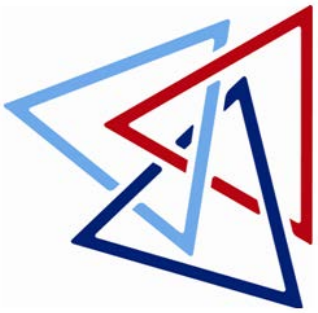

Job-seeker entry into the two-tiered informal sector in South Africa

Nwabisa Makaluza

Abstract

Despite the high open unemployment in South Africa, the informal sector remains small. This study aims to bring us closer to understanding the incentives and constraints of job-seekers who find employment in the heterogeneous informal sector. We do this by identifying two tiers in the sector (the survivalist and the growth oriented micro-enterprises) by using the data-driven k-medians clustering technique which finds natural groupings of jobs in the informal sector by using several work characteristics. Once the members of either tier of the informal sector have been successfully identified we find the properties that facilitate entry into either tier of the informal sector. Growth oriented micro-enterprise entrants are more likely to have access to financial and human capital, which form barriers to entry some job-seekers. On the other hand, survivalist workers have been induced into the informal sector to find employment when other sources of income are lacking.

The **Research Project on Employment, Income Distribution and Inclusive Growth** is based at SALDRU at the University of Cape Town and supported by the National Treasury. Views expressed in REDI3x3 Working Papers are those of the authors and are not to be attributed to any of these institutions.

Job-seeker entry into the two-tiered informal sector in South Africa

Nwabisa Makaluza
(Department of Economics, University of Stellenbosch)

1. Introduction

One of the puzzles of the South African labour market is that it has a small informal sector amid high open unemployment. The small numbers of informal sector entrants challenges the notion of a free entry segment that can absorb surplus job-seekers. This has encouraged study of the incentives and constraints that govern the decision to enter this part of the economy. But of course, the informal sector consists of workers who are engaged in a variety of activities in the unregistered portion of the economy. At the very least, it is worth following Fields (1990) in distinguishing between an easy entry (survivalist) informal sector and a (growth oriented) upper-tier informal sector. Growth-oriented micro-enterprises consists of entrepreneurs who want to take advantage of income opportunities provided by a less regulated sector, whereas survivalists are job-seekers who have been unsuccessful in finding employment in the formal sector and who are looking for a way to cope with poverty. This study will attempt to contribute to our understanding of the South African informal sector by first identifying the members of growth oriented and survivalist tiers, and then proceed to determine whether job-seekers are restricted from joining either informal sector segment due to high entry barriers, or whether individuals try to avoid these jobs until they are faced with the combination of household responsibilities and a lack of other sources of income.

The two tiers within the informal sector are identified by using a data-driven k-medians clustering technique. This approach combines the information about several job characteristics and an automated algorithm to find natural groupings of workers who share very similar work environments, without the need to specify an arbitrary wage cut-off to distinguish survivalist enterprises from growth oriented micro-enterprise workers. Thereafter, the relationship between individual and household characteristics and the probability of being in either informal sector tier is explored in order to determine the type of job seekers who are entrants in either tier. This relationship is modelled using multinomial logit, conditional logit, ordinary least squares and fixed effects estimators, in order to address various confounding factors that may otherwise bias our estimates.

The results show that most of the South African informal sector consists of survivalist enterprises who entered the informal sector as an employment opportunity of last resort. They work in harsh working conditions for a low pay and with poor prospects for upward mobility. Entry into this segment is usually associated with the responsibility of providing for dependent household members and a lack of other household income sources. A smaller portion of the informal sector consists of growth oriented micro-enterprise workers who have the skills and financial means to overcome barriers that prevent entry into this segment. They earn a higher income and do jobs that are a closer to

those found in the formal sector. Another finding is that women are less likely to enter the growth oriented micro-enterprises than men.

The study starts by reviewing the international and South African literature on the informal sector in section 2. This is followed by a description of the panel data used in the empirical analysis in section 3. The statistical technique that is used to sort informal sector workers into their respective tiers is discussed in section 4 before the results are presented. The empirical analysis in section 5 begins with a description of the subgroups of the jobs in the economy that have been identified in the cluster analysis. These clusters are then combined into the two tiers of the informal sector, before analysing the determinants of entering either of these segments. Section 6 concludes.

2. Literature review

2.1. Unemployment in South Africa

After the political transition in South Africa, labour market participation grew at a faster pace than the economy could absorb. The unemployment rate (narrowly defined) stabilised at above 20 per cent with approximately 10 per cent of additional discouraged work seekers. Structural changes in the economy such as changes in production technology and shifts to less labour intensive sectors contributed to the inability of employment growth to keep up with the increased labour market entrants (Bhorat, 2004). In this time, South Africa also experienced a period of feminisation of the labour force which was driven by supply-side push factors (Casale et al. 2002). Women moved to entrepreneurial activities in the informal sector so the feminisation of the labour market was associated with an increase in low paid employment and female unemployment.

Job search has been less successful for women, black people, and the youth. Job-seekers in these groups also experience longer periods of unemployment on average. The majority of the unemployed have never had a job and many of those who have worked before have experienced unemployment for longer than a year (Banerjee et al, 2008). A number of elements contribute to the low success rates of job seekers, one of them being the high search costs that are attached to living in areas that are far from business centres. These factors interact with other structural elements, such as skills inflation, in the economy to produce an unemployment rate that is unlikely to change without policy intervention (Banerjee et al, 2008).

Another possible explanation of high unemployment is that the reservation wages are too high. Although the evidence of the role that reservation wages play on unemployment in South Africa has been mixed. Kingdon and Knight (2004) used a measure of the self-reported reservation wage from the 1993 PSLSD SALDRU dataset. The reservation wages in this study were much higher than the remuneration that respondents could expect to receive in the job market. However, the authors claim that the reservation wages captured the respondent's idea of a fair wage from a bargaining perspective and hence do not interpret their results as indicative of the fact that reservation wages

constrain employment. Rankin and Roberts (2011) found that the youth had reservation wages that were similar to the remuneration that they could expect from larger firms which are above the smaller firms' wages. This could deter the young job seeker from accepting a job from the smaller firms. On the other hand, Nattrass and Walker (2005) found that the reservation wages of working class Khayelitsha and Mitchell's Plein residents (Cape Town) were below their predicted wages. This means that the residents in this area were willing to work for lower wages than what they could expect to be paid in the labour market. The varied results from the studies leaves us with little clarity on whether high reservation wages are the cause for the high unemployment.

Since unemployed job-seekers cannot depend on their earnings to survive, there must be some form of non-wage income that they can rely on. Non-wage income increases job seekers' ability to sustain themselves during the period of unemployment. In the international literature, unemployment insurance is often used as an important source of non-wage income for the (typically small) group of unemployed job-seekers. In South Africa, the UIF is awarded for a limited period when a person is out of work. A person may claim the unemployment insurance fund for three months after they lose their job. The UIF's presence and subsequent absence can be used to measure the effect of non-wage income on search intensity and willingness to accept a job offer. However, using this fund as the measure for non-wage income is not suitable in the South African labour market because of the lack of coverage; less than ten per cent of strictly unemployed people receive the grant (Leibbrandt et al. 2010). The absence of a grant specifically aimed at assisting the unemployed can necessitate job seekers to use other sources of non-wage income.

