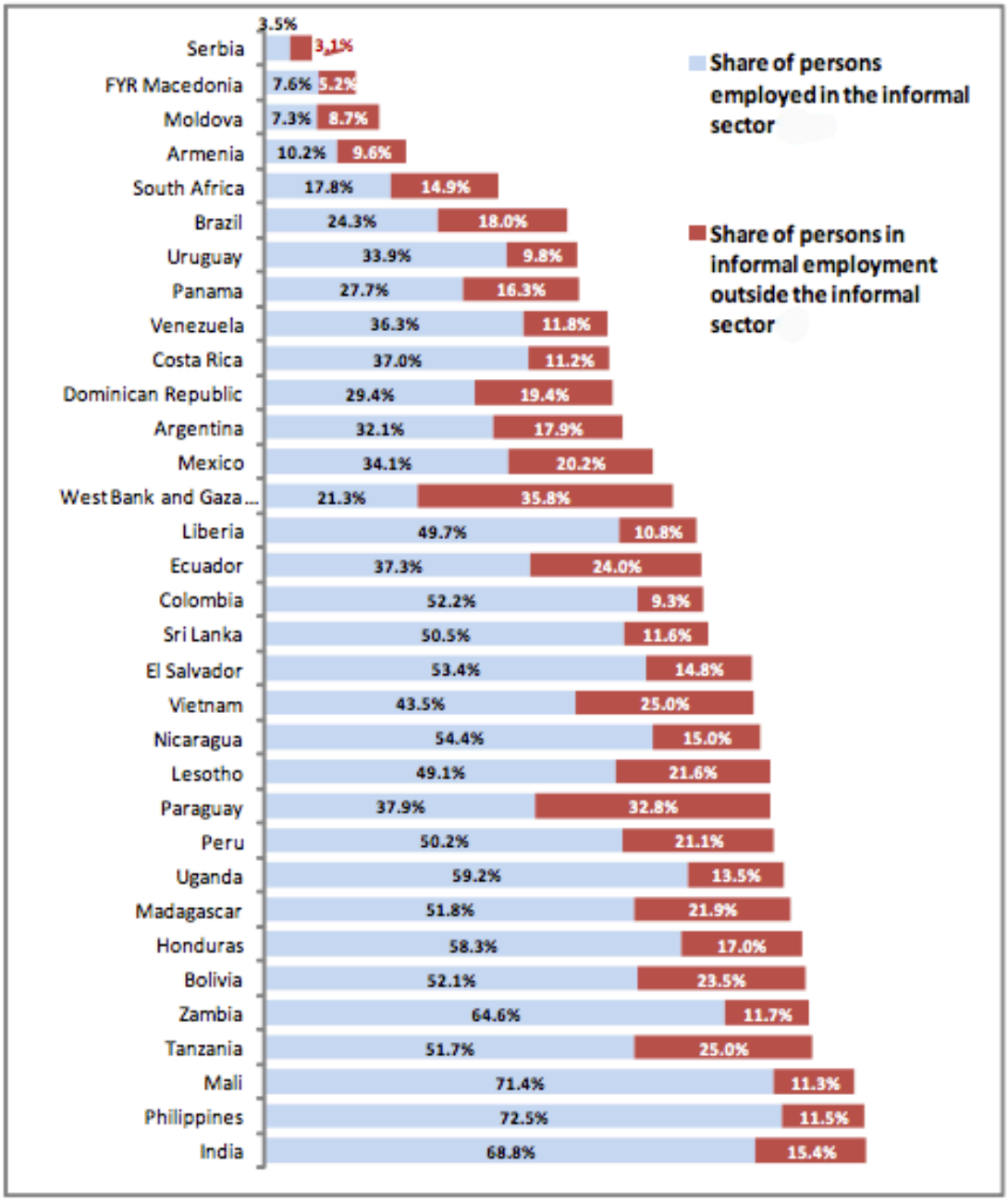
Similar to many other low- and middle-income countries, employment in the formal sector in South Africa is "highly regulated, highly remunerated and scarce" (Magruder, 2010), with formal wage employment representing the most lucrative form of work available (Kingdon and Knight, 2004, 2007a; Ruffer and Knight, 2007). However, it is the co-existence of a high rate of open unemployment and relatively limited informal employment that is perhaps the most salient feature of South Africa's labour market and one of the country's most difficult economic challenges (Kingdon and Knight, 2007a; Posel and Heintz, 2008). In this respect, South Africa is considered an international outlier (see Figure 1 for a global comparison of informality rates). As Kingdon and Knight (2007a) point out, other Sub-Saharan African countries and Latin American and Asian countries with similar levels of per capita income exhibit much higher rates of informal employment, which in turn help to offset higher rates of unemployment. South Africa, on the other hand, consistently records one of the highest unemployment rates in the world, particularly amongst the youth.[2]

While there is an extensive body of literature dedicated to uncovering the factors driving South Africa's low rates of informality, it is still an imperfectly understood phenomenon. Evidence suggests that informal work is preferable to unemployment both in terms of income

[2] According to the latest QLFS in March quarter 2015, the official unemployment rate was 26.4%, while it was 50.3% for those aged from 15-24 (Stats SA, 2015). However, using the expanded definition that includes discouraged workers, the figures rise to 36.1% and 63.1%.

and subjective wellbeing (Kingdon and Knight, 2007a), such that the relatively small size of South Africa's informal sector is frequently attributed to the existence of barriers to entry. However, there does not appear to be a consensus on which specific factors constitute the

Figure 1: Share of persons employed in informal economy, (non-agricultural)



Source: ILO (2012)

dominant barriers to informal employment. In a study of 500 informal firms in Johannesburg[3], Chandra et al. (2002) identify high crime rates, lack of access to credit, lack of access to infrastructure and services, and insufficient training as the four greatest constraints on business growth. The authors suggest that government support in relation to these constraints is inadequate, corroborating more general arguments that state support for small and medium micro-enterprises (SMMEs) is overwhelmingly in favour of formal

[3] Cichello (2005) and Skinner (2005) pinpoint similar constraints in the Cape Town and Durban areas.

enterprises (Xaba et al., 2002). This is in line with the notion of the "second economy"[4], which has dominated government policy discourse since 2003 when, in an address to the National Council of Provinces, the then President Mbeki stated that:

"The second economy is characterised by underdevelopment, contributes little to GDP, contains a big percentage of our population, incorporates the poorest of our rural and urban poor, is structurally disconnected from both the first and the global economy and is incapable of self generated growth and development." (Mbeki, 2003)

The implication of this conceptualisation of the economy is that the poor remain poor because they are trapped in a "structurally disconnected" realm, unable to become part of the "first economy". This notion of poverty quickly pervaded political, journalistic and academic spheres, with most government departments soon adopting policies addressed at shrinking the "second economy" (Valodia and Devey, 2010). However such dualist representations of the labour market have drawn heavy criticism from no shortage of researchers (Devey et al., 2005; Du Toit and Neves, 2007; Phillip, 2009), who argue that the notion of a "second economy" does not accurately capture the dynamics of economic marginalisation and that the economy should instead be conceptualised as a "single internally differentiated and segmented" entity (Du Toit and Neves, 2007). An in depth discussion of dualist theories of labour market segmentation is presented in Chapter III.

While barriers such as a lack of access to credit may play a role in restricting informal activity, Kingdon and Knight (2007a) note that such factors are prevalent in most developing countries and cannot account for the particularly low rates of informal employment in South Africa. There are, however, two key distinctive features that are unique to the South African context. The first is the effectiveness of labour market institutions, particularly the enforcement of labour regulations (*ibid.*). In South Africa, most sectoral minimum wages and employment conditions are prescribed by Bargaining Councils and Wage Boards, which apply across the country regardless of firm size and attract stiff penalties in the event of their breach. Informal enterprises are not exempt from punitive measures[5]. This mechanism can impose a prohibitive burden of high labour costs on small firms, such that it may hinder their entry and growth. Moreover, minimum wage legislation in South Africa has an explicit provision stipulating that wages take into account the wage gap that resulted from the Apartheid period (Arora and Ricci, 2005), which can further lift wages beyond the market-clearing equilibrium wage and heighten the burden of labour costs.

However, while such regulatory mechanisms may impact on small firms, it does not affect individual owner-operated enterprises (Kingdon et al., 2005). This is more likely attributable to a second distinctive feature of the South African economy — the legacy of Apartheid

[4] The second economy is not entirely synonymous with the informal economy as it also includes "the involuntarily and marginalised unemployed and some of those those who are not economically active" (Du Toit and Neves, 2007)

[5] "Informal sector firms that are not party to a bargaining council but which operate within sectors regulated by bargaining councils are incorporated into the Labour Relations Act…[and] the employment conditions of employees in such firms would be regulated by the Basic Conditions of Employment."

policies which, despite being abandoned in 1994, have continued to shape economic and social outcomes (Posel and Heintz, 2008). These policies not only resulted in a lowly educated African population and spatial inequalities, but also suppressed a culture of entrepreneurship through restrictive legislation such as harsh licensing, restrictive bye-laws and strict zoning regulations. For example, the Black Consolidation Act (1945) restricted African urban businesses as a means to discourage rural-urban migration and protect white business from competition, while the Group Areas Act (1950) restricted non-white entrepreneurs from operating in vital areas of the economy (Standing et al.,1996). Importantly, the repression and disempowerment of the African population also stymied the social and entrepreneurial skills that are critical for the success of informal enterprises.

Before moving to a more detailed analysis of South African informal employment, a review of the key theoretical debates and relevant empirical literature will be presented.

III. CONCEPTS AND THEORIES OF INFORMAL LABOUR

a.) What is 'informal'?

Since the term 'informal sector' was coined by Hart (1973), the concept of informality has continually evolved and its appropriate definition and measurement remains an area of strong contestation among researchers. In earlier literature, informality was interpreted using an enterprise approach adopted at the International Labour Organisation's (ILO) 15th International Conference of Labour Statisticians (ICLS) in 1993, whereby informal employment was defined as employment in firms within the informal sector. Specifically, it encompassed "*all individuals who work in small unregistered enterprises, both employers and employees, as well as self-employed persons who work in their own or family businesses"* (ILO, 1993). This was the approach initially adopted by SSA in its regular household surveys in the post-Apartheid period (Neves et al., 2011).

More recent discourse, however, has seen a notable shift away from the idea of an informal *sector* and towards the idea of an informal *economy*, reflecting not only the interdependency and linkages between the formal and informal sectors, but also the heterogeneity of informality[6]. The notion of informality has extended beyond the nature of the enterprise to encompass characteristics of the employment relationship itself. More specifically, the informal economy incorporates both (i) employment in the informal sector and (ii) employment in the formal sector that displays informal characteristics — that is, unprotected or unregulated jobs that lie outside the regulatory framework and are not covered by basic legal and social protection, taxes or employment benefits. As Neves et al. (2011) argues, this shift was necessary to reflect the increasing informality driven by the spread of neoliberalism and globalisation, in the form of "casualisation, outsourcing and industrial outworking within formal enterprises and markets". To reflect this development in theoretical understanding, the employment-based approach was incorporated into an expanded ILO definition at the 17th ICLS in 2005 and was adopted by SSA in 2008 (Yu, 2010). Nevertheless, there remains

[6] However, the idea that entry into informality occurs for various reasons is not new. Hart (1973) never saw the informal sector as intrinsically bad, and Fields (1990) noted that there is an "upper tier" to informal employment that does very well.

inconsistencies regarding precisely which 'informal' characteristics should be included in this expanded definition. This thesis will return to the various approaches used by empirical researchers to define informality in the following chapter.

b.) Theoretical approaches: dualist, structuralist and legalist

There are three main schools of thought on informal labour that have emerged in the literature: the dualist school, the structuralist school, and the legalist school. The core point of distinction between the approaches is the degree to which informality is assumed to be a voluntary choice and the degree to which it is imposed as a result of limited opportunities and constraints in the formal labour market (Bacchetta et al., 2009).

The dualist school, popularised by the ILO during the 1970s, has its theoretical roots in the works of Lewis (1954) and Harris and Todaro (1970) and reflects a dichotomous view in which the informal sector is viewed as the 'residual' of the formal sector — that is, a separate, marginal sector with no direct linkages to the formal sector. Proponents of the dualist view typically characterise informal labour as labour-intensive, survivalist activities, with low wages and low productivity, in firms that are small-scale and undercapitalised (Godfrey, 2011). Under the assumption of zero marginal productivity and an unlimited supply of labour in the informal sector, informal employment is seen as the result of insufficient formal sector employment opportunities to absorb surplus labour and is therefore an involuntary outcome (Chen, 2014). As the economy develops, the informal sector self-eradicates as the formal sector expands. In other words, informality is seen as a sign of economic under-development.

A key proposition of the dualist view is the existence of an informal earnings penalty due to rigidities in the formal labour market, such as minimum wage legislation, that raise wages above a market-clearing level and leave some workers rationed out of formal employment opportunities. These individuals, seeking to avoid involuntary unemployment, enter the informal labour market where there are no institutional arrangements distorting wages, no barriers to entry, and an abundant supply of unskilled workers. The result is lower wages for the informally employed, and a labour market characterised by a dualism in earnings such that workers with similar characteristics can attract different wages depending on the sector they are employed in, rather than their individual skills and characteristics. In the absence of segmentation, competitive forces should equalise wages for identical workers across different segments of the labour market, with differences in earnings solely reflecting compensating differentials, such as returns required for investment in education or to compensate for the riskiness of a job. This is the rationale behind the examination of earnings gaps to uncover evidence for labour market segmentation.

Despite its pervasiveness in South African policy discourse the dualist view has increasingly been regarded as too simplistic, with assumptions that are inconsistent with reality. Labour markets rarely exhibit a distinct divide between the informal and the formal and are more akin to a continuum where aspects of both informality and formality can characterise each employment relationship. Moreover, viewing informal employment as simply a residual category fails to recognise that the informal and formal sectors are intrinsically linked.

These intrinsic linkages are at the core of the structuralist school of thought. Proponents of this approach argue that informal labour is a consequence of capitalist development, where small firms and unregistered workers in the informal economy are subordinate to, and exploited by, large capitalist firms in the formal economy who are seeking to reduce costs (Bacchetta et. al., 2009). For example, informal labour may be integrated into global production systems as a source of low cost labour, or social protections for workers may be reduced in reaction to stronger global competitive pressures, driving the increasing informalisation of labour. Such dynamics perpetuate the unequal and exploitative relationships between the two sectors and account for the persistence of informal labour. In contrast to the dualist view, it is the demand for informal labour rather than the lack of demand in the formal market that drives informality.

The third approach, the legalist school, gained popularity during the 1980s and 1990s with the rise of neoliberalism and sees the informal economy as a "rational response by micro-entrepreneurs to over-regulation by government bureaucracies" (de Soto, 1989). Entrepreneurs seek to reduce costs by avoiding costly government regulation both to formalise, such as business registration requirements, as well as to remain formalised, with costs imposed directly through taxes, compliance with bureaucratic requirements, or as a result of an instable legal system or insecure property rights (*ibid.*). Thus, according to the legalist paradigm, entrepreneurial activity is stifled by formal institutional arrangements and informality is a means for microenterprises to prosper. In contrast to the previous two approaches in which informality is characterised as an imposed constraint, informality in this view is a voluntary outcome.

These divergent approaches to the origins and causes of informality suggest a highly heterogeneous informal economy and an existence of labour market segmentation not only between informal and formal activities, but also *within* the informal economy. Indeed, there have been efforts to synthesise the approaches by recognising the internal duality of informal employment, such that there exists both an "upper tier" of workers who voluntarily choose informality over formal employment, as well as a "lower tier" of workers who see informality as a last resort (Fields, 1990 and 2005; Perry et al., 2007). Whether one tier dominates the other is ultimately an empirical question and, as mentioned previously, one of the key approaches to addressing this question is the estimation of informal-formal earnings gaps. This method is grounded in the revealed preferences principle and the crucial assumption that income is a reasonable proxy for individual utility. All else being equal, if an informal worker attracts an earnings premium vis-à-vis their formal counterpart, informality can be interpreted as a voluntary choice. A review of the empirical literature adopting this approach will now be presented.

IV. LITERATURE REVIEW

There is an extensive body of empirical literature investigating the existence of an informal earnings penalty in developing countries, with Latin American economies attracting a dominant share of research interest. However, the evidence drawn from these studies remains inconclusive. While numerous researchers appear to support the dualist theories of labour

market segmentation by confirming the existence of an informal penalty (Gindling, 1991; Bucheli and Ceni, 2010; Krstic and Sanfey, 2011; Rand and Torm, 2012; Zuo, 2013; Harati, 2013), many find contrary evidence of an informal *premium* (Marcouiller et al., 1997; Maloney, 1999; Arabsheibani and Staneva, 2014, Huber and Rahimov, 2014). Not only is the evidence mixed between countries, it also suggests that earnings differentials can differ according to factors such as gender, education level, or position in the income distribution. The overarching message that can be drawn from the literature, therefore, is that the informal labour market is highly heterogeneous and composed of multiple segments. The observation of an informal-formal earnings gap not only relies on the segment of employment under analysis, but also the precise definition of informality adopted. These issues will now be examined in turn.

a.) Definition and measurement of informality

As mentioned in the previous chapter, there remains contestation amongst researchers on the precise definition and measurement of informality. These divergences are compounded by data limitations, which can hamper the operationalisation of the ILO's recommendations and prevent the widespread uniform use of the term 'informal' in empirical analyses.