Social protection programs such as the unconditional cash transfer provided by the state through means tested grants provide support to the vulnerable and have also been used as an exogenous source of non-wage income in econometric analysis (Klasen and Woolard 2008; Duflo 2003; Van der Berg et al. 2010; Van der Berg and Bredenkamp 2002). Two of the largest grants in South Africa in terms of coverage and value are the child support grant and the old age pension. The impact that these grants have had on members of the household has been positive. For example, Coetzee (2013) found that the recipients of the child support grant have better school outcomes than comparable children who do not receive this grant. The old age pension has been shown to produce favourable welfare outcomes to poor rural households especially when the beneficiary is a woman (Duflo 2003). These grants are an important source of non-wage income for poor South Africans and have also affected how households with the elderly are formed.

One viable strategy to coping with unemployment is to live with someone who receives stable income (Klasen and Woolard 2008). This constant income could be wages from employed household members, remittances from non-household members, or social grants from eligible beneficiaries such as the old age pension or the child support grant. In fact, Bertrand et al. (2003) found a negative relationship between old age pension eligibility and labour supply decisions of prime aged adults. The marked increase in household income due to an elderly member crossing the eligibility threshold was associated with a decrease in hours worked of the employed. Posel et al. (2006)

extended Bertrand et al.'s (2003) research to include the effect of possible labour migration of household members due to the increase in total income. They find that the income from the old age pension helps to relieve the constraints of female labour migration. Neo-classical job search theory postulates that non-wage income makes it possible for an unemployed person to remain so. However, Posel et al. (2006) show that if costs are a binding constraint to job-search then non-wage income may be used to relax this restriction.

The discussion of unemployment extends beyond measures of material well-being to measures of subjective well-being. Kingdon and Knight (2004) found that people who live in households with higher rates of unemployment had lower levels of life satisfaction. Their finding is supported by the literature on subjective well-being that has found that the onset of unemployment lowers the levels of happiness (Clark 2003; Layard 2005; Lucas et al. 2004). This suggests that unemployment is involuntary because nobody with the ability to move out of unemployment would choose this unsatisfactory outcome (Kingdon and Knight 2004). They use this result as well as some evidence on the challenges that are faced in the informal sector to conclude that there are barriers that restrict entry into this form of employment.

2.2. The informal sector

The predominant view in the early literature was that the informal sector is a single, free entry sector (Moser 1978; Fields 1990). Having failed to find employment in the formal private or public sector, the job-seeker would have the option to move from unemployment to underemployment in the informal sector. The role of the informal sector was not only to provide employment for residual labour market participants, but also to act as a transition mechanism into the formal sector (Banerjee et al. 2008). Under these assumptions, the size of the informal sector would diminish as a country develops more formal enterprises. This transition did not occur as hypothesised in developing countries. The informal sector grew in developing countries and it became clear that there were barriers that prevented the entry of informal sector workers into the formal public and private sectors. This solidified the application of dual labour market theory in order to explain the existence of this sector.

The dual labour market theory is based on the premise that it is necessary to dichotomise the labour market into a high wage primary sector with formal labour regulations and a low wage secondary sector with informal hiring practices (Reich et al. 1973; Dickens and Lang 1988). The two segments have different wage structures, and as a result earnings depend on whether the worker is employed in a primary firm (formal sector) or a secondary firm (informal sector). The wage structures of firms in the primary segment depend on various institutional stakeholders, such as trade unions or minimum wages set by the state. Labour codes such as fringe benefits and severance pay, which are not available in the secondary segment, are some of the additional factors that inflate the remuneration earned in the primary segment therefore workers from different segments with the same levels of human capital earn different wages. Job-seekers accept a wage penalty to working in the

informal sector which remains even when skill and unobserved heterogeneity have been controlled for which indicates underemployment (Badaoui et al. 2008). Furthermore, the jobs offered in formal firms are scarce due to skills biased technological change and mobility between sectors is limited due to barriers to entry (Bhorat and Hodge 1999). The workers in informal-sector enterprises do not have the job security of formal sector workers; they are more likely to move into and out of labour force participation and unemployment.

The disparate activities of an informal sector – where labour can be underemployed or can be geared towards growth-oriented profitable businesses – caused scholars such as Fields (1990) to question the validity of the homogeneous sector assumption. In reality, the informal sector is a mixture of underemployed labourers in survivalist enterprises and entrepreneurs in growth-oriented enterprises (Rogerson 2000; Lund 1998; Fields 1990). The recognition of this heterogeneity resulted in a new way of thinking about the informal sector and analyses that neglect this aspect may result in misleading evidence about the incentives and constraints of informal sector entry. However, being able to identify the members of the different segments in the informal sector is an important component of analysis. Before any quantitative analysis is carried out, we need to understand the characteristics of growth-oriented and survivalist micro-enterprise workers.

Survivalist enterprise workers are usually impoverished, and unable to find stable employment in the formal sector. In order to cope with poverty they seek employment with low income and low capital requirements that often offer very few prospects of expansion or upward mobility (Rogerson 2000). There are a large number of entrants into this sector due to the relative ease of access. So the environments that survivalist enterprises compete in are congested markets that trade highly saturated goods and services, for example street vendors selling fruit. One of the consequences of the high competition is that survivalists may find it difficult to upwardly adjust prices in response to increases of costs from their suppliers (Mkhize et al. 2013).

Survivalists have to operate under poor working conditions. A study of street vendors in Durban by Mkhize et al. (2013) found that these persons trade in inadequate business spaces where they are exposed to weather elements – which often leads to damaged stock and negative health effects – and usually have poor access to toilets or rubbish removal. These difficulties are exacerbated if trade takes place in an area where a vending permit is required because this can lead to problems with officials. Failure to produce a permit may result in fines or the goods being confiscated. Sometimes the stocks are returned damaged or are not returned at all.

Entrepreneurs forced to start businesses because of desperation have high risks of failure (Caliendo and Kritikos 2009). The enterprises, started by these entrepreneurs, that do not fail generate a small amount of income. Because the survivalist informal sector acts as the employer of last resort, an exit from this sector typically leads to unemployment or inactivity in the labour market. Any profit that is earned by the owners of these enterprises contributes to the provision of their basic needs as well as that of the household that they belong to. The entrepreneur reinvests insignificant amounts

of capital so the enterprise has little prospect of profit induced growth (Santarelli and Vivarelli 2006) and can do very little to absorb unemployed job-seekers.

Growth-oriented micro-enterprises can emerge because of informalisation, from market opportunities that are not available in the formal sector, or because regulations are costly and are not uniformly enforced. Informalisation occurs when activities that were conducted in the formal sector are outsourced to unregistered businesses. The income generated from these activities is comparable to income from the formal sector (Blunch et al. 2001), which makes these growth oriented micro-enterprises an attractive segment for employment and growth. When unemployed job-seekers enter this part of the informal sector, it would typically lead to an improvement in well-being.

The informal sector can reach a consumer base outside the economic hubs of urban areas which gives it the potential to allocate resources to the poorest parts of the population. Businesses in this sector are in the spatial and economic position to provide goods and services that are accessible to low income earners (Weeks 1975). Brand (1986) found that in Zimbabwe a lot of the skills that were used in informal sector activities were self-taught by observing people work. If the sector grows, it can also provide opportunities for skills transfer in a labour intensive portion of the economy.

South Africa has a history of restrictive laws and practices that made it difficult to work in the informal sector (Kingdon and Knight 2004). Apartheid spatial planning removed marginalised people away from the economic hubs to the outskirts of urban areas (Rogerson 2000). As a result, transport costs have had an important effect on seeking and providing labour. Informal sector enterprises, such as spaza shops (retail outlets) and taverns, have developed within the township economy. These types of traders usually purchase their products from the formal sector and sell them at a mark-up. The goods sold here are more expensive than the formal sector prices but the proximity to the consumers encourages sales which makes the trade a viable employment option.

Growth micro-enterprises have greater physical and human capital requirements and are able to generate higher remuneration for the entrepreneur. These prerequisites limit the ability of unemployed people to start such businesses. Occupations such as those of vehicle mechanics, tailors, and builders depend on the availability of workers with industry-specific skills. These persons can move between the formal and informal sector with more ease than survivalist enterprise workers.

There are various ways to identify workers in the different segments within the informal sector. Several studies use a specific variable to identify the differences between the firms within the informal sector. For example, Grimm et al. (2012) use accumulated capital in a model that sorts firms into either survivalist enterprises (the lower tier), growth-oriented upper-tier enterprises, as well as a group of ‘constrained gazelles’ who have the potential for high returns given their observable characteristics but have not reached the upper tier. Günther and Launov (2012) use earnings to differentiate between the portion of informal employment that arise because of income opportuni-

ties or the type that serves as an employer of last resort. A limitation of focusing on a single feature of an occupation to categorise subgroups within the informal sector is that it ignores other aspects, such as work conditions, that also characterise the formality of the job.

The empirical analysis described in section 4 is of a data-driven clustering technique that uses various job attributes to identify the survivalist and growth-oriented micro-enterprises in the informal sector. The k-medians cluster analysis is an exploratory technique that partitions data by maximising similarity within groups and minimising similarity between groups (Johnson and Wichern 2007). The data-driven nature of this method decreases the need for ad hoc assumptions about the number of subgroups and the fraction of informal workers in each subgroup. Results from this technique are used to identify the two tiers in the South African informal sector.

3. Data description

Statistics South Africa's Labour Force Survey (LFS) is a rotating panel dataset that was collected biannually from 2000 until 2007. The repeated observations (individuals and households) that were accumulated between September 2001 and March 2004 were used to construct a panel dataset. This dataset is used throughout this study.

The LFS was formed by using a two-stage sampling procedure (Statistics South Africa, 2001). In the first stage, the 1999 Master Sample was used to select primary sampling units (PSUs) from the 1996 Census list of enumerator areas (EAs). This Master Sample, which was stratified into nine provinces each with distinct urban and rural areas, did not change throughout the LFS series (Kerr and Wittenberg, 2015). Thereafter, ten dwelling units were sampled from each PSU in the second stage. Each of these households had to complete a module that contained information about the employment status and sector of each working aged adult in the household.

The StatsSA classification of informal sector activities in the LFS was based on whether the individual worked in a business that is not registered for VAT[1]. This information was obtained from the respondent who had to give description of their occupation and firm of employment during the preceding week. Respondents were asked whether the business they worked for was registered for VAT and to identify the sector (formal or informal) they were employed in. The information obtained from these questions was used to classify the respondents into the sector (formal or informal) that they work in. The self-reported nature of this classification is therefore more likely to be an indication of the respondent's perception rather than the actual employment sector (Heintz and Posel 2008).

The classification used in the LFS of the informal sector is focused on the enterprise and not on the type of work carried out in the business. This corresponds to the guidelines for defining the infor-

1 In the more recent labour force surveys (Quarterly Labour Force Survey) employees are identified as informal sector employees if they work in firms that have less than five workers and if no income tax is deducted from their wages.

mal sector that were set out in the 15th International Conference of Labour Statisticians (2000) which was meant to increase the cross-country comparability of informal sector definitions. A business is a part of the informal sector if it is not registered and/or should not employ more than a certain number of workers. The enterprise and its owner(s) cannot be separate legal entities, and the production process should be of non-agricultural activities. At least some of the goods and services that have been produced have to be traded and should not be produced solely for the owner's consumption (Hussmanns 2004). As a result of these guidelines, the informal sector does not include subsistence agriculture or domestic work

The focus of this study is on identifying the members of survivalist and growth oriented micro-enterprises, and then finding the reasons behind the entry into either tier by identifying the properties of out of work job-seekers who enter the informal sector within six months. The term 'out of work' includes the searching unemployed, discouraged work-seekers, and anyone else who is not economically active (NEA) for reasons other than being enrolled in an education institution. In order to be classified as the searching unemployed the respondent must satisfy three conditions. First, the person should not have worked for seven days before the survey interview. Secondly, the individual should want a job and be available to start working within two weeks of the interview. Lastly, the respondent must have conducted active job search or taken steps to start their own business in the four weeks before the interview (Statistics South Africa, 2001).

In the empirical analysis in section 5 below, growth-oriented micro-enterprises are distinguished from survivalist micro-enterprises by using the cluster analysis described in the next section. This data-driven technique requires a set of variables in order to group similar observations together. Care had to be taken to choose variables that would indicate work conditions and not the job-seekers' characteristics because these will be used as explanatory variables in the subsequent analysis of the determinants of informal sector entry.

4. Methodology

While heterogeneity within the informal sector is now widely recognised, there have only been a few studies that identify the different types of workers empirically. Often cut-off points in earnings or capital (Grimm et al. 2012) have been used to distinguish between workers. The position of these cut-offs are always somewhat arbitrary and matter for subsequent analysis of the reasons behind the decision to enter the informal sector. It is therefore advisable to use an approach that relies as little as possible on the inclinations of the econometrician to form the groups of labourers.

Cluster analysis is an exploratory statistical technique that partitions the data by maximising similarity within groups and minimising similarity between groups. The technique was introduced by the behavioural psychologist Tryon (1939). Cluster analysis is used, in this study, to find groups of informal sector labourers who share similar conditions. Once these groups are identified, we can

use our knowledge of the various characteristics of the different informal sector segments to classify each group as either survivalist or growth oriented micro-enterprise labourers.

The two main clustering procedures are the hierarchical and the partition methods. The hierarchical method organises groups in a tree-like structure by using various linking procedures (e.g. nearest neighbour). The partition method separates observations through an iterative process that uses the mean or median (centroid) of the groups. We use the k-medians procedure in this study which is an algorithm that sorts the data into *k* groups based on calculating the medians of the clustering variables. This procedure is well suited for large datasets because of the computational simplicity. Additionally, the k-medians procedure is less sensitive to outliers than hierarchical methods (Anderberg 1973).

K-medians clustering algorithm begins by choosing *k* observations randomly[2] from the dataset. These data points are used to form the first *k* clusters by grouping all other observations with the nearest initial observations. Next, the medians of the variables belonging to each of the *k* groups are calculated, which then become the centroids of the next round of clusters The new set of clusters is formed by grouping the observations with the shortest distance from the new centroids. This process is repeated with the calculation of medians of the current clusters and forming new centroids for another set of clusters. Initially the observations in each group will change a lot as the algorithm tries to find the *k* centroids that are most suitable to separate the data. The process stops when the centroids of the new clusters lead to identical observations as the previous clusters.

We chose variables that ought to help us distinguish the members of the different informal sector segments based on our review of the literature on the characteristics of the jobs in either tier. We had to be careful of choosing variables which will not be the variables of interest for modelling the incentives and constraints of informal sector transitions. For example, we cannot cluster according to education because we want to know how education influences the job-seeker's decision and would not simply want these estimates to reflect our choice of clustering variables. Informal sector workers in the data were identified by self-reported firm size and VAT registration (Statistics South Africa 2001). Occupations in the informal sector are found in small firms which offer their labourers low wages. The state provides little protection for workers in this sector in terms of regulation and the sector rarely has organised labour representation. Based on this information, the following variables should distinguish types of informal sector workers: logged wages, membership of union, firm size, industry, occupation, enterprise registration for VAT, hours worked, and whether the firm is private or public.