Amongst those studies that rely solely on the enterprise-based approach, some have defined the informal sector using firm size as a criteria (Maloney, 1999; Pradhan and Van Soest, 1995), while others use the firm's registration status (Unni, 2005). More recent empirical studies also tend to incorporate the employment-based approach into their definition, with a broad array of characteristics used to identify informality, such as a lack of social security protection in its various forms (Arias and Khamis, 2008; Bucheli and Ceni, 2010; Krstic and Sanfey, 2011; Garcia, 2014), a lack of labour regulation (Pratap and Quintin, 2006), an absence of a written contract, (Arabsheibani and Staneva, 2012; Harati, 2013), an absence of a signed labour card (Tannuri-Pianto and Pianto, 2002; Carneiro and Henley, 2001) or a lack of union membership (Saavedra and Chong, 1999). As Yu (2010) indicates, there are also a few recent studies that adopt an alternative productive approach, using the professional status of workers (occupation or educational attainment) to distinguish informal employment.

Other issues of contention regarding the definition of informality are the treatment of agricultural workers, domestic workers, unpaid family workers and the self-employed. The treatment of the self-employed is particularly crucial to the empirical results as they, along with salaried workers, form the two broad groups of informal employment (Perry et al., 2007). While some empirical studies classify all self-employed workers as informal (Marcouiller et al, 1999; Pradhan and Van Soest, 1995; Guifu and Shigeyuki, 2009), others distinguish between the formal and informal self-employed[7] (Nguyen et al., 2013; Tansel and Kan, 2012), while several omit the self-employed from analysis altogether (Tannuri-Pianto and Pianto, 2002; Pratap and Quintin, 2005; Bucheli and Ceni, 2010). The omission of the self-employed is often justified on the grounds of the unreliability of self-employed earnings,

[7] However, as Nguyen et al. (2013) notes, the failure to segregate the formal self-employed may largely be due to their relatively small number in many developing and emerging economies.

as it is inherently difficult to disentangle returns to capital or to unpaid family work, from returns to labour (Falco et al., 2010; Pratap and Quintin, 2005; Krstic and Sanfey, 2011).

Some researchers, such as Unni (2005), segment the informal self-employed even further by distinguishing those with and without capital, arguing for the need to separate those who operate with a view to maximise profits, and those who view self-employment as a residual, survivalist activity. The findings show that those with capital receive a higher return to their labour than those without capital. This highlights the multiplicity of motivations that can exist even within segments of informal employment. In a similar vein, Posel and Heintz (2010) find variations in the earnings gap between those informal self-employed with and without their own accounts in South Africa (this study is discussed in detail below).

Focusing on definitional issues in South Africa, Yu (2010) reviews a number of different approaches used by both SSA and researchers to measure the size of informal employment. The approaches differ in a number of ways, such as the treatment of agriculture workers or unpaid workers in private household businesses, or the characteristics used to identify an informal employment relationship. It shows that different measurements of informality indeed have important implications for understanding labour market dynamics, underscored by the finding that only 6.9% of workers are defined as informal under all approaches reviewed.

b.) Segmentation within the informal economy: international evidence

As mentioned above, the literature suggests that the empirical ambiguity on the existence and nature of the informal-formal earnings gap stems not only from incongruities in defining informality, but also from the segment of informality under analysis. Most prominently, there appears to be a demarcation of results between informal wage workers and the informal self-employed. As highlighted previously, several studies omit the self-employed from analysis, highlighting a major limitation of their results. Moreover, even when the self-employed are included, many studies subsequently fail to delineate between informal wage workers and the informal self-employed in their estimation of an earnings gap, treating informal employment as homogenous (Gindling, 1991; Pradhan and Van Soest, 1995; Marcouiller et al., 1997; Carneiro and Henley, 2001; Arabsheibani and Staneva, 2012).

The significant implications of this shortcoming is underscored by the number of studies who do examine wage workers and the self-employed separately, and report different findings for the earnings of formal workers vis-à-vis informal wage workers, and vis-à-vis the informal self-employed. For example, Nguyen et al. (2013) find that while overall there is an informal wage penalty (15%) in the Vietnamese labour market, this penalty is attributed solely to informal wage workers, with the informal self-employed attracting a premium vis-à-vis formal wage workers. There is, however, an informal earnings penalty when comparing the informal self-employed to their formal counterparts. Using similar methodology, albeit defining informality under different criteria[8], Nordman et al. (2012) find strikingly similar results for Madagascar, characterising the self-employed as the 'upper tier' of informal employment. This superiority of the self-employed segment is also largely supported by evidence from the

[8] Nordman et al. (2012) use an enterprise approach, while Nguyen et al. (2013) combine both job and enterprise approaches.

Latin American literature (Arias and Khamis (2008) in Argentina, Saavedra and Chong (1999) in Peru and Maloney (1999) in Mexico).

Indeed, it is the Latin American economies in which the evidence for the voluntary nature of informal employment is most compelling. Maloney (1999, 2004) suggests several potential explanations for this, including the flexibility or independence afforded by informal self-employment, or the desire to avoid the relative high cost of low quality social security services associated with formal sector work.[9] It is important to note, however, that while the characterisation of wage workers as the 'lower-tier' of informal employment and the self-employed as the 'upper-tier' is consistent with the majority of the literature, this conclusion cannot necessarily be generalised across all countries. Examining the Turkish labour market, Tansel and Kan (2012) show that it is the informal self-employed that occupy the lower tier of the informal economy, attracting an earnings penalty vis-à-vis informal salaried workers.

The findings in Tansel and Kan (2012) also draw attention to the potential gender variation in informal-formal earnings differentials. They find that the informal penalty for the self-employed is significantly worse for females (40%), and almost insignificant for men. However, when examining wage employment, the results are reversed and it is men who report an informal penalty (10%), while there is no significant earnings gap for women. Observing informal employment as a whole, they find that men suffer a significant informal earnings penalty, while the pay gap is statistically insignificant for women. In contrast, Nguyen et al. (2013) and Nordman et al. (2012) find that women suffer the relatively larger informal penalty in Vietnam and Madagascar respectively. Thus, while there is no consensus on the impact of gender on informal-formal earnings gaps, the available evidence does underscore the importance of including the female population in any empirical analysis. Several studies have excluded female workers, with the decision often justified on the grounds of reducing the complexities of modeling female participation decisions (Maloney 1999, Badaoui et al., 2008). Yet the inclusion of female workers is crucial as one of the key arguments in support of the view of voluntary informality is the advantage of flexible work arrangements, which undoubtedly has a gender dimension. For example, Marcouiller et al. (1997) find that for women, heading a household and having many children decreases the probability of formal sector employment. Thus, the exclusion of women from the sample would not only remove a significant proportion of informal workers, it also removes from analysis a key potential motivation for informal employment.

Earnings differentials can also vary along the income distribution and several studies have investigated this potential structural heterogeneity within sectors by using quantile regression to analyse distributional patterns of the earnings gap at various points on the pay ladder. For example, Nguyen et al. (2013) show that the penalty for informal employment as a whole, once accounting for unobserved individual characteristics, decreases along the income distribution — ranging from 19% for the bottom quantile to 10% for the uppermost quantile. This finding that earnings differentials are smaller among high income groups than among low income groups is corroborated by a number of other studies across several countries

[9] In addition, there may be informal networks that may be able to substitute for formal, government-run social security systems at a lower cost, and that are more trustworthy (Maloney, 2004).

(Tannuri-Pianto and Pianto (2002) for Brazil; Tansel and Kan (2012) for Turkey; Garcia (2014) for Colombia). Studies such as Garcia (2014) further investigate this distributional pattern using decompositional techniques and find that low income informal workers in Colombia earn less not only because they are less skilled, but also because they receive lower returns to these skills, while informal workers at the upper end of the distribution earn less because their formal counterparts are more highly skilled. This is evidence of a segmented informal labour market, in which voluntary and involuntary informal employment can co-exist by choice *and* as a result of labour market segmentation, depending on the worker's relative position in the income distribution.

Individuals can also differ in a number of unobserved characteristics that may be important determinants of their earnings potential, such as their innate ability, entrepreneurial skill, effort or preferences. While the widespread unavailability of panel data limits the number of empirical studies being able to control for unobserved heterogeneity, the few studies that have been able to exploit the availability of panel surveys have shown that unobserved characteristics account for a large portion of the observed wage gap, or in some cases, even erase or reverse the sign of the differential (Pratap and Quintin, 2006; Bargain and Kwenda, 2011; Tansel and Kan, 2012; Nordman et al. 2012; Nguyen et al. 2013). This underscores one of the key advantages of the NIDS dataset— the ability to exploit its panel structure and observe the extent to which earning gaps can be attributed to individual fixed effects.

Overall, the international empirical literature on informal-formal earnings gaps affirms the characterisation of a highly diverse and multi-segmented informal economy, suggesting a degree of validity to each of the theoretical approaches to informality. It underscores the importance of accurately defining informality and the need to separately examine different segments within the informal economy. It also provides ample motivation for conducting distributional analysis of earnings gaps using quantile regression, as well as taking a gender perspective in the analysis.

c.) South African empirical evidence

There are four recent empirical studies that have sought to identify and measure an informal-formal earnings gap in South Africa: Posel and Heintz (2008), Badaoui et al. (2008), Bargain and Kwenda (2011) and Kerr and Teal (2012).

Posel and Heintz (2008) investigate patterns of labour market segmentation in order to seek explanations for the puzzling co-existence of a low rate of informal employment and high rate of open unemployment. Arguing that dualist formulations of labour market segmentation are too narrow, the authors analyse earnings differentials across 11 categories of employment. Persistent earnings differentials across different categories, having controlled for other factors, would support a theory of labour market segmentation, albeit a broader conception than under the dualist school.

By examining cross sectional data from the September 2004 LFS and estimating standard Mincerian earnings equation for each employment category, Posel and Heintz (2008) indeed find evidence of large wage differentials between the formal and informal sectors and within

the informal sector itself. The results thus support arguments of labour market segmentation, whilst also providing evidence that the informal sector itself is multi-segmented, with some types of informal employment facing higher barriers to entry and average wages than others. Specifically, it is found that, amongst informal workers, the non-agricultural informal self-employed (but who are not own-account workers) have the highest predicted earnings, followed by informal wage workers (both private and public), the self-employed own-account workers, and finally the agricultural wage and self-employed workers. A key shortcoming of the study, however, is that the cross-sectional structure of their data does not allow control for unobserved heterogeneity, which may explain some of the observed wage gap. In contrast, this limitation is addressed in Badaoui et al. (2008).

Badaoui et al. (2008) also use LFS data and employ a similar definition of informality as Posel and Heintz (2008), combining both enterprise and employment-based approaches. However, by exploiting the rotating panel nature of the survey and matching individuals within households across five waves (2001 to 2003), Badaoui et al. (2008) are able to create a panel of individuals. Their findings show that while almost 75% of the informal wage penalty can be attributed to observable human capital and job characteristics, there is still a 37% differential once these factors are taken into account. Controlling for time-invariant unobserved heterogeneity further reduces the penalty to around 18% and once formal sector income tax payments are taken into account, the informal-formal wage gap disappears entirely. Badaoui et al. (2008) thus arrive at a different conclusion to Posel and Heintz (2008). Badaoui et al. (2008) argue that one potential explanation for the lack of evidence of a significant earnings differential is that there are social and economic barriers to entry into informal employment, such that the least productive workers remain unemployed. However, it should be noted that both studies have their limitations. While Posel and Heintz (2008) do not account for tax payments[10] nor controls for time-invariant unobserved heterogeneity as in Badaoui et al. (2008), the latter's focus on a subsample of male salaried workers in the private sector is a major shortcoming as it does not provide a picture of the full spectrum of informal activities.

In addition, the findings of Posel and Heintz (2008) and Badaoui et al. (2008) may both be misleading as earnings comparisons at the mean can conceal variations along the income distribution. In contrast, Bargain and Kwenda (2011) estimate fixed effects quantile regressions in order to analyse distributional patterns in the informal-formal earnings gaps in Brazil, Mexico and South Africa. Not only does the study conduct a distributional analysis, it also controls for time-invariant unobserved characteristics using repeated samples of cross-sectional data while being one of the few cross-country empirical studies of informal-formal earnings differentials.

For South Africa, Bargain and Kwenda (2011) examine LFS data from September 2001 to March 2007 and also restrict their analysis to a subsample of urban men working in full-time employment in the private sector. However, unlike Badaoui et al. (2008), Bargain and Kwenda (2011) include the self-employed in their analysis. They find a particularly large informal penalty for the self-employed in South Africa (15%), compared to an earnings

[10] Adjustment of gross wages to take into account taxes paid in the formal sector is rarely considered in the literature.

premium in Mexico. Wage workers, however, are systematically underpaid compared to their formal counterparts in all countries and at almost all conditional quantiles. Yet again, this penalty is substantially larger in South Africa than either Brazil or Mexico, with informal wage workers penalised an average of around 21%. Further distributional analysis shows that while the informal penalty for wage workers is decreasing with income, it occurs at all points along the earnings distribution. Conversely, while the penalty for the informal self-employed workers is also decreasing with earnings, it changes into a moderate premium in the highest quarter of the distribution.

Across the three countries, Bargain and Kwenda (2011) find that the lower tier of informal employment is consistently comprised of wage workers, as well as the young and unskilled self-employed with their own accounts. While the self-employed without their own accounts are found to occupy the upper tier in all three countries, this segment is only marginal in South Africa. In addition to constraints to the entry and growth of informal enterprises discussed in Chapter II, it has been suggested that the relatively small and unproductive South African self-employed sector may be attributable to a lesser incentive to register successful enterprises in South Africa than in Mexico or Brazil (Henley et al., 2009).

In contrast to the previous three papers, Kerr and Teal (2012) do not use LFS data, but draw on the KwaZulu-Natal Income Dynamics Study (KIDS) panel dataset[11] to investigate earnings differentials across different types of employment. In particular, it looks at the roles of human capital and two labour market institutions— trade unions and public sector pay policy — as determinants of earnings. Due to data limitations which do not allow the authors to classify whether an individual is working for a registered enterprise or not, the paper identifies employment categories according to a spectrum of formality, whereby unionisation and public sector work are considered characteristics of highly formal employment, regular work in the private sector less so, and self-employed and casual employment even less formal. It finds evidence of a substantial public sector premium, which cannot be explained by human capital or unobserved characteristics. On the contrary, the private sector union premium is largely explained by these factors, with the authors noting in particular that unionised workers appear to be of unobservably higher quality. While such results should be treated with caution as the analysis is restricted not only to a single province, but also to the African population[12], the findings corroborate the broader empirical literature in rejecting a strictly dichotomous view of the labour market.

d.) Qualitative empirical studies

Relying on quantitative data alone can conceal important labour market dynamics that may be uncovered with qualitative research. Such studies are a useful complement to studies of informal-formal earnings differentials as it provides invaluable insights into the non-income dimensions of informal activity. As pointed out by several authors, including Posel and Heintz (2008), an important caveat regarding studies on earnings gaps is that evidence of earnings differentials may not necessarily be attributable to segmentation. An alternate possibility is that there are non-pecuniary benefits to certain types of employment that may

[11] KIDS is a three wave panel conducted in 1993, 1998 and 2004 in KwaZulu-Natal province.
[12] KIDS data is restricted to African and Indian households and the authors choose to abstract from racial differences in earnings determination by removing Indian households from analysis.

compensate for wage differentials, such as the flexibility or independence associated with self-employment, as mentioned above. While such benefits are promulgated by several authors, and in particular by Maloney (1999, 2004), it is the developed country literature that has tended to emphasise the desirability of independence and flexibility (Falco et al. 2013). Evidence of a positive impact of self-employment on job satisfaction tends to be in western, developed countries, while evidence from developing and transition economies are more mixed (Markussen et al. 2014). Further, as noted by Kucera and Roncolato (2008), the Latin American countries from which the bulk of the evidence for voluntary informal employment is derived, are among the most developed of developing countries. On the contrary, there is a large body of evidence highlighting the precariousness of informal work in developing countries, showing that it is often associated with poorer employment conditions, such as insecurity of employment or occupational health hazards (Chang, 2009; Barrientos and Barrientos, 2002; Unni and Rani, 2002; Du Toit and Neves, 2007). It is thus important to consider the non-pecuniary aspects of informal employment and the possibility that earnings gaps may be driven by differences in intrinsic non-pecuniary utility, rather than labour market segmentation. To this end, subjective measures of job and more broadly life satisfaction may provide a good proxy for job quality.

While the empirical literature exploring this avenue remains limited, there are a few studies investigating the relationship between measures of subjective wellbeing and informal employment. Falco et al. (2012) look at determinants of subjective wellbeing in urban Ghana and find no evidence of a wellbeing premium for formal employment, vis-à-vis both informal self-employed and wage workers. They do, however, find that the self-employed who employ others report higher levels of satisfaction than formal wage workers. Razafindrakoto et al. (2012) also look at the Vietnamese labour market, analysing job satisfaction survey responses to reveal a labour market characterised by a strong hierarchical structure depending on institutional sector, with the public sector sitting at the pinnacle, followed by the private formal sector, then informal workers and finally those in agriculture. Their study thus corroborates the characterisation of informal employment as inferior to formal work, with informal workers reporting lower levels of job satisfaction (although it is seen as a means to escape agricultural work). As pointed out by De Vreyer and Roubaud (2013), however, a caveat with studies of job satisfaction is that it is difficult to establish a causal link between job satisfaction and employment outcomes because job satisfaction is likely to be correlated with aspirations. Highly educated workers, for example, are more likely to have higher aspirations and may report lower job satisfaction levels, despite holding better jobs. These studies are also plagued by the intrinsic methodological problem of relying on subjective responses. This constraint, coupled with the lack of quality data, means that it has been difficult to reach a consensus in the international literature on the claim that non-pecuniary aspects of the labour market can compensate for earnings differentials. Furthermore, to the author's knowledge, there exists no study that empirically models the impact of informality on job satisfaction, or broader measures of wellbeing, in the South African context.[13]

[13] Kingdon and Knight (2004) show that informal workers have a significantly higher perceived quality of life than the unemployed, thus supporting their hypothesis of the existence of barriers to informal employment. However, they do not compare the wellbeing of informal workers vis-à-vis formal workers.

Overall, this review has revealed several gaps in the literature worthy of further investigation. Firstly, there are significant shortcomings to each previous empirical study on South Africa. While Badaoui et al. (2008), Bargain and Kwenda (2011) and Kerr and Teal (2012) considerably restrict their analysis to a subsample of workers, Posel and Heintz (2008) use cross-sectional data and cannot control for individual fixed effects. Further, no consensus has emerged on the existence and nature of an informal-formal earnings gap. Secondly, with Bargain and Kwenda (2011) being the only study to investigate distributional patterns of the informal-formal earnings differential in South Africa to date, there is a dearth of empirical evidence on how earnings gaps can vary with income. Thirdly, previous studies have not investigated the gender dimension of earnings gaps in the South African context, beyond the inclusion of a gender dummy. Studies such as Nguyen et al. (2013) and Nordman et al. (2012) estimate regressions for separate subsamples of male and female workers in Vietnam and Madagascar respectively, allowing a more in-depth gender analysis to be conducted. Lastly, the NIDS dataset contains a wealth of information on non-pecuniary wellbeing variables, which can help shed light on a relatively unexplored dimension of informality in South Africa.

V. DATA AND DESCRIPTIVE STATISTICS

a.) Data description and defining informality

Data for this study is drawn from all three waves of the NIDS, conducted in 2008, 2010 and 2012, with the survey consisting of both household and individual level interviews. As South Africa's first nationwide panel data survey, NIDS' longitudinal character and wide coverage are two of its key advantages.

The NIDS' labour market module contains data on an individual's employment status, earnings and hours worked as well as other occupational information including occupation type, industry, type of contract, unionisation status, job history, and amount of extra benefits received, such as bonuses. While the LFS data is limited to a respondent's primary work, NIDS allows an individual to provide information on primary and secondary waged jobs, as well as any casual employment, self-employment, subsistence agriculture or help provided to others (whether family or not). Unlike the LFS, the NIDS does not collect information on the registration status of enterprises for wage workers (although it does for the self-employed). Thus, NIDS does not allow informal labour to be defined in line with the ILO definition of the informal *sector,* but rather in line with its broader definition of informal *employment*— that is, using an employment-based approach. Specifically, this thesis defines a wage worker as formal if they have a written contract or receive both pension and medical aid contributions from their employer. For self-employed workers, an enterprise approach is used, with respondents classified as formal if their business is registered for income tax and/or VAT.

One of the issues that frequently arise in defining informality is the treatment of agriculture work. The ILO definition of informal employment does not dictate either its inclusion or exclusion, allowing a degree of flexibility for its treatment (Wills, 2009). The NIDS dataset identifies agricultural work in two areas: a respondent can nominate "agriculture, hunting,

forestry or fishing" as their sector of employment [14] or they can identify any subsistence agricultural activities not previously nominated as wage or self-employment. This approach reflects the dualistic nature of agricultural employment in South Africa, comprising of a highly capitalised commercial sector and a subsistence sector largely concentrated in the former 'homeland' areas (May and Carter, 2009). However, this paper excludes all types of agricultural employment from regression analyses, in line with the existing literature that focuses on non-agricultural work. This is primarily due to a lack of quality earnings data, stemming particularly from the inherent difficulties in measuring the value of agricultural production for own consumption.

Another issue of contention for researchers is the treatment of domestic workers. While such workers are usually treated as informal, this paper will define the informality of domestic work on the basis of contract status and employer contributions, consistent with the employment-based approach outlined above. Domestic workers are identified in the NIDS dataset as those wage workers or self-employed who report their sector of employment as "in a private household".

b.) Descriptive statistics

Table 1 reports the distribution of employment by employment type from 2008 to 2012, for individuals aged 15-64 years. It focuses on wage employment, self-employment and casual employment and excludes those in subsistence agricultural or helping others in their businesses.[15] The size of informal employment is consistent with previous studies that use the LFS, such as Posel and Heintz (2008). On average across the three waves, informal employment represents around 35% of total employment, compared to around 37% in Posel and Heintz (2008). Table 1 also shows that formal non-agricultural wage employment is the most important category of employment in South Africa, accounting for around half of total employment, followed by informal non-agricultural wage employment (around a quarter). Consistent with previous accounts of a relatively small self-employed sector, data also shows that self-employment accounts for an average of only 13% of total employment, with the overwhelming majority of these jobs being informal (around 80%). Finally, agricultural wage employment — that is, agricultural work not including subsistence activities— represents around 10% of total employment.

Examining the evolution of employment between waves provides further insights that are particularly interesting as this period captures the immediate effects of the global financial crisis of 2008-09. The labour market impact of the crisis-induced recession in South Africa was, as noted in a 2010 ILO report— "multifaceted and in some respects unexpected" (Verick, 2010). According to the dualist paradigm, an economic downturn should be accompanied by a contraction of formal sector employment and a concomitant expansion of informal sector employment, with the informal sector absorbing those laid off in the formal sector. However, as Table 1 shows, the share of informal employment fell in the immediate

[14] The sector question for the self-employed was only introduced in 2012.
[15] Cichello et al. (2012) highlights the data quality issues associated with these categories in the 2010 wave.

Table 1: Distribution of employment, by employment type

	2008		2010		2012	
FORMAL	**N**	**%**	**N**	**%**	**N**	**%**
Wage workers	2,385	45.0	2,692	53.2	3,390	67.0
Non-agricultural	1,963	37.1	2,370	46.8	2,876	56.8
of which: private household	90	1.7	94	1.9	86	1.7
Agricultural	239	4.5	260	5.1	358	7.1
Self-employed	176	3.3	100	2.0	122	2.4
Non-agricultural					104	2.1
of which: private households					12	0.2
Agricultural					12	0.2
Total formal	**2,561**	**48.3**	**2,792**	**55.2**	**3,512**	**58.6**
INFORMAL						
Wage workers	1,384	26.1	1,290	25.5	1,268	25.1
Non-agricultural	984	18.6	1,005	19.9	987	19.5
of which: private household	311	5.9	292	5.8	344	6.8
Agricultural	290	5.5	247	4.9	221	4.4
Self-employed	652	12.3	448	8.9	612	12.1
Non-agricultural					560	11.1
of which: private households					154	3.0
Agricultural					28	0.6
Total informal	**2,036**	**38.4**	**1,738**	**34.3**	**1,880**	**31.4**
Casual employment	701	13.2	531	10.5	597	10.0
TOTAL	**5,298**	**100.0**	**5,061**	**100.0**	**5,989**	**100.0**

Source: NIDS and author's own calculations.

Notes: 1/ Question to identify sector of self-employment only available in 2012. 2/ Agricultural wage work does not include subsistence agricultural activities. 3/ Totals may not match individual sum of figures due to individuals who did not identify their sector of employment.

away from self-employment and casual employment, and into regular wage employment, suggesting the precariousness of the former two forms of employment. Further, as noted in Verick (2010), the fall in total employment levels in 2010 did not initially translate into an increase in official unemployment. This is illustrated in a comparison of labour market status between the three waves, which shows that the reduction in employment was instead manifested in a sharp spike in labour market inactivity (Table 2).

Table 2: Changes in labour market status, 2008-2010

	2008	**%**	**2010**	**%**	**2012**	**%**
Not economically active	4,854	37.6	7,839	51.1	7,108	44.0
Unemployed (strict)	1,819	14.1	1,625	10.6	2,624	16.2
Unemployed (discouraged)	944	7.3	826	5.4	427	2.6
Employed	5,298	41.0	5,061	33.0	5,989	37.1
Total	12,915	100	15,351	100	16,148	100

Source: NIDS and author's own calculations.

Table 3 presents average real net monthly earnings and weekly hours worked, for all individuals who reported positive earnings and hours worked. This earnings measure is constructed using net monthly earnings, adjusted using the national consumer price index reported by SSA. The NIDS also includes imputed earnings figures, which have been used to address the issue of high rates of missing earnings data (which is particularly acute for the self-employed). Imputed earnings figures are used for all earnings calculations in this paper. Finally, taking into account severe issues of misreporting in 2010 in which a number of interviewers reported *daily* hours worked, rather than *weekly* hours worked (Cichello et al.,

2012), it is monthly rather than hourly earnings that is chosen as the benchmark earnings measure for analysis.

Table 3: Average real net monthly earnings and weekly hours worked

	2008		2010		2012	
Formal	**Earnings (R)**	**Hours**	**Earnings (R)**	**Hours**	**Earnings (R)**	**Hours**
Wage workers	4,171	41.6	4,147	36.4	4,065	42.2
Non-agricultural	4,529	41.3	4,452	35.9	4,418	42.0
of which: private households	1,209	37.5	1,236	33.4	1,347	37.6
Agricultural	1,442	43.9	1,621	39.9	1,580	43.9
Self-employed	7,713	42.9	5,219	43.0	6,346	45.1
Non-agricultural					6,373	*43.6*
of which: private households					3,769	*52.7*
Agricultural					7,533	*58.4*
Total formal	**4,411**	**41.7**	**4,210**	**36.5**	**4,179**	**42.3**
Informal						
Wage workers	1,378	40.9	1,557	35.5	1,471	40.8
Non-agricultural	1,453	40.2	1,674	35.1	1,524	39.7
of which: private households	863	36.5	929	30.9	1,030	35.4
Agricultural	1,036	43.4	1,033	37.0	1,115	45.3
Self-employed	1,439	30.5	1,172	26.4	1,325	33.3
Non-agricultural					1,444	33.5
of which: private households					1,390	*32.6*
Agricultural					2,102	*37.3*
Total informal	**1,381**	**37.6**	**1,447**	**33.3**	**1,409**	**38.5**

Source: NIDS and author's own
Notes: 1/Earnings are net of taxation, medical aid and pension contributions. 2/ Individuals with more than one job are assigned their main job. 3/ Figures for totals are weighted. 4/ Agricultural wage work does not include subsistence agricultural activities.

On an aggregate level, the data reveals evidence of an informal earnings penalty, with formal workers earning around three times higher than informal workers, on average. Moreover, average earnings for each sub-category of formal employment — non-agricultural wage employment, agricultural wage employment and self-employment — are each found to be higher vis-a-vis their informal counterparts. However, examining the segments within 'formal' and 'informal' shows that there is substantial variation between different employment states. In other words, not all workers who hold formal jobs report higher average earnings than all workers in informal employment. Domestic workers, whether formal or informal, are revealed to be the lowest paid segments, while formal self-employed workers are systematically paid the highest.

There are also clear patterns emerging from the hours worked data, notwithstanding the serious issues of misreporting in 2010 as noted previously. Formal workers tend to work more hours than informal workers, although there is substantial variation between segments. In general, it is agricultural workers, both formal and informal, and the formal self-employed who work the longest hours, with the informal self-employed working the least.

Table 4 presents descriptive statistics for a number of individual and job-specific characteristics, including indicators of health and wellbeing. Data shows a stark age differential when examining workers along the self-employment/wage work divide. The self-employed are found to be much older than their wage earning counterparts, which provide some support to Maloney's (2004) argument that older workers may be more likely to enter

Table 4: Descriptive statistics, non-agricultural employment

Variable	FORMAL						INFORMAL					
	All workers		Wage		Self-employed		All workers		Wage		Self-employed	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Demographic and human capital characteristics												
Age	37.90	10.85	37.93	10.69	43.17	10.77	38.17	11.56	37.18	11.44	40.17	11.56
Male	0.54	0.50	0.51	0.50	0.65	0.48	0.42	0.49	0.41	0.49	0.44	0.50
Married	0.39	0.49	0.41	0.49	0.65	0.48	0.30	0.46	0.27	0.45	0.35	0.48
Number of children living in household	1.85	2.10	1.85	2.06	1.47	1.59	2.10	2.37	2.13	2.40	2.04	2.29
Rural	0.36	0.48	0.27	0.45	0.28	0.45	0.43	0.49	0.40	0.49	0.48	0.50
African	0.68	0.47	0.71	0.45	0.41	0.49	0.81	0.39	0.78	0.41	0.87	0.34
White	0.08	0.27	0.08	0.27	0.39	0.49	0.04	0.20	0.04	0.19	0.05	0.23
Coloured	0.22	0.42	0.19	0.39	0.11	0.31	0.13	0.34	0.16	0.37	0.06	0.24
Asian/Indian	0.02	0.13	0.02	0.14	0.09	0.29	0.01	0.11	0.01	0.11	0.01	0.12
Education level:												
No Schooling or did not complete primary	0.15	0.36	0.09	0.29	0.05	0.22	0.24	0.43	0.23	0.42	0.27	0.44
Completed primary (only)	0.37	0.48	0.34	0.47	0.23	0.42	0.51	0.50	0.53	0.50	0.47	0.50
Completed secondary (only)	0.21	0.41	0.24	0.43	0.21	0.41	0.16	0.37	0.16	0.37	0.16	0.36
Completed tertiary	0.25	0.43	0.30	0.46	0.42	0.49	0.07	0.26	0.07	0.25	0.09	0.29
Literacy	0.92	0.27	0.96	0.20	0.99	0.11	0.86	0.35	0.86	0.35	0.85	0.35
Computer literacy	0.22	0.42	0.26	0.44	0.35	0.48	0.07	0.25	0.06	0.24	0.07	0.26
Job characteristics:												
Skilled	0.28	0.45	0.30	0.46	0.15	0.36	0.09	0.28	0.10	0.31	0.05	0.21
Semi-skilled	0.41	0.49	0.47	0.50	0.29	0.46	0.45	0.50	0.42	0.49	0.50	0.50
Elementary occupation	0.27	0.44	0.20	0.40	0.14	0.35	0.37	0.48	0.46	0.50	0.18	0.38
Industry:												
Private households	0.03	0.17	0.04	0.19	0.02	0.15	0.24	0.43	0.32	0.47	0.10	0.29
Mining and quarrying	0.04	0.20	0.06	0.23	0.00	0.05	0.01	0.11	0.02	0.13	0.00	0.05
Manufacturing	0.09	0.29	0.12	0.33	0.01	0.11	0.10	0.29	0.12	0.33	0.04	0.20
Electricity, gas, water	0.01	0.10	0.01	0.12	0.01	0.08	0.01	0.09	0.01	0.10	0.01	0.07
Construction	0.04	0.18	0.05	0.21	0.04	0.19	0.07	0.25	0.08	0.28	0.04	0.19
Wholesale and retail trade	0.15	0.35	0.19	0.40	0.04	0.19	0.15	0.36	0.20	0.40	0.06	0.23
Transport, storage and communication	0.04	0.19	0.05	0.22	0.03	0.18	0.04	0.18	0.04	0.21	0.02	0.13
Finance, real estate and business services	0.07	0.26	0.10	0.30	0.02	0.14	0.04	0.19	0.05	0.22	0.01	0.10
Community, social and personal services	0.29	0.45	0.38	0.49	0.08	0.27	0.13	0.33	0.16	0.37	0.06	0.24
Health and wellbeing:												
Self-reported health status	2.03	1.00	1.99	0.99	2.06	1.02	2.22	1.08	2.16	1.04	2.35	1.15
Has been diagnosed with:												
Tuberculosis	0.04	0.20	0.03	0.18	0.03	0.16	0.04	0.20	0.04	0.20	0.05	0.21
High blood pressure	0.13	0.34	0.13	0.33	0.20	0.40	0.14	0.35	0.13	0.34	0.16	0.36
Diabetes	0.04	0.19	0.04	0.19	0.09	0.29	0.04	0.18	0.02	0.15	0.06	0.24
Stroke	0.00	0.06	0.00	0.06	0.00	0.00	0.01	0.08	0.00	0.07	0.01	0.08
Heart problems	0.02	0.13	0.02	0.12	0.03	0.16	0.02	0.15	0.02	0.13	0.04	0.19
Cancer	0.00	0.06	0.00	0.06	0.00	0.05	0.00	0.06	0.00	0.05	0.01	0.08
Smokes	0.26	0.44	0.22	0.42	0.21	0.41	0.24	0.43	0.26	0.44	0.21	0.41
Consumes alcohol at least 3 times a week	0.04	0.18	0.03	0.17	0.09	0.29	0.04	0.19	0.04	0.19	0.04	0.19
Current life satisfaction	0.26	0.44	0.26	0.44	0.20	0.40	0.31	0.46	0.30	0.46	0.34	0.47
Felt lonely, depressed, fearful or had restless sleep at least 3 times over the past week	0.76	0.43	0.76	0.43	0.78	0.41	0.72	0.45	0.72	0.45	0.71	0.45
Felt happy or hopeful about the future more than 3 times over the past week	5.71	2.41	5.73	2.38	6.66	2.23	4.94	2.35	5.00	2.34	4.82	2.38

Source: NIDS and author's own calculations.
Notes: 1/ Literacy rates are defined as the ability to read and write in at least one language. 2/ Skilled occupations include managers, skilled occupations include the armed forces clerical support workers, service professionals and technicians. Semi-and sale workers, craft and related trade workers and plant and machinery operators.
3/ Self-reported health status renages from 1= very satisfied to 5= poor. 4/ Current life satisfaction ranges from 1= very satisfied to 10= very dissatisfied.

self-employment after either 1.) accumulating the necessary knowledge and experience from salaried work to open up their own businesses or 2.) having been fired from formal salaried work and are seeking a 'safety net' where age is no limitation to finding work again.