Once the clustering variables have been chosen, the dissimilarity (distance) measure must be chosen as well as the optimal number of clusters. The choice of the distance measure is based on the type of clustering variables that are chosen with the Minkowski metric, $d(x,y) = \left[\sum_{i=1}^{p} |x_i - y_i|^m\right]^{1/m}$,

[2] A consequence of this is that the final groups depend on what the initial observations were.

being the most commonly used metric (Anderberg 1973). When using continuous measures, the most popular types of the Minkowski metric are the absolute-value distance (m=1) and the Euclidean distance (m=2). Observations that are clustered according to binary variables are grouped through matching scores (Johnson and Wichern 2007). Consider a pair of observations that are described by a set of p binary variables. The Euclidean distance $d(x,y) = \sum_p (x_i - y_i)^2$ would then simply count the number of mismatches of zeros or ones.

Gower (1971) developed a distance measure that was suitable for both discrete and continuous data. This is the distance measure that is used in the analysis. The Gower dissimilarity coefficient, $\frac{\sum_p \delta_i(x,y) d_i(x,y)}{\sum_p \delta_i(x,y)}$, weights the distance $d_i(x,y)$ of the non-missing variables by the inverse of the number of variables $\left[\sum_p \delta_i(x,y)\right]^{-1}$ used to cluster the observations in the analysis. The distance measure for binary variables is the matching measure and the distance measure for continuous variables is the absolute-value distance divided by the range of the variable.

The choice of the number of groups (*k*) is based on how distinct the clusters are from each other. More groups generally yield more discrete clusters. Caliński and Harabasz (1974) derived a stopping rule based on the variance ratio criterion $VRC_k = \frac{SS_B/(k-1)}{SS_W/(n-k)}$ that can assist in the choice of *k*. Larger values of VRC_k indicate clearer groupings. A possible shortcoming of this method (and of choosing a large *k*) is that it becomes less informative when there are groups with a few observations.

5. Results

5.1. Description of the clusters

The k-medians clustering technique is applied to all of the employed respondents in the LFS dataset i.e. both formal and informal sector employees. The algorithm found that the best way to partition the data is to cluster at k=15 groups which gives the highest variance ratio criterion ($VRC_{15} = 2\,860.85$) within the range of $k = [2, 20]$ clusters. Each of the 15 groups has a unique combination of medians of the clustering variables. The descriptive statistics of some of these variables are shown in Table 1, while Table 2 shows how the 15 clusters are distributed across sectors in the economy.

Firstly, we will look at the first five groups which contain own account workers. Clusters 1 and 2 consist of workers who receive a low income for their labour. Most of the workers in these clusters are domestic workers. Approximately one in four workers in the first cluster are underemployed[3] and 42% of workers in the second cluster are underemployed. About half of the labourers work for

[3] Time related unemployment refers to the situation where workers who are willing and are able to work for a longer period, are constrained to less than 35 hours of labour in a week.

periods that are longer than 48 hours in a week. Nearly one in two of the workers in the Cluster 3 are street vendors. The majority of the workers in Clusters 4 and 5 are domestic workers who also receive low compensation for their labour. Almost 90% of the workers in the Cluster are low paid agricultural workers.

In sum, first 5 clusters are strongly dominated by typical characteristics of informal enterprises, although the first 5 clusters are not exclusively informal-sector jobs. The few formal sector workers in the first 5 clusters are shop attendants, cleaning staff, and have occupations in the taxi industry. Workers in these clusters are typically own account workers and they receive low hourly wages. The sixth cluster is mostly made up of commercial agriculture labourers who, unlike the domestic and informal sector workers, are employed in larger firms (or farms) that are registered for VAT. However, like the informal sector and domestic workers, these labourers receive little remuneration.

The remaining clusters are mainly public and private formal sector employees. These labourers work in more favourable conditions; they have higher wages, have written contracts, are more likely to be members of a union, and are able to make contributions towards their retirement. The majority of workers in clusters 7 to 15 work in enterprises that have been registered for VAT which leaves little room for informal sector workers. Any informal sector jobs that end up being categorised in these clusters have been grouped together with formal sector work because of the similarity in working conditions. Some of the most common occupations among the informal sector labourers in Clusters 7 to 15, i.e. those with earnings and working conditions similar to formal-sector workers are in the building trade, hairdressers and barbers, mechanics, and spaza (shop) workers.

Tables 1 show a detailed overview of the remaining nine clusters. Starting our description from Cluster 7 we can see that about three in four employees work in the construction industry. Approximately 57% of the workers in Cluster 7 are in firms that are registered for VAT and there is little worker protection in the form of unionisation (8%). The majority of workers in Cluster 8 are in the wholesale and retail industry. Cluster 9 consists of workers in mining companies. The mining industry is highly unionised which is shown by approximately 90% of workers being members of a union. All of the workers in Cluster 10 are in the manufacturing industry with the most common occupation in this cluster being plant and machinery operations at approximately 35% of the employees. The majority of Cluster 11 employees are in the financial services industry and work in a range of occupations like technicians (15%), clerks (26%), and sales workers (22%). Employees in Cluster 12 earn almost ten times as much as the median worker in Cluster 1. The median income for employees in the rest of the clusters is even higher. The majority of workers in Clusters 12, 13, and 14 are public sector workers in the community and social services industry. The 15th cluster has the highest median wage with employees who have jobs in large private sector firms in the transport (23%), financial (16%) and communications (22%) industries.

Table 1: Summary statistics of clusters

	Income	Wage	Weekly hours	Under-employed	VAT	Public	Self-employed	Firm size (<5)	Firm size (<10)	Union	Contract	Pension
1	511.88	4.22	30	0.26	0.03	0.00	0.17	0.98	0.99	0.01	0.10	0.04
2	643.23	10.62	15	0.42	0.02	0.00	0.27	0.99	1.00	0.01	0.08	0.03
3	716.25	4	49	0.11	0.00	0.00	0.77	0.90	0.94	0.01	0.04	0.07
4	741.28	2.86	60	0.01	0.03	0.00	0.15	0.95	0.98	0.01	0.09	0.03
5	1070.74	5.83	40	0.02	0.03	0.00	0.09	0.96	0.98	0.02	0.20	0.08
6	1073.67	5.17	50	0.02	0.91	0.00	0.08	0.17	0.34	0.07	0.42	0.18
7	1986.34	10.18	45	0.06	0.57	0.01	0.25	0.35	0.52	0.08	0.33	0.27
8	2383.61	12.22	45	0.04	0.93	0.00	0.13	0.25	0.48	0.16	0.53	0.42
9	3514.92	17.36	48	0.01	0.98	0.00	0.00	0.02	0.05	0.90	0.89	0.87
10	3709.43	19.94	45	0.02	0.95	0.00	0.04	0.06	0.12	0.40	0.73	0.69
11	4202.93	22.06	45	0.02	0.91	0.01	0.13	0.22	0.40	0.15	0.67	0.55
12	5045.21	28.30	40	0.03	1.00	0.76	0.00	0.08	0.17	0.70	0.81	0.87
13	6648.25	38.14	40	0.04	0.00	0.90	0.02	0.09	0.19	0.70	0.76	0.88
14	7525.21	42.96	40	0.01	1.00	0.82	0.00	0.03	0.09	0.77	0.85	0.93
15	8514.01	44.00	45	0.02	0.95	0.00	0.09	0.07	0.12	0.17	0.74	0.74
Total	2383.61	12.73	45	0.05	0.58	0.15	0.15	0.33	0.44	0.27	0.51	0.51

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
*Values expressed as proportions
*Median hourly wage in 2012 prices and median hours spent at work in a week.

Table 2: Employment sectors of clusters (distribution across sectors)

	Domestic workers	Informal sector	Formal sector	Subsistence Agriculture	Commercial Agriculture	(mv)	Total	Informal sector totals	Tier totals
1	68.66	27.75	0.93	2.55	0.04	0.07	2 789	774	
2	58.65	36.03	2.15	3.16	0.00	0.00	2 878	1 037	
3	0.00	86.47	9.54	3.27	0.18	0.54	11 718	10 133	
4	65.62	27.59	5.12	1.27	0.06	0.34	4 962	1 369	
5	78.66	18.77	1.51	0.89	0.07	0.10	6 751	1 267	
6	0.03	0.51	9.37	9.06	80.14	0.89	14 998	76	**14656**
7	0.00	38.73	59.06	0.05	0.65	1.52	6 453	2 499	
8	0.10	6.85	92.40	0.00	0.00	0.66	15 703	1 075	
9	0.07	0.63	97.90	0.00	1.28	0.11	7 088	45	
10	0.00	3.25	96.42	0.00	0.00	0.34	13 307	432	
11	0.08	5.78	93.72	0.00	0.00	0.41	7 712	446	
12	0.08	0.99	98.34	0.00	0.47	0.12	2 532	25	
13	0.03	5.17	94.17	0.02	0.44	0.17	15 311	791	
14	0.00	0.62	98.94	0.00	0.26	0.18	3 402	21	
15	0.31	1.86	87.46	0.42	9.73	0.23	7 844	146	**5481**

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
*Percentage of workers in each sector per cluster
*(mv) denotes missing values
*Total denotes the number of workers in each cluster

5.2. Survivalist micro-enterprises and growth-oriented micro-enterprises (informal sector)[4]

Cluster analysis has sorted the jobs according to their wages, weekly hours worked, firm size, industry, occupation, union membership, registration for VAT, and whether the firm is private or public. These variables are correlated to the types of sectors in the labour market, and it follows that some clusters are dominated by certain sectors. For example, the twelfth cluster consists almost entirely of formal sector workers. The sectors, however, are not perfectly distinguishable by the clusters. There are some informal sector jobs that are very similar to formal sector and have been sorted by the data-driven technique accordingly. In this section we exploit the relationship between the sectors and the clusters as well as the prominent characteristics of the 15 groups to simplify the diverse nature of informal sector jobs (column 2 of Table 2) into two tiers; the survivalist and the growth-oriented enterprises.

Survivalist enterprises are typically smaller and offer lower wages. These jobs have lower entry requirements and do not provide a secure form of employment (Fields 1990). Survivalist enterprises are usually single person firms therefore they are more likely to conduct small scale activities with low start-up costs and capital requirements. The hours worked are variable, depending on the weekly market and other non-market conditions. For example, street vendors which are some of the more common survivalist enterprises are often exposed to weather conditions that affect the hours worked and remuneration earned. The typical informal sector workers in clusters 1 to 6 earn low wages, work in in small firms, are frequently self-employed (especially cluster 5), do not have written contracts of employment and have variable weekly hours, all of which are common characteristics of survivalist jobs.

Growth oriented micro-enterprises, on the other hand, are larger, have more capital and skills requirements, and maintain more formal labour agreements. For example, spaza owners need more costly infrastructure to set up shop and they are more likely to have fixed operating times. Vehicle mechanics require knowledge and skills that have been accumulated from some form of training or apprenticeship in order to operate. The more stringent financial and human capital barriers to entering the growth oriented micro-enterprise tier allow these workers to earn higher wages that their counterparts in the survivalist sector. By and large, informal growth oriented micro-enterprises operate like formal sector firms that have not been registered for VAT. The informal sector workers in clusters 7 to 15 resemble the labourers that are expected in growth-oriented enterprises.

Based on this information, the informal sector workers in clusters 1 to 6 are classified as survivalist micro-enterprise workers and the informal sector labourers in clusters 7 to 15 are classified as growth-oriented micro-enterprise workers. The focus of this study is based on the heterogeneity within the informal sector therefore the diverse nature of the formal sector is not addressed beyond this point.

4 This section only applies to the informal sector i.e. only the workers in column 2 of Table 2.

Tables 3 to 5 show some summary statistics of survivalist enterprises, growth-oriented enterprise workers, formal sector labourers, and people who are out of work. The purpose of these summary statistics is twofold. One is to describe the personal and household characteristics of each segment of the working aged population. Another is to verify that there are sufficient differences across clusters to justify the clustering and subsequent allocation of informal sector workers into the two tiers.

The majority of the informal sector consists of survivalist enterprise workers who earn low hourly income in single person firms. In the growth-oriented enterprise segment, a majority (54%) of workers are employees, with 46% being self-employed owner/employers.

Despite the low wages, close to one in five survivalists would be willing to work for more hours in a day. Workers in growth-oriented enterprises are more likely to have somewhat more secure hiring practices with written contracts of employment (though still a much lower proportion than workers in the formal sector). We can also see that fewer growth-oriented enterprise workers are dissatisfied with their current job than survivalist enterprise workers – but much more than those in the formal sector jobs.

Table 3: Summary statistics of employed labour force

	Observations	Per cent	Wage	Self-employed	Contract	Hours (weekly)	Work longer	Different job
Survivalist enterprise	14 656	10.5	4.54	0.74	0.02	46.80	0.28	0.24
Growth oriented enterprise	5 480	3.9	10.57	0.46	0.15	46.38	0.22	0.14
Formal sector	85 691	61.5	21.44	0.06	0.70	46.13	0.11	0.07
Domestic & agri workers	33 621	24.1	4.42	0.16	0.26	43.96	0.13	0.10
Total	139 448	100	12.15	0.17	0.51	45.70	0.14	0.10

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
*Median wage in 2012 prices.
*Self-employed, and contract are expressed as proportions of individuals who are self-employed or have written contracts in each row. Hours are expressed as weekly averages.
*'Work longer' is the proportion of labourers in a row category who are willing to work for more hours in a week. 'Different job' is the proportion of labourers who want to work for more hours in a different job.

There are gender biases in the types of occupations in the informal sector (Manning 1993). As a result of the differences of occupations in either tier, the membership of the survivalist and growth oriented micro-enterprise tiers is also gendered. A detailed look at the occupations in the survivalist tier shows that about 39.65%[5] of workers are street vendors, which are jobs that are usually carried out by women. On the other hand, approximately 28.41%[6] of the occupations in the growth oriented micro-enterprise segment of the informal sector are related to the building trade which is dominated by men. A broader look at the segments within the informal sector in Table 4 shows us that the majority of workers in the growth-oriented enterprises are men whereas the workers in the survivalist enterprises are evenly represented.

5 This figure is calculated by using the detailed occupation categories in the dataset.
6 This figure is calculated by using the detailed occupation categories in the dataset.

The household characteristics of the working age population reveal how vulnerable survivalist enterprise labourers are relative to growth-oriented enterprise workers. Table 4 shows us that one in two survivalist enterprise workers reported that they never or seldom have trouble satisfying their food consumption needs. The inability to satisfy food consumption needs is less of a problem for growth oriented micro-enterprise workers who, on average, live in slightly smaller households. Survivalist enterprise workers tend to be old and are heads of their households so they may have taken on roles in which they are required to provide for their families.