In line with previous studies (Devey et al. 2008, Valodia, 2007; Valodia et al. 2005), informal workers are shown to be disproportionately African and female, with women also more likely to be in wage employment than self-employment. Household characteristics also differ according to informality status, with formal workers, and the formal self-employed in particular, more likely to be married but with fewer children living at home. In line with conventional wisdom, workers with higher levels of education are more likely to be engaged in formal employment and vice versa. While almost half of formal workers have completed at least secondary education, this figure is only around a quarter for those employed informally. This trend is reflected in the allocation of workers across occupation types, with informal

workers much less likely to be employed in skilled jobs, and more likely to hold elementary occupations. There is also substantial variation in sectors of activity between formal and informal workers. While community, social and personal services is the largest industry for both formal wage and self- employment, the dominant sector of employment for informal workers is private households, followed by wholesale and retail trade.

Finally, descriptive statistics of health and wellbeing measures reveal a number of interesting insights. While there are no substantial differences in the likelihood of being diagnosed with major illnesses or health problems, workers in informal employment are likely to self-report poorer health statuses and lower life satisfaction, are more prone to feelings of loneliness, depression or sleeplessness, and are less likely to feel happy or hopeful about the future. While such findings are undoubtedly correlated to an individual's level of earnings, they nonetheless underscore the importance of delving deeper into the characteristics of informal employment beyond monetary measures, in order to examine the nexus between life satisfaction and informality.

Overall, the descriptive statistics indicate that the informality status of a job is correlated with a number of individual characteristics and is therefore unlikely to be randomly assigned across the different employment categories. Controlling for such differences in characteristics is imperative as it is one potential explanation for the observed raw earnings gaps. The empirical strategy used in this paper seeks to not only control for observed characteristics, but also time-invariant unobservables through the use of fixed effects regression. This will shed light on the extent to which such characteristics can account for the observed raw earnings differential, and the extent to which this gap may instead support theories of labour market segmentation.

However, before detailing the empirical methodology, it is imperative to look at employment flows between different employment categories across the three waves. Not only do transition flows provide valuable information on labour market dynamics, their analysis is an important validity check when using fixed effects panel regression, as the estimation strategy requires there be a substantial amount of 'movers' between employment states over time (Nguyen et al., 2013).

c.) Transition matrices

Table 5 shows transition matrices of labour market status between 2008-2010, 2010-2012 and 2008-2012, obtained from the unbalanced panel dataset of all individuals aged 15-64 years. The shaded cells highlight the flows between formal and informal employment and represent the basis for estimating formal-informal earnings gaps. Before examining these patterns of mobility, however, a number of features deserve mention here. Firstly, there is a particularly strong probability of becoming unemployed if a respondent was previously in informal employment, which suggests that both forms of informal employment are relatively more unstable and precarious than formal employment. However, a second and more striking feature of the matrices is the high volume of workers moving from informal self-employment into inactivity. While movement into unemployment (both strict and discouraged) is similar for both types of informal employment (around 10-15%), movement into inactivity is much

Table 5: Transition matrices

2008	Inactive	Unemployed (strict or discouraged)	Casual	Agricultural employment	Formal non-ag. wage employment	Formal self-employment	Informal non-ag. wage employment	Informal self-employment	Total
					2010				
Inactive	70.3	16.3	2.6	1.1	4.3	0.3	3.2	2.0	100.0
Unemployed (strict or discouraged)	45.4	27.9	5.1	1.6	10.0	0.3	6.4	3.3	100.0
Casual	35.1	21.0	7.4	3.0	12.4	1.1	14.3	5.6	100.0
Agricultural	17.3	6.8	2.2	51.5	11.4	0.3	9.6	0.9	100.0
Formal non-ag. wage employment	11.9	6.6	2.2	2.3	**66.7**	0.8	8.0	1.5	100.0
Formal self-employment	19.5	7.8	1.3	2.6	9.1	**40.3**	11.7	7.8	100.0
Informal non-ag. wage employment	18.7	11.5	4.7	6.4	27.5	0.5	**28.0**	2.8	100.0
Informal self-employment	46.7	10.9	4.8	1.4	5.9	2.3	4.6	**23.5**	100.0
Total	46.5	16.7	3.6	3.8	17.5	0.9	7.5	3.6	100.0
					2012				
2010									
Inactive	61.2	21.6	3.5	1.5	5.5	0.2	3.5	3.0	100.0
Unemployed (strict or discouraged)	37.0	31.6	4.9	3.2	11.6	0.4	7.4	4.0	100.0
Casual	28.0	22.3	11.0	2.7	16.7	0.3	12.8	6.3	100.0
Agricultural	11.1	7.7	2.8	58.8	11.9	0.6	5.4	1.7	100.0
Formal non-ag. wage employment	8.8	7.8	2.4	2.2	**70.4**	0.6	6.2	1.5	100.0
Formal self-employment	18.5	6.2	4.6	4.6	7.7	**43.1**	1.5	13.9	100.0
Informal non-ag. wage employment	18.0	12.9	6.5	3.6	27.7	0.7	**27.3**	3.3	100.0
Informal self-employment	34.7	16.7	4.0	0.3	7.3	3.3	5.2	**28.6**	100.0
Total	42.3	19.7	4.0	4.0	18.9	0.7	6.6	3.9	100.0
					2012				
2008									
Inactive	63.8	19.8	3.2	1.4	5.3	0.3	3.5	2.6	100.0
Unemployed (strict or discouraged)	40.8	29.6	5.1	2.4	11.1	0.3	7.1	3.6	100.0
Casual	32.5	21.9	8.6	3.1	13.8	0.8	13.7	5.6	100.0
Agricultural	14.0	7.3	3.3	54.9	11.8	0.5	7.1	1.2	100.0
Formal non-ag. wage employment	10.3	7.2	2.2	2.2	**68.9**	0.7	7.0	1.5	100.0
Formal self-employment	18.1	7.2	2.4	3.0	9.6	**42.8**	7.8	9.0	100.0
Informal non-ag. wage employment	18.6	12.1	5.9	4.9	27.7	0.7	**27.1**	3.0	100.0
Informal self-employment	40.7	13.7	4.6	0.9	7.0	2.8	4.9	**25.5**	100.0
Total	43.36	18.51	3.9	3.97	18.62	0.85	7.06	3.73	100.0

Source: NIDS and author's own calculations.
Note: Agricultural employment does not include subsistence agricultural activities.

higher for the informal self-employed — around 35-45% of the informally self-employed drop out of the labour market between waves, compared to around only 20% for informal wage workers.

Analysis of the characteristics of the informal self-employed suggests that this finding is likely to be attributable to a number of factors. Firstly, as mentioned previously, the informal self-employed are, on average, older than wage employees and this partially explains the substantial flow of workers dropping out of the labour market as they move into retirement. However, this factor cannot wholly explain the volume of movements, as the formally self-employed exhibit a similar age profile and do not suffer the same magnitude of flows into labour market inactivity. Instead, the phenomenon is also likely to be partly attributable to gender and family dynamics, with data indicating that the informal self-employed are much more likely to be female than the formal self-employed. In fact, 70% of the informally self-employed that become inactive between 2008 to 2010 are found to be female workers. While it is difficult to determine with certainty the motivations lying behind these transitions, it is likely that family and child-care responsibilities play a considerable role. Unmarried women comprise over 40% of the informally self-employed dropping out of the labour market between 2008 and 2010, of which over 90% have children. While such figures are only suggestive, it is likely that these individuals, without the assistance of a spouse, are more likely to exit the labour market to fulfill child care responsibilities. This is reinforced by the fact that, for married women with children, there appears a relatively greater tendency to shift into other types of employment, rather than exit the labour market. This is consistent with the findings presented in Mlatsheni and Leibbrandt (2001), who study the effects of motherhood on labour force participation and suggests that South African women (like many mothers in developing countries) are more likely to stay or search for work in order to support their children, because they cannot afford the luxury of staying at home with their young children.

Inextricably entwined with this female labour force participation decision is the level of household income and interactions with other members of the household. Examining of household income reveals that of those informal-self employed who exited the labour market from 2008 to 2010, more than 40% experienced at least a 20% increase in household income. This suggests that any fall in income as a result of dropping out of the labour market may have, at least to some extent, been compensated by an increase in government grants, or other non-labour market income. Indeed, an examination of the level of household government grants suggests this is the case — around 40% of those informally self-employed who left the labour force resided in households where there was at least a 20% increase in government payments. Clearly the interplay between labour market and household decisions is a complex dynamic that is deserving of further research, particularly in the context of informal self-employment.

Turning to transition flows between employed formal and informal workers (shaded cells), there are a number of remarkable patterns that emerge. Firstly, supporting the validity of a fixed effects panel regression, the matrices show that there are indeed substantial flows of workers between employment states. Between each wave, more than a quarter of employed workers changed their employment status, with informal workers more likely to move than

formal workers. There is also significant fluidity in informality status, with substantial movement between formal and informal work.

Secondly, focusing on movements within self-employment and wage employment indicates that informality status is far more fluid for wage workers than for the self-employed. Between 2008 and 2012, around 88% of informal wage employee movements are into formal wage employment, while around 76% of formal wage employees move in the opposite direction. In the self-employed sector, very few informal workers formalise their businesses (around 7% of movers) and while a considerably higher proportion of formal businesses shift into informality (around 35% of movers), the flows are not of the same magnitude as those observed in wage employment.

Thirdly, a more detailed examination of employment categories reveals the highly lucrative nature of formal non-agricultural wage employment, a feature of the South African labour market noted in Chapter II. This category is the most stable of employment states, with workers the most likely to remain in the same category over time (around 70% of workers stay within this segment between each wave). Moreover, formal non-agricultural wage employment represents the most likely destination for all movers, regardless of their original employment status. Indeed over 75% of movers who did not start off in formal non-agricultural wage work, move into this category between 2008 and 2012.

Having conducted a preliminary analysis of the NIDS data, this paper will now detail the empirical strategy used to estimate informal-formal earnings gaps.

VI. METHODOLOGY

In order to investigate the impact of informal employment on earnings, standard Mincerian-type wage equations are estimated using OLS and fixed effects regressions, both at the mean and at various points along the earnings distribution.

First, OLS regression is conducted on a pooled set of individuals in non-agricultural employment, considering informality status (I_{it}) on an aggregate level. The estimated equation is specified as follows:

$$y_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 I_{it} + \varepsilon_{it} \quad (1)$$

where y_{it} denotes log real monthly earnings, X_{it} a vector of socio-demographic and human capital characteristics for individual i, observed at time t and Z_{it} a vector of job specific characteristics. I_{it} is a dummy variable equal to 1 if the individual is employed informally, and 0 otherwise, and ε_{it} is an i.i.d normally distributed error term. The covariates include variables indicating age[16], gender, marital status, number of children in the household, ethnic

[16] Age and age-squared are included instead of variables to indicate potential years of experience, as used in many standard Mincerian earnings equations. This is due to the particular nature of the South African labour market, which is characterised by high levels and long spells of unemployment as well as an education system plagued by high repetition rates.

population group, education level, literacy in reading, writing and computer skills, occupation skill level, and two year dummies to control for macroeconomic trend effects on earnings.

Following Nguyen et al. (2013), formal and informal workers are then divided into sub-categories with the construction of four dummy variables, taking a value of 1 if the individual is a formal non-agricultural wage worker (FW_{it}), formal self-employed (FSE_{it}), informal non-agricultural wage worker (IW_{it}), or informal self-employed (ISE_{it}). The model is estimated using the following specification, with informal self-employed workers as the reference category:

$$y_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 FW_{it} + \beta_4 FSE_{it} + \beta_5 IW_{it} + \varepsilon_{it} \quad (2)$$

The estimated coefficients β_3 to β_5 can be interpreted as the conditional earnings gap for wages in formal non-agricultural wage employment, formal self-employment, and informal non-agricultural employment, respectively, compared to wages in informal self-employment.

Next, in order to allow earnings gaps to differ along the income distribution, quantile regression analysis is conducted with the following model specifications:

$$q_\varrho(y_{it}) = \alpha + \beta_1(\varrho) X_{it} + \beta_2(\varrho) Z_{it} + \beta_3(\varrho) I_{it} + \varepsilon_{it}, \forall\, \varrho \in [0,1] \quad (3)$$

$$q_\varrho(y_{it}) = \alpha + \beta_1(\varrho) X_{it} + \beta_2(\varrho) Z_{it} + \beta_3(\varrho) FW_{it} + \beta_4(\varrho) FSE_{it} + \beta_5(\varrho) IW_{it} + \varepsilon_{it}\ , \forall\, \varrho \in [0,1] \quad (4)$$

where $q_\varrho(y_{it})$ is the ϱth quantile of log real hourly earnings, the coefficients $\beta_1(\varrho)$ and $\beta_2(\varrho)$ are the returns to each characteristic at the ϱth quantile, and the set of coefficients from $\beta_3(\varrho)$ to $\beta_5(\varrho)$ measure the earnings gap attributable to different employment statuses at the various quantiles. This captures the heterogeneity in returns to observed characteristics as well as allows the size of the earnings gap to differ along the earnings distribution, providing more complete results than those derived from OLS.

To account for time-invariant unobserved heterogeneity, a fixed effects regression is then run on a panel of employed individuals who appear in at least two out of the three waves. Using the common personal identifier across years to match respondents, this leaves a sample of 6,359 individuals. The estimated fixed effects equations are as follows:

$$y_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 I_{it} + \mu_i + \varepsilon_{it} \quad (5)$$

$$y_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 FW_{it} + \beta_4 FS_{it} + \beta_5 IW_{it} + \mu_i + \varepsilon_{it} \quad (6)$$

where notation follows the specifications outlined previously[17], and μ_i is an time-invariant

[17] Time-invariant control variables are dropped in the fixed effects regressions.

unobserved individual fixed effect. Using an extension of the traditional quantile regression model originally developed by Koenker (2004), fixed effects quantile regressions are also run under the assumption that the unobserved heterogeneity term affects all quantiles in the same manner. Under this simpler assumption proposed by Canay (2011), the unobserved individual fixed effect, μ_i, can be estimated using traditional fixed effects OLS regression, and subsequently used to adjust earnings derived from a traditional quantile regression, such that $\hat{y}_i = y_i - \hat{\mu}_i$.