Table 4: Household characteristics of the working age population

	Women	Food satisfaction	Household size	Education	Age	Household head	Urban
Out of work	0.62	0.53	5.7	9	32.5	0.27	0.56
Survivalist enterprise	0.52	0.53	4.6	8	39	0.62	0.56
Growth oriented enterprise	0.27	0.62	4.4	9	37.5	0.63	0.62
Formal	0.39	0.83	4.1	12	37.5	0.60	0.80
Domestic & agri workers	0.58	0.64	4.3	6	38	0.58	0.31
Total	0.53	0.62		9	31.5	0.36	0.59

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
*Food satisfaction: proportion of respondents who reported to never or seldom having trouble satisfying their food needs.
*Urban, Women, Household head are expressed as proportions of individuals who have these characteristics in each segment.
*Age and Education are expressed in the median. Household size is expressed at the mean

Private and public financial safety nets play an important role as non-wage income for individuals. The largest social grants in terms of coverage and value are the old age pension and the child support grant. These are means tested and people who are out of work can be seen to be more likely to live in households that have beneficiaries of these grants (Table 5). The differences between survivalist and growth oriented micro-enterprise workers are more apparent for financial assets. Growth oriented micro-enterprise workers are more likely to have saving in a bank than survivalist enterprise workers. Saving for long term contingencies is much more difficult for survivalist enterprise workers as we can see from the lower proportion of people with some retirement plan and funeral cover. Given the lack of ability to satisfy immediate food consumption needs, the lack of savings for future contingencies probably indicates inability rather than unwillingness to save.

Table 5: Financial characteristics of the working age population

	Old age pension	No. of Elderly	Child support grant	No. of children	Bank	Retirement plan	Funeral cover
Out of work	0.30	0.4	0.15	1.9	0.34	0.09	0.21
Survivalist enterprise	0.13	0.2	0.14	1.7	0.33	0.04	0.14
Growth oriented enterpr	0.12	0.2	0.10	1.4	0.45	0.11	0.23
Formal	0.09	0.2	0.05	1.2	0.71	0.31	0.49
Domestic & agri workers	0.11	0.2	0.10	1.4	0.26	0.07	0.18
Total	0.21	0.3	0.12	1.7	0.44	0.14	0.28

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
*Old age pension and Child support grant7 are the proportions of households with at least one beneficiary.
*Elderly and Children are the average members of the household who are older than 60 and younger than 15 years respectively.
*Bank, Retirement plan, and Funeral cover are expressed as proportions of individuals who have these types of savings.

7 Data for the old age pension and the child support grant was collected only in the September rounds 2001-2003 dataset

The characteristics of the individuals in either tier correspond to the features of informal-sector survivalist and growth-oriented micro-enterprises. Therefore, the data-driven technique seems to have partitioned the data in a sensible manner. It seems like informal-sector upper-tier work has higher returns than survivalist work and the lives of survivalists are more difficult. One would expect job-seekers who are willing (and able) to work in the informal sector to have a preference for upper-tier work. However, there are more survivalist enterprise workers in the informal sector, which is a result of higher retention and entry rates into the tier. The next section takes a look at labour market transitions with a specific look at the movements of people who are informal-sector survivalist enterprise workers and/or growth oriented micro-enterprise workers at some point.

5.3. Transitions

The South African informal sector is known to experience a lot of churning, but these dynamics may be very different for survivalist and growth-oriented enterprise workers. Workers in growth-oriented micro-enterprises are in a better position in terms of material well-being than survivalist enterprise workers. So job-seekers who were not able to find employment in the formal sector ought to prefer the jobs in this tier of the informal sector. However, we may witness higher entry rates in the survivalist enterprise part of the informal sector which may indicate the existence of barriers to entry into the preferred growth oriented enterprise tier.

To analyse this, we constructed a transition matrix using the LFS panel dataset[8] collected in the period between September 2001 and March 2004. The transition matrix in Table 6 shows the probabilities of changing from one state (row) to the next state (column) in a period of six months. For example, a person who was not working had a 79.58% chance of being out of work, and a 9.74% chance of working in the formal sector within the next six months.

The probability of finding employment in either a survivalist or a growth oriented enterprise was low. This is illustrated in the transition matrix where only 1.21% of people who were out of work[9] were able to find a job in the growth-oriented enterprise tier. The probability of entering the survivalist tier for a person who was not working was almost three times that of entering the growth oriented enterprise. This is an indication of the relative ease with which out-of-work searchers can enter the survivalist tier compared to the growth-oriented enterprise tier. Movements within the informal sector illustrate a similar pattern; it is less difficult to enter the survivalist tier than the growth oriented tier (assuming that the better work characteristics would make the growth oriented enterprise more attractive). The probability of a survivalist worker moving to a growth-oriented enterprise is about 3.9% whereas the probability of a movement in the opposite direction is 10.6%.

Searchers who were not able to meet the financial and/or human capital requirements to work in growth-oriented enterprises or the formal sector were likely to continue with out of work job search.

8 A dataset that is collected by surveying the same individuals over multiple periods of time
9 Strictly unemployed, discouraged work seekers, and NEA

On the other hand, those who have been unsuccessful in finding a formal sector or informal sector growth-oriented enterprise job and have accepted a survivalist job probably do so because they simply cannot afford to stay out of work; even for a marginal income.

Table 6: Transition matrix with a six month period

		Period t+1				
		Out of work	**Survivalist enterprise**	**Growth oriented enterprise**	**Formal**	**Domestic & agri-cultural workers**
Period t	**Out of work**	79.20	3.92	1.21	10.06	5.62
	Survivalist enterprise	38.77	36.77	3.93	13.73	6.80
	Growth oriented enterprise	32.60	10.64	17.87	33.51	5.38
	Formal	16.39	2.18	1.87	76.66	2.91
	Domestic & agri workers	25.83	3.08	0.88	8.93	61.27

Source: Own calculations using September 2001 to March 2004 data for persons aged 15 to 65

The informal sector segment that the job-seeker is in has important implications about the likely duration of their employment as well as the likely direction of their labour market mobility. Survivalist enterprise workers are more likely than growth oriented enterprise workers to transition into non-employment. Growth oriented enterprise workers are likely to spend a shorter time in their tier than survivalist enterprise workers because they have a much higher probability of making the transition to formal sector work in six months' time. Yet a third of workers in growth-oriented enterprises lost their job in the relevant period; but this is lower than for those in survivalist enterprises.

The descriptive statistics in the previous section showed that a higher proportion of survivalist enterprise labourers wanted to work in a different job yet the transition matrix illustrates that they are not as successful at upward mobility. It seems that a lot (about a third) of the exits from growth oriented enterprises appear to be voluntary (moving into a 'better' cluster in the formal sector) whereas the higher retention that we see in survivalist enterprises is indicative of lower upward mobility and persistent low earnings.

5.4. Correlates of informality

Amidst the objectionable working conditions and poor prospects of upward mobility, unemployed people have a higher probability of finding a job in the survivalist tier than in growth oriented enterprises. A possible explanation is that job-seekers who become survivalist enterprise workers were not sufficiently protected by private or public safety nets and their household responsibilities induce them to accept the harsh work conditions. On the other hand, the more appealing growth oriented enterprises are not able to accommodate enough seekers from the job queue. This could be due to a lack of income opportunities in the market or high barriers to entry.

In the following analysis, we attempt to explain the labour market transitions of out of work job-seekers by modelling the relationship between worker and household characteristics and the probability of transitioning into either of the two informal sectors. Specifically, we investigate how the

individual's human capital, access to household income, and demographics affect their likelihood of entering the survivalist or growth oriented enterprise sectors. This relationship is modelled using multinomial logit, the linear probability model, and fixed effects estimators.

The regression analyses in Table 7 indicate that there may be capital and skills requirements in the growth oriented enterprises that create barriers to entry. Job-seekers who entered the upper (growth oriented) tier had more access to financial capital, whereas those who entered the survivalist tier came from households with lower income. Survivalist enterprise workers were job-seekers who had fewer years of schooling; job-seekers who had some secondary education had a lower probability of entering this tier than those who had completed their schooling at the primary level. Survivalist enterprise workers were also less likely to be a part of the racial population groups that could have accumulated historical capital.

Table 7: Multinomial logistic regressions (MNL) and Linear probability models (LPM) of job-seekers who were out of work in the previous six months

	MNL		LPM[10]		LPM	
	Survivalist	**Growth oriented**	**Survivalist**	**(FE)**	**Growth oriented**	**(FE)**
HH income(-1)	-0.155	0.047	-0.006	-0.003	0.0003	0.004
	(5.16)***	(0.89)	(5.76)***	(0.89)	(0.49)	(2.38)**
Primary	0.039	0.056	0.002		0.001	
	(2.99)***	(2.13)**	(3.10)***		(2.08)**	
Secondary	-0.052	0.005	-0.002		0.0001	
	(2.37)**	(0.14)	(2.83)***		(0.24)	
Tertiary	-0.446	0.125	-0.007		0.002	
	(1.63)	(0.66)	(1.37)		(0.74)	
Age	0.148	0.151	0.005	0.005	0.002	0.002
	(9.62)***	(5.28)***	(10.01)***	(2.59)***	(5.77)***	(1.61)
Age squared	-0.002	-0.002	-0.000	-0.000	-0.000	<-0.001
	(10.30)***	(6.39)***	(11.14)***	(2.76)***	(7.25)***	(2.07)**
Coloured	-0.721	-0.050	-0.020		-0.0004	
	(4.91)***	(0.23)	(4.56)***		(0.19)	
Indian	-1.063	0.372	-0.024		0.005	
	(3.37)***	(1.24)	(3.32)***		(1.25)	
White	-1.277	-0.054	-0.027		-0.001	
	(4.56)***	(0.20)	(4.36)***		(0.23)	
Head	0.741	1.256	0.040		0.020	
	(9.80)***	(9.25)***	(12.36)***		(10.44)***	
Married	0.446	0.533	0.021	0.007	0.009	0.002
	(7.05)***	(4.61)***	(8.48)***	(0.78)	(5.81)***	(0.33)
HH size	-0.044	-0.030	-0.002	-0.001	-0.0003	0.001
	(4.01)***	(1.48)	(4.26)***	(0.70)	(1.65)*	(1.49)
Constant	-5.189	-7.606	-0.016	-0.029	-0.021	-0.039
	(12.85)***	(10.64)***	(1.15)	(0.69)	(2.66)***	(1.78)*
N	33,895		33,498	33,666	32,577	32,736

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
* p<0.1; ** p<0.05; *** p<0.01
MNL (multinomial logistic regression) and LPM (linear probability model). Base category: Black women who are out of work. Logged per capita household income in 2012 prices. Primary, secondary, tertiary: splines of years completed schooling. Coefficients for province, rural, and period are not shown

10 Results from logit and fixed effects logits are in the Appendix

Other demographics indicate the job-seeker's role in the household. Searchers who are married or heads of households (i.e. hold positions that are associated with more household responsibilities) are more likely to enter the informal sector. The individuals who have these roles have a higher probability of entering the growth oriented tier than the survivalist tier. So from the regression analyses it seems like the entrants of the survivalist tier share characteristics with the vulnerable portion of the population, whereas the upper-tier workers have access to factors (such as more household income and education) that are often the source of barriers to higher earning jobs.

The summary statistics of survivalist and growth oriented micro-enterprise workers showed a gendered informal sector.[11] Gender roles also extend to the household responsibilities which influence the job-seekers' decision, which suggests that different processes may be determining the outcomes of women and men.

Table 8 Multinomial logistic regressions (MNL) and Linear probability models (LPM) of women who were out of work in the previous six months

	MNL		LPM[12]		LPM	
	Survivalist	**Growth oriented**	**Survivalist**	**(FE)**	**Growth oriented**	**(FE)**
HH income (-1)	-0.161	0.124	-0.007	-0.002	0.001	0.002
	(4.26)***	(1.32)	(4.96)***	(0.54)	(1.63)	(1.57)
Primary	0.041	-0.007	0.002		-0.000	
	(2.60)***	(0.14)	(2.58)***		(0.15)	
Secondary	-0.023	0.303	-0.001		0.002	
	(0.82)	(4.34)***	(1.07)		(4.74)***	
Tertiary	-0.423	0.238	-0.007		0.006	
	(1.26)	(1.05)	(1.03)		(2.21)**	
Age	0.174	0.195	0.006	0.009	0.001	0.001
	(8.58)***	(3.54)***	(8.48)***	(2.87)***	(3.35)***	(0.80)
Age squared	-0.002	-0.002	-0.000	-0.000	-0.000	-0.000
	(8.70)***	(3.56)***	(8.96)***	(2.66)***	(3.46)***	(0.68)
Coloured	-1.109	-0.600	-0.026		-0.003	
	(5.35)***	(1.38)	(4.79)***		(1.33)	
Indian	-1.681	-0.077	-0.032		0.000	
	(3.64)***	(0.16)	(3.76)***		(0.03)	
White	-1.319	-1.380	-0.028		-0.009	
	(4.13)***	(2.70)***	(3.86)***		(3.23)***	
Head	0.800	0.484	0.045		0.003	
	(7.87)***	(1.51)	(9.88)***		(1.63)	
Married	0.471	0.561	0.023	-0.001	0.004	0.004
	(5.67)***	(2.50)**	(6.84)***	(0.08)	(2.92)***	(0.95)
HH size	-0.040	-0.019	-0.001	-0.002	-0.000	-0.000
	(2.90)***	(0.52)	(2.90)***	(1.28)	(0.44)	(0.25)
Constant	-5.826	-10.047	-0.032	-0.102	-0.019	-0.025
	(11.13)***	(7.49)***	(1.85)*	(1.63)	(2.70)***	(1.16)
N	21,534		21,412	21,509	20,691	20,782

Source: Own calculations from the September 2001 to March 2004 StatSA LFS. * p<0.1; ** p<0.05; *** p<0.01
MNL (multinomial logistic regression) and LPM (linear probability model). Base category: Black women who are out of work. Logged per capita household income in 2012 prices. Primary, secondary, tertiary: splines of years completed schooling. Coefficients for province, rural, and period are not shown.

11 Refer to Table 4 for proportion of women in each segment of the economy
12 Results from logit and fixed effects logits are in the Appendix

According to the regression coefficients in Table 8 women who start work as survivalist enterprise workers tend to live in households with lower per capita income and have completed fewer years of school. The probability of working in survivalist enterprises is the highest for women who had only studied up to primary school. The race effect also demonstrates that survivalist employment is negatively correlated to privilege. On the other hand, women who lived in households with higher per capita income and had some secondary or tertiary education were more likely to find work in the upper-tier of the informal sector. This illustrates the barriers that prevent women with lower capital and academic skills from entering growth oriented micro-enterprises.

Entry into either informal sector is more likely if the individual has a role of responsibility for the household. Women who are married and/or household heads have a higher probability of working in the informal sector than other household members. The role that the person plays in the household is more important for entry into the informal sector than the household size. Women from larger households are less likely to work in the informal sector. Therefore the push factor into the informal sector is not necessarily derived from how many people there in the household but rather whether the woman is responsible for providing for them.

A comparison of Table 8 and Table 9 shows us that the barriers for growth oriented micro-enterprise entry are less pronounced for men than they are for women. Men do not need to complete as high a level of schooling as women in order to find work in the growth oriented enterprises. Added to that, the income effect for growth oriented enterprises is insignificant. If a man was part of the race group that could accumulate more historical capital he would be more likely to work in a growth oriented micro-enterprise but not as survivalist enterprises.