Finally, all regression analyses are replicated on subsamples of men and women, with the rationale that motivations for informality may be driven by gender-specific considerations. For example, with non-market childcare work, women may have a welfare function that takes non-pecuniary aspects into greater consideration such that informal work may become more attractive (Nguyen et al. 2012). This is suggested by the data, which shows that women are more likely to be employed informally. Further, the raw gender earnings gap is more pronounced for informal employment than it is for formal employment. Together, these observations provide ample motivation for a deeper exploration into the gender dimension of informal-formal earnings gaps.

Before proceeding to estimation results, the issue of potential sample selection bias should be addressed, as it is often raised as a challenge in studies of the informal-formal earnings gap, and on analyses of non-experimental data more broadly (Badaoui et al. 2008; Pratap and Quintin, 2006). However, as Posel and Heintz (2008) point out, correcting for selection bias requires that "the variables that explain selection into a particular employment type (e.g. informal employment) are independent of those variables used to explain earnings". The variables used in this study to explain selection would also affect earnings and therefore a strategy to correct for selection bias is not pursued. Following several other studies, the strategy used here reasons that there is sufficient information available on workers to control for potential selection bias, including the ability to control for individual fixed effects and to take into account intrinsic gender differentials in labour force participation rates by running separate regressions for men and women.

VII. RESULTS

a.) Informal vs formal employment (aggregate analysis)

Table A1 (columns 1 and 2) reports the OLS estimates for the pooled earnings function, with informal employment measured on an aggregate level. The earnings gap is first estimated using a model specification that includes only the informal dummy and year dummies, with results confirming the existence of a significant and large informal wage penalty of 105%. However, when the model is re-estimated to include socio-demographic and human capital characteristics, this penalty drops considerably to 66%, and further again to 56% once job-specific characteristics are also controlled for. In other words, more than 45% of the raw earnings differential can be attributed to observable individual and job specific characteristics. While the informal penalty of 56% is larger than that reported by Badaoui et

al. (2008)[18], it is consistent with other studies such as Bargain and Kwenda (2011) and Kingdon and Knight (2007a) who both find a penalty of around 60%.

The results also confirm that the standard explanatory variables included in Mincerian equations are significant and have coefficients with expected signs. Age is significantly and positively correlated with earnings and those who have completed secondary and tertiary schooling attract an increasing premium of 32% and 68% respectively. Being male, married, having the ability to read and write in at least one language, or being computer literate, are also found to have a significant and positive effect on earnings. As expected, white workers attract a large and significant premium of 61% vis-à-vis African workers, as does being Asian or Indian (37%), while being Coloured has no significant effect. The coefficients for job specific characteristics are also consistent with a priori expectations, with those working in elementary or semi-skilled occupations earning less, on average, than those in skilled work.

After controlling for time-invariant unobserved heterogeneity using fixed effects panel regression (Table A1, column 4), the informal-formal earnings gap falls considerably to 20%, in line with Badaoui et al.'s finding of 18%. This narrowing of the earnings gap is consistent with the international literature that similarly finds that the informal wage penalty can largely be attributed to the poor unobserved characteristics of informal workers. However, the size of South Africa's informal penalty is relatively large in the international context, with Nguyen et al. (2012) finding a penalty of 15% in Vietnam, Tansel and Kan (2012) a penalty of 7% in Turkey and Nordman et al. (2013) a penalty of 3% in Madagascar. In their cross country study, Bargain and Kwenda (2011) also note the relatively large size of South Africa's informal penalty and argue that this is unsurprising in a labour market where the legal institutions of formal employment, such as unionisation and the minimum wage, have been found to be particularly effective.

Furthermore, distributional analysis shows that informal workers suffer a significant wage penalty at all income levels (Table A4, columns 1 to 5). The largest penalty is suffered by the bottom 10% of earners (65%) and the smallest by those in the top 10% (48%). This finding that earnings differentials are smaller among high income groups than among low income groups is consistent with the findings in the international literature discussed in Chapter IV. Even when taking into account time-invariant unobservables, the informal penalty remains at all income levels, although the penalty is now of a similar magnitude along the entire distribution (Table A7, columns 1 to 5). This suggests that lower income earners have worse unobserved characteristics than those at the top end of the scale. It is also interesting to note that the average formal worker has more productive unobserved individual characteristics than the average informal worker regardless of their level of income, with the informal penalty narrowing considerably at all points in the distribution in the fixed effects regression.

Thus, there is evidence of a significant informal earnings penalty, which supports dualist theories of labour market segmentation. However, aggregate analysis may conceal structural

[18] Badaoui et al. (2008) focus solely on a subsample of male, private salaried employees and find a penalty of 37%, While a public and private split is unavailable in the NIDS, re-estimating the model on a sample of male, wage employees produces a similar informal penalty of 44%.

heterogeneity within formal and informal employment. There may also be earnings differentials between different types of employment that suggest barriers of entry to and mobility within informal employment, not just barriers into formal employment.

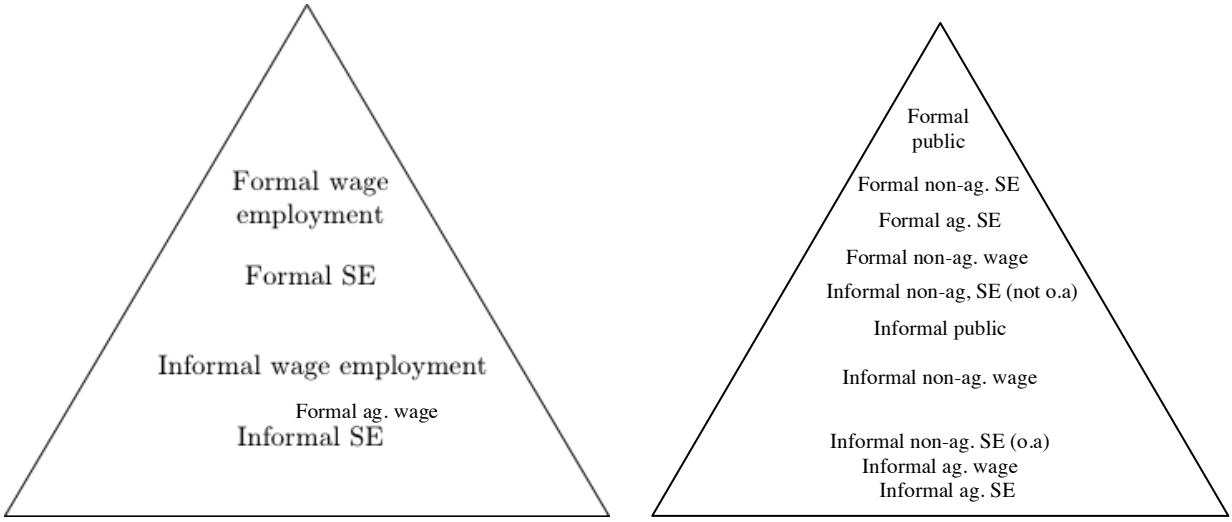
b.) Informal vs formal employment (disaggregate analysis)

The results of the pooled OLS for the disaggregate analysis are shown in Table A1 (column 3). Informal self-employment is taken as the reference category and results are interpreted accordingly.

Results show that all types of employment attract a premium vis-à-vis informal self-employment —69% for formal wage employment, 32% for formal self-employment and 24% for informal wage employment[19]. The implications are twofold: firstly, that there is an informal penalty in both wage employment and self-employment and secondly, that the self-employed occupy the lower tier in both formal and informal employment (contrary to the bulk of existing empirical literature).

The South African labour market can thus be characterised by a hierarchical structure depending on employment status, as summarised in the left panel of Figure 2. It is, however, important to note that such broad divisions between wage employment and self-employment are in reality far more nuanced, as suggested by Posel and Heintz (2008) who undertake a more granular dissection of the labour market. Their results indicate that while the formally self-employed, on average, earn less than formal wage workers in the public sector, they earn more than those employed privately. Further, while the informally self-employed with their own account (o.a) earn less than informal non-agricultural wage workers, the informally self-employed without their own accounts earn more. The structure as suggested by the findings of Posel and Heintz (2008) is represented in the right panel of Figure 2.

Figure 2: Hierarchical structure of South African employment



Source: Posel and Heintz (2008), NIDS and author's own calculations.

[19] This is broadly consistent with Bargain and Kwenda (2011) who find an informal penalty of around 62% for wage workers and 30% for the self-employed.

However, Posel and Heintz (2008) do not account for individual fixed effects and, once included, the earnings gap between the informal self-employed and informal wage workers narrows (Table Al, column 5). This suggests that the informal self-employed also have poorer unobserved characteristics than wage workers. There is also a fall in the penalty for the formal self-employed vis-à-vis formal wage workers (results not shown). Overall, therefore, the evidence suggests that the self-employed have less favourable unobserved skills than wage employees, regardless of informality status.

Comparing informal wage workers to their formal counterparts, the informal penalty drops to 12% when unobserved individual heterogeneity is accounted for (results not shown)[20]. This suggests that informal workers, on average, have detrimental unobserved characteristics relative to formal wage workers. Again, the magnitude of this informal penalty is relatively high in the international context, with 11%, 9% and 5% found for Vietnam, Madagascar and Turkey respectively (Nguyen et al. 2012; Nordman et al. 2013; Tansel and Kan, 2012). While the informal penalty also narrows in the self-employed sector, the earnings gap between formal and informal self-employment becomes insignificant.

Quantile regression results (Table A7, columns 6 to 10) shed further light on the structure of employment in South Africa. When fixed effects are taken into account, the premium observed by informal wage workers over the informal self-employed is decreasing with income, turning into a penalty at the 75th quantile (although it is insignificant). This is mirrored in the formal market, with wage workers also attracting a decreasing earnings premium vis-à-vis the self-employed (results not shown), with the earnings gap no longer significant for higher income earners.

Another interesting insight is revealed when comparing estimates for formal and informal self-employment. While estimates at the mean showed an insignificant pay gap between the two segments, the quantile results reveal a significant premium for formalisation for the top 10% when controlling for unobserved individual attributes (Table A7). This is in contrast to findings in Vietnam and Madagascar, where formalisation is found to be beneficial at all income levels (Nguyen et al. 2012; Nordman et al. 2013). For South African wage employees, however, formal employment is systematically better than informal employment, attracting a significant premium at all quantiles, even when time-invariant unobservables are taken into account (results not shown). Moreover, the premium is generally increasing with income, with the earnings advantage starkest for the top 10% of earners (21%). This distributional pattern is in contrast to that observed by Bargain and Kwenda (2011), who find a decreasing premium.

Overall, the evidence strongly suggests a high degree of heterogeneity in the structure of formal and informal employment, with earnings gaps varying both between types of employment and along the income distribution. The findings can be summarised with two broad conclusions. Firstly, there is strong evidence of a persistent informal penalty for wage employment. In terms of monetary gains, formal wage employment is systematically seen as a preferable option to informal wage employment. This finding is consistent with a segmented

[20] This compares to Bargain and Kwenda's (2011) result of a 21% average informal penalty for wage employees.

labour market in South Africa. However, a similar broad conclusion does not hold for the self-employed sector, with formalisation attracting a significant premium only for the highest income earners. Secondly, while the OLS regression results dispute the widely claimed characterisation of informal self-employment occupying the "upper-tier" of employment, quantile regression results reveal a far more nuanced picture, suggesting this is not the case for the top 25% of income earners.

c.) Gender analysis

The above regression analyses are replicated for separate subsamples of men and women, relaxing the assumption that there exists a common earnings function across both genders.

Tables A2 and A3 show that when the model is re-estimated without controlling for individual or job characteristics, females suffer a greater informal earnings penalty (114%) than men (85%). This differential persists but narrows when individual characteristics are introduced (72% vs. 55%). Estimates indicate that women receive a higher return to education, but a considerably lower return to being white, suggestive of a "white male" bias in the South African labour market. When job specific characteristics are taken into account, the gender gap again narrows, with informal female workers suffering a 60% penalty and male workers a 50% penalty. The penalty narrows more for women than for men, which may reflect the fact that women tend to be employed in lower skilled jobs than men.

However, when unobserved characteristics are taken into account, there is a negligible difference in the magnitude of the informal penalty across genders (21% for women and 20% for men). In other words, accounting for both observed and unobserved differences, there is no gender differential in the size of the informal earnings penalty. However, distributional analysis reveals a more detailed picture(Tables A8 and A9). An informal penalty is observed for both genders at all levels of income, but women suffer a greater penalty in the upper half of the pay scale, while men suffer a greater penalty in the bottom half. Further, the informal penalty for women is increasing with income, while the penalty for men is decreasing.

Disaggregate analysis shows that for wage employment, both men and women experience significant informal penalties, even when accounting for individual fixed effects and at all income levels (results not shown). Again, while men suffer a greater penalty than women in the bottom half of the distribution, women suffer a larger penalty in the top half. Assessing earning gaps for the self-employed, the results are similar to the aggregate sample, with the informal penalty only significant for the top 10% of both female and male workers (Tables A8 and A9).

Analysis of the aggregate sample also showed that wage employees held an earnings advantage over self-employed workers, regardless of informality status, but not at the upper end of the income distribution. A gender breakdown reveals an even more nuanced picture. In the informal labour market, the penalty for self-employment is much larger for women, with the differential most acute for the poorest 10% of women (Tables A8 and A9). These workers attract a penalty of 63%, compared to a penalty of only 23% for the equivalent male worker. Women also suffer a greater penalty for self-employment in the formal market,

regardless of income level (results not shown). These findings suggest that women are at a relatively greater disadvantage when in self-employment than men, regardless of informality status.

Overall, gender analysis shows that while there is little evidence of gender discrimination on an aggregate level and when assessing earnings at the mean, delving deeper using distributional analysis on different employment types reveals a somewhat different picture. Examining informal employment as a whole, men are found to suffer a greater earnings penalty than women at the lower end of the pay scale. This is shown to be driven by a greater penalty for informal wage earners vis-à-vis their formal counterparts. In other words, in the bottom half of the income distribution, men benefit more from formality than women. On the contrary, in the top half of the distribution, it is women who would benefit more from becoming formal. For both genders, there is less compelling evidence of a significant effect of formality for self-employed enterprises. However, women are found to be relatively much better off in wage employment compared to self-employment, regardless of informality status, and particularly for those in lower income brackets.

VIII. DISCUSSION

a.) Further discussion of key results

Using NIDS data from 2008-2012, this paper finds evidence in support of theories of labour market segmentation in South Africa, with a significant earnings penalty observed for informal employment, on average, at all levels of income, for both men and women. Fixed effects estimation confirms the existence of this informal penalty even when time-invariant unobserved heterogeneity is taken into account, while also suggesting that informal workers, on average, hold less productive unobserved attributes than formal workers at all income levels. However, the evidence also indicates that earnings differentials can differ according to gender, income level and type of employment, thus challenging a strict dualist representation of the South African labour market.

In the self-employed sector, an informal penalty is observed in the pooled OLS results, although this becomes insignificant at almost all quintiles once individual fixed effects are included. This insignificance may, however, be attributed to the relatively small sample size of the self-employed. Examining wage employment, the evidence for an informal penalty is far more compelling, with a significant penalty observed for both genders, at all income levels. It should also be noted that the analysis was conducted using monthly earnings, net of taxation, medical aid and pension contributions. As these benefits are afforded only to the formally employed, inclusion of these in earnings would be expected to only exacerbate the informal-formal earnings gap.

Another remarkable finding is the observation of a larger informal penalty for female wage employees in the upper income brackets, but a larger penalty for men at lower incomes. Analysis of earnings by industry and gender sheds some light on this result. Women in lower paying jobs tend to be domestic workers in private households, an industry where earning gaps between being the formally and informally employed are small relative to other industries (Table 6). Meanwhile, male informal workers tend to be less concentrated on

particular industries, such that those on the lower end of the income scale are in sectors of activity where the benefits of being formally employed are greater, such as trade and manufacturing. In the top half of the earnings distribution, over 60% of women are in skilled occupations, compared to only around 30% of men. This could be consistent with a number of explanations. For example, it may suggest that women in skilled occupations are relatively more willing to accept lower informal wages in return for greater non-pecuniary benefits. Alternatively, it may also reflect the existence of barriers into skilled occupations in the formal labour market, which are more prohibitive for women than it is for men. Further investigation is needed to shed light on these potential drivers of gender differences in the informal wage penalty.

Table 6: Average real monthly earnings by industry and gender (Rand)

	Male			Female		
Sector	Informal	Formal	Difference	Informal	Formal	Difference
Private households	867	1,274	407	835	1,107	272
Mining and quarrying	1,867	4,215	2,348	2,696	4,408	1,712
Manufacturing	1,379	3,096	1,716	1,007	1,846	839
Electricity, gas, water	2,299	4,023	1,724	1,283	5,382	4,099
Construction	1,624	2,860	1,236	1,269	1,937	668
Wholesale and retail trade	1,870	4,252	2,381	1,328	3,159	1,831
Transport, storage and communication	1,430	2,285	855	733	1,301	568
Finance, real estate and business services	2,356	3,421	1,064	1,365	3,708	2,343
Community, social and personal services	2,081	4,624	2,542	1,282	3,988	2,705

Source: NIDS and author's own calculations

The empirical results also indicate a degree of structural heterogeneity within informal employment, with both male and female workers in wage employment attracting an earnings premium over those who are self-employed. However, this is only true for those in the bottom 75% of income earners, with the premium turning into a (insignificant) penalty thereafter. Nevertheless, as this result is not in line with the bulk of international evidence, it is deserving of further investigation.

Firstly, it is notable that the earnings penalty for the self-employed vis-à-vis wage workers only becomes significant when individual and job characteristics are included, suggesting that wage employees have more productive observed characteristics. This suggests that if informal wage workers held identical attributes to their self-employed counterparts, they would in fact be significantly lower paid. Indeed, this finding is consistent with Tansel and Kan (2012), who similarly find that the informal self-employed attract a penalty vis-à-vis informal wage workers in Turkey only when individual and job characteristics are introduced.

The use of monthly instead of hourly earnings may have also played a vital role in the observed earnings gap between wage workers and the self-employed. As the descriptive statistics showed, the informal self-employed work, on average, fewer weekly hours than informal wage employees, contributing to a higher average *hourly* wage for the self-employed, despite a lower *monthly* income. Replicating regression analyses using real hourly earnings as the dependent variable shows that while narrowing considerably, there is still a significant wage employment premium over the self-employed for lower income earners (Table A10). This premium again decreases with income and turns into a significant penalty for the top

25% of the income scale. Overall, the distributional patterns are consistent with the previously reported findings that used monthly earnings[21]. Thus, differences in hours worked, while partly accounting for the size of the wage employment premium, cannot wholly explain the result.

Rather, the finding may also be partly attributable to another limitation of the NIDS data — the inability to distinguish between employment in the formal and informal sector due to the absence of a question on firm registration status, as mentioned previously. It is impossible to determine, for example, whether an informal wage employee is working in a formal or informal firm. Therefore, it is likely that the premium for informal wage employment over the informal self-employed stems from those informal employees working in the formal sector.

The shift from a wage employment premium into a penalty along the earnings distribution also suggests a degree of heterogeneity within the informal self-employment sector itself and indeed, the literature finds ample evidence supporting this claim. Studies highlight the existence of both survivalist operations and aspirant entrepreneurs within the informal self-employed sector (Neves et al. 2011). Nevertheless, the evidence suggests that the majority of informal enterprises are of the small scale, survivalist type, established to avoid unemployment rather than to exploit a promising business opportunity (Ligthelm, 2008). Ligthelm (2008) estimates that only between 10% and 15% of informal entrepreneurs have the business acumen to expand and develop their businesses. With such limited scope for growth or upscaling, it is not surprising to find such high rates of business mortality amongst informal entrepreneurs in South Africa, with only an estimated quarter of all new businesses surviving beyond 3.5 years (*ibid.*). This may not only help to explain the earnings penalties observed by the self-employed, but also corroborates the results of the transition matrices presented in Chapter V, which showed a low proportion of informal self-employed remaining in this same employment category between waves. Those entrepreneurs that do survive are likely to be those located at the upper end of the earnings distribution, attracting a premium over their salaried counterparts.

b.) The relationship between informality and subjective wellbeing

A heterogeneous informal self-employed sector also indicates a diversity of motivations that go beyond profit-maximising objectives. Research suggests that the informally self-employed in South Africa may also take into account factors such as "ensuring food security, mitigating risk, and repositioning themselves within their social networks" (Neves et al., 2011). This indicates the role of non-pecuniary considerations in an individual's labour market decisions. As mentioned previously, one of the key strengths of the NIDS dataset is the inclusion of several measures of health and wellbeing, which allows the relationship between informal labour and non-monetary measures of utility to be explored. Theoretically, it may be the case that informal workers voluntarily accept lower wages in return for non-monetary advantages that informal employment might bring, and are thus more satisfied with life despite having a

[21] The analysis was also replicated on a sample omitting all observations that worked 10 or less weekly hours, as well as only on observations in the 2008 and 2012 waves. These strategies were to circumvent the issue of severe misreporting of the weekly hours worked variable in 2010. However, both estimations produced similar results, with a decreasing wage employment premium that turns into a penalty at the upper end of the income distribution.

lower income. The descriptive statistics shown in Chapter V, however, suggested that this is not the case, with the average informal worker reporting lower life satisfaction.

In order to more rigorously investigate the impact of informality on wellbeing, an ordered logit model is estimated on pooled data using life satisfaction as the dependent variable and controlling for a number of individual and household characteristics including age, gender, race, marital status, number of children in the household, household income, education level, employment status and self-reported health. It also takes advantage of a question included in the NIDS asking respondents to rank their perceived relative economic status on a six-step ladder, with the bottom rung representing the poorest and the top rung the richest. Respondents are asked to assess their relative position both in the present and in two years' time. As noted in Posel (2012), there is a substantial body of literature stemming from both psychology and economic fields showing that self-assessed life satisfaction is influenced not simply by an individual's actual income, but also how they perceive their economist status relative to others (Luttmer, 2005; Kingdon and Knight, 2007b; Bookwalter & Dalenberg, 2009).

The results, shown in Table A11, indicate that informal employment overall has a significant and negative impact on life satisfaction, challenging the hypothesis that informal workers might be voluntarily accepting lower wages in return for non-pecuniary advantages. Further analysis by employment category shows that formal self-employment has the greatest positive impact on life satisfaction, followed by formal wage employment, informal self-employment, and finally informal wage employment. In other words, self-employment has a more positive effect on life satisfaction than wage employment, regardless of informality status. This is in contrast to the findings of the earnings analysis, where self-employment was generally found to occupy the lower tier of employment. Thus, the findings of the wellbeing analysis suggest that there may be some non-pecuniary advantages offered by self-employment that improve life satisfaction, that may compensate the self-employed for lower earnings. Moreover, the sign and coefficients of the explanatory variables generally accord with a priori expectations, with marriage, good health, household income and expectations of upward income mobility all having a positive effect on current life satisfaction. Moreover, all else equal, all population groups are expected to have a higher life satisfaction than Africans, educated individuals are more satisfied than those with no schooling, and those who perceive themselves to be of lower economic status than the top third of income earners are less satisfied in life.

However, this analysis does not account for unobserved individual heterogeneity, which has been found to play a crucial role in studies of subjective wellbeing (Alem, 2013). For example, it does not control for intrinsic personality differences, such as how optimistic an individual may respond to questions of self-assessed satisfaction or differences in family background and history. While there is no consensus in the literature on an empirical strategy to consistently estimate an ordered logit model with fixed effects, a number of approaches have been adopted, such as converting the ordinal dependent variable into a binary variable and applying a conditional logit estimator, or assuming cardinality of the ordered variable and using a first differencing technique (Baestchmann et al. 2015). This is

an avenue of research worthy of rigorous further investigation, particularly given the preliminary results that suggest varying motivations according to employment type.

IX. CONCLUSION

This paper draws on a previously unused dataset, the NIDS, in order to investigate the existence of an informal earnings penalty in South Africa and assess the validity of labour market segmentation theories. It finds evidence of an overall earnings penalty for informal employment at all levels of income, for both men and women, even when accounting for time-invariant unobserved heterogeneity. This supports the existence of segmentation.

However, using fixed effects quantile regression and disaggregating employment into four categories allows for a more complete picture to emerge. A strict dualist representation is found to be inappropriate, with earnings differentials found to differ by gender, income level and employment type. There is strong evidence of a systematic and persistent informal penalty for wage employment, although evidence in the self-employed sector is less compelling. The results also challenge the widespread characterisation of informal self-employment occupying the "upper-tier" of employment, with the self-employed found to suffer a penalty for all but the top 25% of income earners.

A detailed gender analysis of earnings gaps is also conducted, a previously unexplored avenue in the South African context. While there is little evidence of gender discrimination on an aggregate level, delving deeper shows that men suffer a greater earnings penalty in the lower half of the pay scale while the opposite holds true in the top half of the distribution. Women are also found to be relatively much better off in wage employment compared to self-employment, regardless of informality status, and particularly for those in lower income brackets.

Finally, the paper is the first to investigate the impact of informality on wellbeing, finding that informal employment overall has a significant and negative impact on life satisfaction. This challenges the hypothesis that informal workers might voluntarily accept lower wages in return for non-pecuniary advantages. Further, the self-employed, both formal and informal, are expected to experience greater life satisfaction vis-à-vis their salaried counterparts, suggesting that there may be some non-pecuniary benefits attached to self-employment.

However, a number of limitations should be noted. Firstly, the NIDS does not allow for the identification of formal and informal firm. It also does not allow for self-employment to be disaggregated according to those with and without their own account, which may conceal some heterogeneity within informal self-employment. Secondly, the regressions do not control for other job characteristics that may affect motivations for informality, such as riskiness, safety or other fringe benefits. Thirdly, earnings do not take into account longer run social welfare benefits such as pension or unemployment payments, which are accrued only to those employed formally. Fourthly, unobserved individual heterogeneity is not controlled for in the analysis of the impact of informal employment on wellbeing. This represents a potentially fruitful avenue for future research. Further exploration into the factors driving the self-employment penalty, particularly for women, would also be invaluable.

Notwithstanding these limitations, the findings presented in this paper considerably broaden the existing evidence base on the nature of the informal economy in South Africa and have important implications for how the role of the informal economy is viewed within South Africa's development strategies. It has revealed a highly heterogeneous characterisation of informal employment that challenges the dualist notion of the "second economy" and emphasised the importance for policymakers to recognise the needs of different segments in order to identify appropriate measures to support the livelihood strategies of those living on the margins.

REFERENCES

Alem, Y. (2013). *Relative Standing and Life-Satisfaction: Does Unobserved Heterogeneity Matter?* Working Papers in Economics No. 579. University of Gothenburg.

Arabsheibani, G. and Staneva, A. (2012). *Is there an Informal Employment wage Premium? Evidence from Tajikistan, IZA* Discussion Paper, No. 6727, IZA, Bonn.

Arias, O., and Khamis, M. (2008). *Comparative advantage, segmentation and informal earnings: A marginal treatment effects approach*. IZA Discussion Paper No. 3916.

Arora, V., and Ricci, L. A. (2005). Unemployment and the labor market. *Post-apartheid South Africa: the first ten years*, pp. 23-47.

Bacchetta, M., Ernst, E., and Bustamante, J. P. (2009). *Globalization and informal jobs in developing countries*. International Labour Organization.

Badaoui, E.E, Strobl, E., and Walsh, F. (2008). Is there an informal employment wage penalty? Evidence from South Africa. *Economic Development and Cultural Change*, *56*(3), pp. 683-710.

Baetschmann, G., Staub, K. E., and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *178*(3), pp. 685-703.

Bargain, O., and Kwenda, P. (2011). Earnings Structures, Informal Employment, and Self-Employment: New Evidence From Brazil, Mexico, and South Africa. *Review of income and wealth*, *57*(1), pp.100-122.

Barrientos, A., and Barrientos, S. W. (2002). *Extending social protection to informal workers in the horticulture global value chain*. Washington, DC: World Bank.

Bhorat, H., Lundall, P., and Rospabe, S. (2002). *The South African labour market in a globalizing world: Economic and legislative considerations*. International Labour Organisation.

Bookwalter, J. T. and Dalenberg, D. R. (2009). Relative to what or whom? The importance of norms and relative standing to well-being in South Africa. *World Development* 38(3) pp. 345-355.

Bucheli, M., and Ceni, R. (2010). Informality sectoral selection and earnings in Uruguay. *Estudios Económicos*, pp. 281-307.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, *14*(3), pp. 368-386.

Carneiro, F. G., and Henley, A. (2001). Modelling formal vs. informal employment and earnings: micro-econometric evidence for Brazil. *University of Wales at Aberystwyth Management & Business Working Paper*, (2001-15).

Chandra, V., and Nganou, J. (2001). Obstacles to formal employment creation in South Africa: Evidence from recent firm surveys. In *DPRU Conference, Johannesburg, November*.

Chang, D. O. (2009). Informalising labour in Asia's global factory. *Journal of Contemporary Asia*, *39*(2), 161-179.

Chen, M. (2014). Informal employment and development: patterns of inclusion and

exclusion. *European Journal of Development Research*, *26*(4), pp.397-418.

Cichello, P. L. (2005). *Hindrances to self-employment activity: evidence from the 2000 Khayelitsha/Mitchell's Plain survey*. Centre for Social Science Research, University of Cape Town.

Cichello, P., Leibbrandt, M. and Woolard, I. (2012). Labour Market: Analysis of the NIDS Wave 1 and 2 Datasets. *SALDRU Working Papers 78*, Southern Africa Labour and Development Research Unit, University of Cape Town.

Cohen, J. (2010). How the global economic crisis reaches marginalised workers: the case of street traders in Johannesburg, South Africa. *Gender & Development*, 18(2) pp. 277-289.

De Soto, H. (1990). *The Other Path: The Invisible Revolution in the Third World*. New York: Harper and Row.

De Vreyer, P.and Roubaud, F. (2013). *Urban labour markets in sub-Saharan Africa*. Africa Development Forum Series, Washington, DC:World Bank.

Devey R, Skinner, C. and Valodia, I. (2005). The state of the informal economy. In Buhlungu, S., Daniel, J. Latchman, J. and Southall, R. (eds) *The State of the Nation 2005-2006* .HSRC Press.

Devey, R., Lebani, L., Skinner, C. and I. Valodia (2008). The Informal Economy. In: Kraak, A. and K. Press (eds.) *Human Resources Development Review 2008: Education, Employment and Skills in South Africa*. Human Sciences Research Council Press, Pretoria, South Africa.

Du Toit, A., and Neves, D. (2007). In search of South Africa's 'second economy''. *Africanus*, *37*(2), pp. 145-174.

Falco, P., Kerr, A., Rankin, N., Sandefur, J. and Teal, F. (2010).The Returns to Formality and Informality in Urban Africa. CSAE WPS 2010-03, Oxford.

Falco, P., Maloney, W. F., Rijkers, B., and Sarrias, M. (2013). Heterogeneity in subjective wellbeing: an application to occupational allocation in Africa.*Documento CEDE*, (2013-01).

Fields, G. S. (1990). Labour market modelling and the urban informal sector: theory and evidence. In: OECD, *The Informal Sector Revisited,* Paris, France

— (2005) A guide to multisector labor market models. *Social Protection Discussion Paper Series No. 0505.* Washington, DC, World Bank.

García, G. A. (2014). *Labor informality: choice or sign of segmentation? A Quantile Regression Approach at the Regional Level for Colombia* (No. 011213). Departamento Nacional de Planeacion

Godfrey, P. C. (2011). Toward a theory of the informal economy. *The Academy of Management Annals*, *5*(1), pp. 231-277.

Gindling, T. H. (1991). Labor market segmentation and the determination of wages in the public, private-formal, and informal sectors in San Jose, Costa Rica. *Economic Development and Cultural Change*, pp. 585-605.

Guifu, C., and Shigeyuki, H. (2009). Formal employment, informal employment and income differentials in urban China. *MPRA Paper No. 17585.*

Harati, R. (2013). Heterogeneity in the Egyptian informal labour market: choice or

obligation?. *Revue d'économie politique*, *123*(4), pp. 623-639.

Harris, J. R., and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American Economic Review*, 126-142

Hart, K. (1973). Informal income opportunities and urban employment in Ghana.*The journal of modern African studies*, *11*(01), pp. 61-89.

Henley, A., Arabsheibani, G. R., and Carneiro, F. G. (2009). On defining and measuring the informal sector: Evidence from Brazil. *World Development*,*37*(5), pp. 992-1003.

Huber, P., and Rahimov, U. (2014). Formal and Informal Sector Wage Differences in Transition Economies: Evidence from Tajikistan. *MENDELU Working Papers in Business and Economics 48/2014*. Mendel University in Brno.

ILO (1993).Resolution concerning statistics of employment in the informal sector, adopted by the Fifteenth International Conference of Labour Statisticians. In *Fifteenth International Conference of Labour Statisticians (Geneva, 19-28 January 1993),* Report of the Conference, Geneva: ILO.

— (2012). Statistical update on employment in the informal economy,June 2012.Geneva:ILO.

Kerr, A. and Teal, F. (2012). The Determinants of Earnings Inequalities: Panel data evidence from South Africa. *CSAE Working Paper Series 2012-04*, Centre for the Study of African Economies, University of Oxford.

Kingdon, G. G., and Knight, J. (2004). Unemployment in South Africa: The nature of the beast. *World development*, *32*(3), pp.391-408.

—(2007a). Unemployment in South Africa, 1995–2003: causes, problems and policies. *Journal of African Economies*, *16*(5), pp.813-848.

— (2007b) Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior and Organization,* 64,pp. 69-90.

Kingdon, G., Sandefur, J., and Teal, F. (2005). Patterns of Labor Demand in Sub-Saharan Africa: A Review Paper. *GPRG, Centre for the Study of African Economies, University of Oxford.*

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, *91*(1), pp. 74-89.

Krstić, G., and Sanfey, P. (2011). Earnings inequality and the informal economy.*Economics of transition*, *19*(1), pp. 179-199.

Kucera, D., and Roncolato, L. (2008). Informal employment: Two contested policy issues. *International Labour Review*, *147*(4), pp. 321-348.

Lewis, W.A. (1954). *Economic development with unlimited supplies of labour.* Manchester School of Economic and Social Studies, 22(2), pp. 139–91.

Ligthelm, A. A. (2008). A targeted approach to informal business development: the entrepreneurial route. *Development Southern Africa*, *25*(4), pp.367-382.

Luttmer, E. (2005) . Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics.* 120(3). pp. 963-1002.

Magruder, J. R. (2012). High unemployment yet few small firms: The role of centralized bargaining in South Africa. *American Economic Journal: Applied Economics*, pp.138-166.

Maloney, W. F. (1999). Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *The World Bank Economic Review*, *13*(2), 275-302.

— (2004). Informality Revisited. *World Development*, 32(7), pp. 1159-1178.

Marcouiller, D., de Castilla, V. R., and Woodruff, C. (1997). Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic development and cultural change*, pp. 367-392.

Markussen, T., Fibæk, M., Tarp, F., and Tuan, N. D. A. (2014). *The happy farmer: Self-employment and subjective well-being in rural Vietnam* (No. 2014/108). WIDER Working Paper.

May, J. and Carter M (2009). *Agriculture: analysis of the NIDS Wave 1 dataset*. Discussion paper no. 6, National Income Dynamics Study. Cape Town:University of Cape Town.

Mbeki. T. (2003) *Speech to the National Council of Provinces*, August 2003. Available at: http://www.sarpn.org.za/documents/d0000830/index.php.

Philip, K. (2009).*Second economy strategy: addressing inequality and economic marginalisation.* A strategic framework document for a Presidency TIPS Online www.tips.org.za

Mlatsheni, C and Leibbrandt, M. (2001). *The Role of Education and Fertility in the Participation and Employment of African Women in South Africa*. DPRU Working Papers No 01/54, September 2001.

Neves, D., Aliber, Mogaladi, J., and du Toit, Andries (2011). Research Report to Programme to Support Pro-Poor Policy Development in South Africa (PSPPD). *Understanding Informal Self-Employment: A qualitative-quantitative integrated study.*

Nguyen, H. C., Nordman, C. J., and Roubaud, F. (2013). Who suffers the penalty?: A panel data analysis of earnings gaps in Vietnam. *Journal of development studies*, *49*(12), pp. 1694-1710.

Nordman, C. J., Rakotomanana, F., and Roubaud, F. (2012). *Informal versus Formal: A Panel Data Analysis of Earnings Gaps in Madagascar.*CSAE Conference, March.

Perry, G. E.; Maloney, W. F.; Arias, O.; Fajnzylber, P.; Mason, A. D.; Saavedra-Chanduvi, J.(2007). *Informality: Exit and exclusion*. Washington, DC, World Bank.

Posel, D. (2012). *Self-Assessed Wellbeing: Analysis of the NIDS Wave 1 and 2 Datasets*. Working Paper 79, SALDRU, University of Cape Town.

Posel, D. and Heintz. J. (2008). Revisiting informal employment and segmentation in the South African labour market.*South African Journal of Economics*,*76*(1), pp.26-44.

Pradhan, M., and Van Soest, A. (1995). Formal and informal sector employment in urban areas of Bolivia. *Labour economics*, *2*(3), pp. 275-297.

Pratap, S., & Quintin, E. (2006). Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review*, *50*(7), pp. 1817-1841.

Rand, J., and Torm, N. (2012). The informal sector wage gap among Vietnamese micro-firms. *Journal of the Asia Pacific Economy*, *17*(4), pp. 560-577.

Razafindrakoto, M., Roubaud, F., and Wachsberger, J-M (2012). Travailler dans le secteur

informel : choix ou contrainte ? Une analyse de la satisfaction dans l'emploi au Vietnam. In Cling, J.-P., Lagrée, S., Razafindrakoto, M., Roubaud, F. (eds.), L'économie informelle dans les pays en développement, Paris: AFD Editions, Chapter 1.2, pp. 47-66

Ruffer, T., and Knight, J. (2007). *Informal Sector Labour Markets in Developing Countries.* University of Oxford.

Saavedra, J. and Chong, A. (1999). Structural Reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru. *The Journal of Developments Studies*, Vol.35, N.4, 95-116

Skinner, C. (2005). *Constraints to growth and employment in Durban: Evidence from the informal economy*. School of Development Studies, University of KwaZulu-Natal.

Standing, G., Sender, J., and Weeks, J. (1996). *Restructuring the labour market: the South African challenge*. International Labour Organization.

Stats SA (Statistics South Africa) Quarterly Labour Force Survey, Quarter 1, 2015. Pretoria: Stats SA.

Tannuri-Pianto, M., and Pianto, D. (2002). Informal employment in Brazil-a choice at the top and segmentation at the bottom: a quantile regression approach.*Anais do XXIV Encontro Brasileiro de Econometria*, *2*.

Tansel, A. (2000). *Wage earners, self-employment and gender in the informal sector in Turkey.* Policy Research Report on Gender and Development, working paper series, (24)

Aysit, T., and Elif Oznur, K. A. N. (2012). The Formal/Informal Employment Earnings Gap: Evidence from Turkey. *Ankara, Middle East Technical University/Cankaya University, May*, pp. 1-45.

Unni, J. (2005). Wages and incomes in formal and informal sectors in India.*Indian Journal of Labour Economics*, *48*(2), pp. 311-17.

Unni, J., and Rani, U. (2002). *Social protection for informal workers: insecurities, instruments and institutional mechanisms*. International Labour Office.

Valodia, I. (2001). Economic policy and women's informal work in South Africa. *Development and Change*. 32. pp. 871-892.

Valodia, I., Lebani, L. and C. Skinner (2005).Low-waged and informal employment in South Africa. *Review of Labour Markets in South Africa*, Human Sciences Research Council, Pretoria, SouthAfrica.

Valodia, I. (2007). *The Informal Economy in South Africa: Issues, Debates and Policies: Reflections after an Exposure Dialogue Programme with Informal Workers in Durban, South Africa*, Research Report, No. 75, March 2007.

Valodia, I., and Devey, R. (2010). Formal-informal economy linkages: What implications for poverty in South Africa?. *Law, democracy & development*,*14*(1).

Verick, S. (2010). *Unravelling the impact of the global financial crisis on the South African labour market.* Employment Working Paper, No. 48. Geneva:ILO

Wills, G. (2009). South Africa's informal economy: a statistical profile. *Women in Informal Employment Organizing and Globalizing (WIEGO) Working Paper (Urban Policies) No. 6*.

Xaba, J., Horn, P., Motala, S., & Singh, A. (2002). *Informal Sector in Sub-Saharan Africa* (No. 355190). International Labour Organization.

Yu, D. (2010). *Defining and measuring informal employment in South Africa: A review of recent approaches* . Stellenbosch Economic Working Papers. No. 09/2010. University of Stellenbosch.

Zuo, H. (2013). The Earnings Differential between Formal and Informal Employees in Urban China. *Contemporary Economics*, *7*(3), pp. 115-124.

APPENDIX

TABLE A.1 Mean earnings regressions for all workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) Fixed Effects	(5) Fixed Effects
Informal worker	-1.05***	-0.56***		-0.20***	
	(0.03)	(0.03)		(0.03)	
Informal wage worker			0.24***		0.18***
			(0.05)		(0.05)
Formal self-employed			0.32***		0.09
			(0.09)		(0.09)
Formal wage-worker			0.69***		0.30***
			(0.05)		(0.04)
Age		0.04***	0.04***	0.14***	0.14***
		(0.01)	(0.01)	(0.03)	(0.03)
Age squared		-0.00***	-0.00***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)	(0.00)
Male		0.35***	0.38***		
		(0.02)	(0.02)		
Married		0.14***	0.14***		
		(0.02)	(0.02)		
Children		-0.01**	-0.01**	-0.01	-0.01
		(0.00)	(0.00)	(0.01)	(0.01)
Rural		-0.13***	-0.14***		
		(0.02)	(0.03)		
White		0.59***	0.61***		
		(0.05)	(0.05)		
Coloured		0.02	0.02		
		(0.03)	(0.03)		
Asian or Indian		0.36***	0.37***		
		(0.09)	(0.08)		
Primary education		0.05	0.05		
		(0.04)	(0.04)		
Secondary education		0.32***	0.32***		
		(0.04)	(0.04)		
Tertiary education		0.69***	0.68***		
		(0.04)	(0.04)		
Literate		0.32***	0.34***		
		(0.05)	(0.05)		
Computer literate		0.28***	0.29***		
		(0.03)	(0.03)		
Elementary (wage)		-0.38***	-0.44***	-0.06	-0.10**
		(0.03)	(0.03)	(0.04)	(0.04)
Semi-skilled (wage)		-0.15***	-0.20***	0.01	-0.03
		(0.03)	(0.03)	(0.03)	(0.03)
Elementary (SE)		-0.72***	-0.59***	-0.32***	-0.26***
		(0.11)	(0.11)	(0.08)	(0.08)
Semi-skilled (SE)		-0.57***	-0.45***	-0.18***	-0.12**
		(0.06)	(0.07)	(0.05)	(0.06)
Year dummy 2010	0.05**	0.08***	0.06***	0.03	0.02
	(0.02)	(0.02)	(0.02)	(0.07)	(0.07)
Year dummy 2012	0.14***	0.16***	0.15***	0.05	0.07
	(0.02)	(0.02)	(0.02)	(0.11)	(0.11)
Constant	8.02***	6.10***	5.47***	4.07***	3.89***
	(0.02)	(0.16)	(0.17)	(1.04)	(1.03)
Observations	6,359	6,359	6,359	6,359	6,359
R-squared	0.21	0.51	0.51	0.83	0.83

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Notes: Omitted education variable is "no education or did not complete primary", omitted race variable is "African", and omitted job skill level is "skilled".

TABLE A.2 Mean earnings regressions for female workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) Fixed Effects	(5) Fixed Effects
Informal worker	-1.14***	-0.60***		-0.21***	
	(0.04)	(0.04)		(0.04)	
Informal wage worker			0.44***		0.32***
			(0.08)		(0.07)
Formal self-employed			0.37**		0.05
			(0.15)		(0.14)
Formal wage worker			0.90***		0.44***
			(0.08)		(0.06)
Age		0.04***	0.03***	0.12***	0.11***
		(0.01)	(0.01)	(0.04)	(0.04)
Age squared		-0.00*	-0.00*	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)	(0.00)
Married		0.07**	0.08**		
		(0.03)	(0.03)		
Children		-0.01*	-0.01*	-0.01	-0.02
		(0.01)	(0.01)	(0.01)	(0.01)
Rural		-0.13***	-0.12***		
		(0.03)	(0.03)		
White		0.48***	0.50***		
		(0.07)	(0.07)		
Coloured		0.01	0.02		
		(0.04)	(0.04)		
Asian or Indian		0.30**	0.28**		
		(0.13)	(0.12)		
Primary education		0.07	0.08		
		(0.05)	(0.05)		
Secondary education		0.37***	0.39***		
		(0.06)	(0.06)		
Tertiary education		0.78***	0.75***		
		(0.06)	(0.06)		
Literate		0.36***	0.36***		
		(0.07)	(0.07)		
Computer literate		0.27***	0.27***		
		(0.04)	(0.04)		
Elementary (wage)		-0.36***	-0.48***	-0.06	-0.14**
		(0.04)	(0.04)	(0.06)	(0.06)
Semi-skilled (wage)		-0.16***	-0.25***	0.02	-0.05
		(0.04)	(0.04)	(0.04)	(0.05)
Elementary (SE)		-0.74***	-0.52***	-0.37***	-0.24*
		(0.14)	(0.15)	(0.12)	(0.12)
Semi-skilled (SE)		-0.68***	-0.43***	-0.18**	-0.07
		(0.09)	(0.10)	(0.08)	(0.08)
Year dummy 2010	0.06*	0.07***	0.05*	0.08	0.08
	(0.03)	(0.03)	(0.03)	(0.09)	(0.09)
Year dummy 2012	0.16***	0.15***	0.15***	0.16	0.17
	(0.03)	(0.03)	(0.03)	(0.15)	(0.15)
Constant	7.90***	6.14***	5.39***	4.54***	4.45***
	(0.03)	(0.24)	(0.25)	(1.39)	(1.38)
Observations	3,313	3,313	3,313	3,313	3,313
R-squared	0.26	0.54	0.54	0.85	0.85

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE A.3 Mean earnings regressions for male workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled OLS	(2) Pooled OLS	(3) Pooled OLS	(4) Fixed Effects	(5) Fixed Effects
Informal worker	-0.85***	-0.50***		-0.20***	
	(0.04)	(0.04)		(0.04)	
Informal wage worker			0.05		0.07
			(0.07)		(0.07)
Formal self-employed			0.20*		0.08
			(0.12)		(0.11)
Formal wage worker			0.48***		0.21***
			(0.06)		(0.06)
Age		0.05***	0.05***	0.16***	0.16***
		(0.01)	(0.01)	(0.05)	(0.05)
Age squared		-0.00***	-0.00***	-0.00***	-0.00***
		(0.00)	(0.00)	(0.00)	(0.00)
Married		0.23***	0.23***		
		(0.04)	(0.04)		
Children		-0.01**	-0.01*	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)
Rural		-0.13***	-0.14***		
		(0.04)	(0.04)		
White		0.71***	0.73***		
		(0.06)	(0.06)		
Coloured		0.03	0.03		
		(0.04)	(0.04)		
Asian Indian		0.40***	0.42***		
		(0.11)	(0.11)		
Primary education		0.03	0.02		
		(0.06)	(0.06)		
Secondary education		0.25***	0.25***		
		(0.06)	(0.06)		
Tertiary education		0.56***	0.56***		
		(0.06)	(0.06)		
Literate		0.25***	0.27***		
		(0.08)	(0.08)		
Computer literate		0.31***	0.32***		
		(0.04)	(0.04)		
Elementary (wage)		-0.36***	-0.38***	-0.06	-0.08
		(0.05)	(0.05)	(0.06)	(0.06)
Semi-skilled (wage)		-0.12***	-0.14***	-0.00	-0.02
		(0.04)	(0.04)	(0.04)	(0.05)
Elementary (SE)		-0.67***	-0.61***	-0.26**	-0.25**
		(0.16)	(0.16)	(0.11)	(0.12)
Semi-skilled (SE)		-0.43***	-0.41***	-0.18**	-0.16*
		(0.08)	(0.08)	(0.08)	(0.08)
Year dummy 2010	0.05	0.10***	0.07**	-0.04	-0.05
	(0.03)	(0.03)	(0.03)	(0.10)	(0.10)
Year dummy 2012	0.13***	0.17***	0.16***	-0.08	-0.07
	(0.03)	(0.03)	(0.03)	(0.17)	(0.17)
Constant	8.13***	6.40***	5.93***	3.38**	3.20**
	(0.03)	(0.23)	(0.24)	(1.56)	(1.56)
Observations	3,046	3,046	3,046	3,046	3,046
R-squared	0.14	0.46	0.45	0.80	0.80

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE A.4 Pooled quantile regressions for all workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled 0.1	(2) Pooled 0.25	(3) Pooled 0.5	(4) Pooled 0.75	(5) Pooled 0.9	(6) Pooled 0.1	(7) Pooled 0.25	(8) Pooled 0.5	(9) Pooled 0.75	(10) Pooled 0.9
Informal worker	-0.65***	-0.52***	0.51***	-0.56***	-0.48***					
	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)					
Formal self-employed						0.33	0.11	0.26*	0.39***	0.43***
						(0.20)	(0.11)	(0.15)	(0.07)	(0.04)
Informal wage worker						0.76***	0.56***	0.15***	-0.04	-0.14***
						(0.09)	(0.06)	(0.05)	(0.05)	(0.04)
Formal wage worker						1.17***	0.96***	0.59***	0.46***	0.31***
						(0.07)	(0.06)	(0.05)	(0.05)	(0.03)
Age	0.03***	0.04***	0.04***	0.04***	0.03***	0.04***	0.04***	0.04***	0.05***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age squared	-0.00	-0.00***	0.00***	-0.00***	-0.00**	-0.00**	0.00***	0.00***	-0.00***	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	0.36***	0.33***	0.33***	0.36***	0.36***	0.37***	0.35***	0.35***	0.38***	0.37***
	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Married	0.23***	0.15***	0.13***	0.10***	0.14***	0.23***	0.16***	0.13***	0.11***	0.14***
	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)
Children	-0.00	-0.00	0.01***	-0.01**	-0.02***	-0.01**	-0.00	0.01***	-0.01**	-0.02***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)
Rural	-0.22***	-0.20***	0.13***	-0.09***	-0.07**	0.23***	0.21***	0.14***	-0.11***	-0.08***
	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
White	0.60***	0.65***	0.60***	0.59***	0.53***	0.55***	0.62***	0.64***	0.59***	0.51***
	(0.08)	(0.04)	(0.04)	(0.04)	(0.04)	(0.09)	(0.04)	(0.04)	(0.04)	(0.03)
Coloured	0.04	0.03	0.02	-0.01	0.02	0.04	0.02	0.02	-0.01	0.02
	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)	(0.02)	(0.02)	(0.02)
Asian or Indian	0.36**	0.39***	0.31***	0.34***	0.41***	0.48***	0.40***	0.32***	0.33***	0.43***
	(0.15)	(0.11)	(0.06)	(0.08)	(0.12)	(0.13)	(0.07)	(0.05)	(0.08)	(0.11)
Primary education	0.15**	0.11**	0.07**	0.02	-0.05	0.14**	0.11**	0.07**	-0.02	-0.05
	(0.06)	(0.05)	(0.03)	(0.04)	(0.05)	(0.07)	(0.04)	(0.03)	(0.03)	(0.04)
Secondary education	0.46***	0.35***	0.31***	0.26***	0.20***	0.42***	0.34***	0.32***	0.24***	0.21***
	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)
Tertiary education	0.83***	0.79***	0.76***	0.61***	0.44***	0.79***	0.77***	0.75***	0.58***	0.47***
	(0.07)	(0.06)	(0.04)	(0.04)	(0.05)	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)
Literate	0.18***	0.30***	0.38***	0.37***	0.33***	0.19**	0.29***	0.38***	0.42***	0.35***
	(0.06)	(0.07)	(0.05)	(0.05)	(0.08)	(0.08)	(0.05)	(0.05)	(0.06)	(0.06)
Computer literate	0.29***	0.25***	0.24***	0.31***	0.36***	0.31***	0.26***	0.24***	0.32***	0.37***
	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Elementary (wage)	-0.26***	-0.35***	0.39***	-0.44***	-0.54***	0.48***	0.46***	0.43***	-0.46***	-0.49***
	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)
Semi-skilled (wage)	0.01	-0.12***	0.19***	-0.24***	-0.33***	0.20***	0.19***	0.20***	-0.25***	-0.28***
	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)
Elementary (SE)	-0.84***	-1.05***	0.95***	-0.54**	-0.21	-0.63	0.51***	0.78***	-0.59***	-0.33
	(0.10)	(0.13)	(0.16)	(0.22)	(0.24)	(0.42)	(0.14)	(0.12)	(0.17)	(0.22)
Semi-skilled (SE)	-0.66***	-0.69***	0.71***	-0.43***	-0.32***	-0.21*	0.26***	0.65***	-0.46***	-0.42***
	(0.08)	(0.06)	(0.05)	(0.10)	(0.05)	(0.12)	(0.05)	(0.06)	(0.10)	(0.05)
Year dummy 2010	0.02	0.07**	0.10***	0.12***	0.14***	-0.05	0.07***	0.08***	0.10***	0.11***
	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)	(0.02)	(0.02)	(0.03)
Year dummy 2012	0.14***	0.14***	0.15***	0.16***	0.18***	0.14***	0.14***	0.14***	0.15***	0.17***
	(0.05)	(0.03)	(0.02)	(0.02)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
Observations	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE A.5 Pooled Quantile Regressions for Female Workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled 0.1	(2) Pooled 0.25	(3) Pooled 0.5	(4) Pooled 0.75	(5) Pooled 0.9	(6) Pooled 0.1	(7) Pooled 0.25	(8) Pooled 0.5	(9) Pooled 0.75	(10) Pooled 0.9
Informal worker	-0.68***	-0.54***	-0.53***	-0.61***	-0.58***					
	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)					
Informal wage worker						0.91***	0.89***	0.40***	0.02	-0.02
						(0.06)	(0.09)	(0.10)	(0.07)	(0.08)
Formal self-employed						0.49*	0.24**	0.17	0.41***	0.42***
						(0.29)	(0.12)	(0.16)	(0.12)	(0.15)
Formal wage worker						1.31***	1.28***	0.85***	0.57***	0.49***
						(0.05)	(0.09)	(0.10)	(0.07)	(0.07)
Year dummy 2010	0.02	0.03	0.10***	0.11***	0.14***	-0.08	0.04	0.09***	0.11***	0.14***
	(0.06)	(0.04)	(0.03)	(0.03)	(0.03)	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)
Year dummy 2012	0.17**	0.14***	0.15***	0.13***	0.15***	0.12**	0.12***	0.15***	0.13***	0.17***
	(0.07)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)
Observations	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Regressions also include set of control variables in Table A.4.

TABLE A.6 Pooled quantile regressions for male workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) Pooled 0.1	(2) Pooled 0.25	(3) Pooled 0.5	(4) Pooled 0.75	(5) Pooled 0.9	(6) Pooled 0.1	(7) Pooled 0.25	(8) Pooled 0.5	(9) Pooled 0.75	(10) Pooled 0.9
Informal worker	-0.62***	-0.52***	-0.47***	-0.43***	-0.37***					
	(0.06)	(0.05)	(0.03)	(0.04)	(0.04)					
Informal wage worker						0.59***	0.18*	-0.03	-0.17**	-0.16***
						(0.17)	(0.10)	(0.07)	(0.08)	(0.05)
Formal self-employed						0.24	0.01	0.19	0.22	0.36***
						(0.16)	(0.14)	(0.18)	(0.14)	(0.07)
Formal wage worker						1.09***	0.61***	0.38***	0.22***	0.21***
						(0.15)	(0.10)	(0.07)	(0.08)	(0.04)
Year dummy 2010	0.07	0.11***	0.11***	0.15***	0.13***	0.03	0.10***	0.10***	0.14***	0.11***
	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)
Year dummy 2012	0.16***	0.14***	0.16***	0.19***	0.22***	0.16***	0.16***	0.16***	0.19***	0.21***
	(0.06)	(0.04)	(0.03)	(0.04)	(0.05)	(0.05)	(0.04)	(0.03)	(0.04)	(0.05)
Observations	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Regressions also include set of control variables in Table A.4.

TABLE A.7 Fixed effects quantile regressions for all workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) FE 0.1	(2) FE 0.25	(3) FE 0.5	(4) FE 0.75	(5) FE 0.9	(6) FE 0.1	(7) FE 0.25	(8) FE 0.5	(9) FE 0.75	(10) FE 0.9
Informal worker	-0.21***	-0.20***	0.23***	-0.20***	-0.24***					
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)					
Informal wage worker						0.46***	0.33***	0.19***	0.02	-0.12
						(0.05)	(0.05)	(0.04)	(0.05)	(0.10)
Formal self-employed						-0.05	0.12	0.06	0.11	0.22**
						(0.07)	(0.13)	(0.08)	(0.07)	(0.10)
Formal wage worker						0.54***	0.42***	0.33***	0.15***	0.06
						(0.04)	(0.04)	(0.04)	(0.05)	(0.10)
Age	0.14***	0.14***	0.14***	0.13***	0.15***	0.14***	0.14***	0.14***	0.13***	0.15***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age squared	-0.00***	-0.00***	0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Male	-0.00	-0.00	0.03	0.02	-0.00	0.01	-0.01	0.03*	0.02	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Married	0.00	0.02	-0.01	0.01	-0.04	0.03	0.02	-0.01	0.00	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Children	-0.00	-0.01*	-0.01*	-0.02***	-0.02***	-0.00	-0.01**	-0.01	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Rural	-0.04	-0.00	-0.00	0.02	0.04	-0.03	0.01	-0.01	0.03	0.03
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
White	0.02	-0.00	0.02	-0.08**	-0.05	0.05	0.05	-0.01	-0.04	-0.00
	(0.04)	(0.04)	(0.03)	(0.04)	(0.06)	(0.06)	(0.04)	(0.03)	(0.04)	(0.08)
Coloured	0.03	0.03	0.01	-0.03	-0.03	0.02	0.02	0.01	-0.01	-0.02
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
Asian or Indian	-0.22*	0.04	0.04	-0.06	0.09	-0.13	0.08	0.04	-0.05	-0.04
	(0.13)	(0.04)	(0.06)	(0.05)	(0.08)	(0.14)	(0.05)	(0.06)	(0.09)	(0.21)
Primary education	0.10**	0.06*	-0.02	-0.10***	-0.14***	0.11***	0.06**	-0.02	-0.08**	-0.14**
	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)
Secondary education	0.11**	0.06	-0.03	-0.11***	-0.22***	0.12***	0.05	-0.03	-0.09**	-0.18***
	(0.05)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.03)	(0.03)	(0.04)	(0.06)
Tertiary education	0.19***	0.11***	-0.03	-0.11***	-0.17***	0.20***	0.09**	-0.02	-0.10**	-0.14**
	(0.05)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.06)
Literate	0.05	-0.03	0.01	0.02	-0.02	0.02	-0.04	0.04	0.02	0.02
	(0.05)	(0.05)	(0.06)	(0.07)	(0.05)	(0.04)	(0.05)	(0.06)	(0.05)	(0.07)
Computer literate	-0.05*	-0.02	0.03	0.04	0.01	-0.05	-0.01	0.01	0.04	-0.02
	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)
Elementary (wage)	0.06*	-0.00	-0.02	-0.07**	-0.20***	-0.03	-0.08***	-0.08***	-0.07**	-0.17***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Semi-skilled (wage)	0.10***	0.07***	0.05**	-0.01	-0.09**	0.02	0.02	0.00	-0.02	-0.07*
	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Elementary (SE)	-0.65***	-0.45**	-0.30**	-0.18	-0.15*	-0.41***	-0.39	-0.19*	-0.21	-0.25**
	(0.16)	(0.21)	(0.14)	(0.12)	(0.08)	(0.08)	(0.28)	(0.10)	(0.14)	(0.10)
Semi-skilled (SE)	-0.37***	-0.15**	-0.09*	-0.13**	-0.22***	-0.14*	-0.03	-0.03	-0.19***	-0.39***
	(0.07)	(0.06)	(0.05)	(0.06)	(0.08)	(0.07)	(0.06)	(0.04)	(0.06)	(0.06)
Year dummy 2010	0.01	0.03	0.03	0.02	0.02	-0.00	0.03	0.02	0.02	0.06**
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Year dummy 2012	0.04*	0.07***	0.06***	0.04	0.02	0.05*	0.08***	0.07***	0.06**	0.06*
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Observations	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359

Source: NIDS and author's own calculations. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE A.8 Fixed effects quantile regressions for female workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) FE 0.1	(2) FE 0.25	(3) FE 0.5	(4) FE 0.75	(5) FE 0.9	(6) FE 0.1	(7) FE 0.25	(8) FE 0.5	(9) FE 0.75	(10) FE 0.9
Informal worker	-0.15***	-0.20***	-0.23***	-0.22***	-0.25***					
	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)					
Informal wage worker						0.63***	0.54***	0.29***	0.13**	-0.03
						(0.09)	(0.06)	(0.06)	(0.07)	(0.07)
Formal self-employed						-0.22	-0.13	0.05	0.23	0.23***
						(0.09)	(0.20)	(0.09)	(0.14)	(0.07)
Formal wage worker						0.71***	0.63***	0.40***	0.25***	0.17***
						(0.08)	(0.06)	(0.06)	(0.06)	(0.06)
Year dummy 2010	0.09**	0.10***	0.08**	0.07**	0.10**	0.07*	0.09***	0.09***	0.08**	0.11**
	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Year dummy 2012	0.17***	0.19***	0.16***	0.11***	0.15***	0.18***	0.20***	0.19***	0.17***	0.18***
	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Observations	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313	3,313

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Regressions also include set of control variables in Table A.7.

TABLE A.9 Fixed effects quantile regressions for male workers
Dependent variable: Log real monthly earnings

VARIABLES	(1) FE 0.1	(2) FE 0.25	(3) FE 0.5	(4) FE 0.75	(5) FE 0.9	(6) FE 0.1	(7) FE 0.25	(8) FE 0.5	(9) FE 0.75	(10) FE 0.9
Informal worker	-0.28***	-0.23***	-0.23***	-0.19***	-0.21***					
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)					
Informal wage worker						0.23***	0.15**	0.11*	0.01	-0.07
						(0.05)	(0.07)	(0.06)	(0.08)	(0.07)
Formal self-employed						0.06	0.14	0.06	0.09	0.24**
						(0.06)	(0.15)	(0.08)	(0.14)	(0.10)
Formal wage worker						0.37***	0.29***	0.26***	0.14*	0.08
						(0.05)	(0.06)	(0.05)	(0.07)	(0.07)
Year dummy 2010	-0.05	-0.06*	-0.04	-0.07**	-0.06*	-0.05	-0.05	-0.06**	-0.07*	-0.08*
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Year dummy 2012	-0.09**	-0.10***	-0.08***	-0.06**	-0.10***	-0.06**	-0.08**	-0.07**	-0.05	-0.10**
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Observations	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046	3,046

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Regressions also include set of control variables in Table A.7.

TABLE A.10 Fixed effects quantile regressions for all workers (hourly earnings)
Dependent variable: Log real hourly earnings

VARIABLES	(1) FE 0.1	(2) FE 0.25	(3) FE 0.5	(4) FE 0.75	(5) FE 0.9	(6) FE 0.1	(7) FE 0.25	(8) FE 0.5	(9) FE 0.75	(10) FE 0.9
Informal worker	-0.20***	-0.15***	-0.10***	-0.10***	-0.12***					
	(0.04)	(0.03)	(0.02)	(0.02)	(0.04)					
Informal wage worker						0.16**	0.13**	0.05	-0.09*	-0.20***
						(0.07)	(0.06)	(0.05)	(0.05)	(0.06)
Formal self-employed						0.00	-0.03	-0.10	0.03	0.14*
						(0.07)	(0.10)	(0.09)	(0.14)	(0.08)
Formal wage worker						0.34***	0.27***	0.16***	0.05	-0.04
						(0.06)	(0.05)	(0.05)	(0.05)	(0.05)
Year dummy 2010	0.37***	0.37***	0.40***	0.41***	0.36***	0.34***	0.35***	0.39***	0.40***	0.37***
	(0.04)	(0.03)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Year dummy 2012	0.21***	0.25***	0.26***	0.24***	0.20***	0.21***	0.23***	0.27***	0.25***	0.22***
	(0.04)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)
Observations	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359	6,359

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Regressions also include set of control variables in Table A.7.

Table A.11 Pooled OLS Life Satisfaction regressions for all workers
Dependent variable: Self-perceived current life satisfaction

VARIABLES	(1) Pooled OLS	(2) Pooled OLS
Informal worker	-0.17***	
	(0.14)	
Informal wage worker		-0.10*
		(0.05)
Formal self-employed		0.23*
		(0.12)
Formal wage worker		0.10**
		(0.04)
Age	0.00*	0.00*
	(0.00)	(0.00)
Male	-0.03	-0.03
	(0.02)	(0.02)
Married	0.12***	0.12***
	(0.03)	(0.03)
Children	-0.02***	-0.02***
	(0.00)	(0.00)
Employed	0.11***	0.16***
	(0.04)	(0.04)
Not economically active	0.17***	0.17***
	(0.03)	(0.03)
White	1.15***	1.14***
	(0.06)	(0.06)
Coloured	1.33***	1.34***
	(0.04)	(0.04)
Asian or Indian	1.58***	1.57***
	(0.10)	(0.10)
Primary education	0.12***	0.13***
	(0.03)	(0.03)
Secondary education	0.11***	0.11***
	(0.04)	(0.04)
Tertiary education	0.20***	0.21***
	(0.05)	(0.05)
Health	0.46***	0.46***
	(0.02)	(0.02)
Household income	0.33***	0.33***
	(0.01)	(0.01)
Economic status (bottom third)	-0.87***	-0.87***
	(0.06)	(0.06)
Economic status (middle third)	-0.04	-0.04
	(0.06)	(0.06)
Upward mobility	0.32***	0.32***
	(0.02)	(0.02)
Year dummy 2010	-0.72***	-0.72***
	(0.03)	(0.03)
Year dummy 2012	-0.49***	-0.49***
	(0.03)	(0.03)
Constant	2.42***	2.23***
	(0.14)	(0.13)
Observations	41,538	41,538
R-squared	0.17	0.17

Source: NIDS and author's own calculations.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Notes: 1/ Health is a dummy variable set equal to 1 if respondent self-reported their health as either "excellent" or "very good". 2/ Household income refers to the log of real household income. 3/ Economic status is current perceived relative economic status, with the top third being the reference category. 4/ Upward mobility is a dummy variable set equal to 1 if the respondent expects their relative economic status to be higher in two years' time compared to today.