The effects of job-seekers' characteristics can be more persuasively identified by taking advantage of the panel component of the data and allowing for fixed effects in the regressions. Fixed effects estimation removes the time invariant unobserved heterogeneity that biases the results of the observed characteristics. The change of the direction for the income effect for men in Table 9 illustrates that the bias appears to be a larger problem for the growth oriented micro-enterprise regressions of men. Fixed effects remove this bias and by doing so reveal that men who have experienced an increase in their per capita household income are more likely to join a growth oriented micro-enterprise. The fact that men who have gained access to financial capital are also more likely to enter the growth oriented enterprises shows that capital may also be a barrier to entering the upper-tier for men.

Men who become survivalist enterprise workers react to a change in household attributes in a similar manner as women entering the informal sector; the role is more important for entry than the size. On the other hand, men who have experienced an increase in their household size or get married have a higher probability of working in growth oriented micro-enterprises[13]. The information about

13 Once the fixed effects have been taken into account

how men are able to secure better jobs in the informal sector if the household size increases when women cannot, illustrates the gender roles in the informal sector and expectations from the household.

Table 9 Multinomial logistic regressions (MNL) and Linear probability models (LPM) of men who were out of work in the previous six months

	MNL		LPM		LPM	
	Survivalist	**Growth oriented**	**Survivalist**	**(FE)**	**Growth oriented**	**(FE)**
HH income (-1)	-0.128	-0.017	-0.005	-0.000	-0.001	0.008
	(2.53)**	(0.25)	(2.87)***	(0.01)	(0.39)	(1.76)*
Primary	0.051	0.067	0.002		0.001	
	(2.18)**	(2.10)**	(2.30)**		(2.09)**	
Secondary	-0.098	-0.119	-0.004		-0.003	
	(2.72)***	(2.55)**	(3.10)***		(2.54)**	
Tertiary	-0.500	-0.195	-0.007		-0.004	
	(1.05)	(0.50)	(0.85)		(0.62)	
Age	0.124	0.172	0.005	0.011	0.004	0.001
	(4.91)***	(5.07)***	(6.14)***	(2.41)**	(6.46)***	(0.20)
Age squared	-0.002	-0.003	-0.000	-0.000	-0.000	-0.000
	(6.14)***	(6.23)***	(7.75)***	(2.44)**	(7.97)***	(0.90)
Coloured	-0.256	0.164	-0.008		0.004	
	(1.20)	(0.64)	(1.15)		(0.64)	
Indian	-0.194	0.835	-0.008		0.022	
	(0.44)	(2.13)**	(0.55)		(2.00)**	
White	-1.352	0.823	-0.027		0.020	
	(2.27)**	(2.48)**	(2.31)**		(2.25)**	
Head	0.699	0.666	0.039		0.023	
	(4.32)***	(3.29)***	(5.78)***		(4.39)***	
Married	0.772	0.956	0.036	0.030	0.027	0.055
	(5.29)***	(5.19)***	(5.98)***	(1.30)	(5.81)***	(3.28)***
HH size	-0.047	-0.056	-0.002	-0.004	-0.001	0.004
	(2.53)**	(2.23)**	(2.89)***	(1.51)	(2.32)**	(1.97)**
Constant	-4.806	-6.593	-0.013	-0.146	-0.039	-0.032
	(7.44)***	(7.78)***	(0.60)	(1.59)	(2.27)**	(0.48)
N	12,361		12,086	12,157	11,886	11,954

Source: Own calculations from the September 2001 to March 2004 StatSA LFS
* p<0.1; ** p<0.05; *** p<0.01
Base category: Black men who are out of work
Logged per capita household income in 2012 prices
Primary, secondary, tertiary: splines of years completed schooling
Coefficients for province, rural, and period are not shown

6. Conclusion

The informal sector refers to a set of unregistered enterprises that carry out various activities. Some of these small businesses have been created as a means for the job-seeker to cope with their poverty; these are survivalist enterprises. Other small businesses within the informal sector have been established in order to take advantage, as a business owner and entrepreneur, of income opportunities, and require more initial physical and human capital; these are growth oriented micro-enter-

prises. One of the challenges to furthering analysis in the informal sector lies in identifying the two tiers empirically. In this study we use a data-driven clustering technique to find fifteen natural subgroups in the informal sector jobs. This enables us to classify the jobs/workers into two tiers.

Most of the informal sector (approximately 75%) comprises survivalist enterprise labourers that work in low wage, often single person firms. Despite low wages, survivalist enterprise workers are willing to work for more hours in a day albeit in a different job. However, the probability of making the transition to informal growth oriented micro-enterprises or the formal sector is low.

The growth oriented micro-enterprise tier consists of employees in more established, sometimes larger firms who are able to earn higher wages. The level of satisfaction in this tier is higher and job-seekers have been able to use it as a spring-board into the formal sector. Entry rates into this portion of the sector are low regardless, indicating possible barriers to entry.

Successful allocations of workers into their respective tiers have allowed us to gain more insight into the characteristics of out of work job-seekers who find employment in survivalist or growth oriented micro-enterprises within six months. Job-seekers who started working in the survivalist tier had experienced a decrease in household income while they were out of work and had completed fewer years of schooling. They were heads of their household or were married which is indicative of the responsibility they have to provide for their family.

Heterogeneity within the informal sector is gendered; a lot of the jobs in the survivalist tier are carried out by women while a lot of the jobs in the growth oriented micro-enterprises are carried out by men. Men who had experienced an increase in their household income were more likely to start working in growth oriented micro-enterprises. Women who found work in growth oriented micro-enterprises came from households with more income and had attained a higher level of schooling which helped them deal with the capital and skills requirements that prevent entry into this tier.

Understanding the heterogeneity within the informal sector provides important information about the large and open unemployment in South Africa. A large part of the informal sector acts as an employer of last resort while there is a smaller, more entrepreneurial portion that generates higher income – but job-seekers cannot enter it easily because of the barriers to entry.

* * *

Bibiography

Anderberg, M.R., 1973. *Cluster analysis for applications*, New York: Academic Press.

Badaoui, E. El, Strobl, E. & Walsh, F., 2008. Is There an Informal Employment Wage Penalty ? Evidence from South Africa. *Economic Development and Cultural Change*, 56(3), pp.683–710. Available at: http://www.jstor.org/stable/10.1086/533550.

Banerjee, A. et al., 2008. Why has unemployment risen in the New South Africa? *Economics of Transition*, 16(4), pp.715–740. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0351.2008.00340.x/full.

Bertrand, M., Mullianathan, S. & Miller, D., 2003. Public Policy and Extended Families: Evidence from Pensions in South Africa. *The World Bank Economic Review*, 17(1), pp.27–50. Available at: http://wber.oupjournals.org/cgi/doi/10.1093/wber/lhg014 [Accessed November 23, 2014].

Bhorat, H. & Hodge, J., 1999. Decomposing Shifts in Labour Demand in South Africa. *The South African Journal of Economics*, 67(3), pp.155–168. Available at: http://doi.wiley.com/10.1111/j.1813-6982.1999.tb01146.x.

Blunch, N.-H., Canagarajah, S. & Raju, D., 2001. The Informal Sector Revisited: A Synthesis Across Space and Time. *Social Protection Discussion Paper Series No. 0119*, (0119), pp.1–31. Available at: http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SPDiscussion-papers/Labor-Market-DP/0119.pdf.

Brand, V., 1986. One Dollar Workplaces : A Study of Informal Sector Activities in Magaba , Harare. , pp.53–74.

Caliendo, M. & Kritikos, A.S., 2009. *I Want to , But I Also Need to': Start-Ups Resulting from Opportunity and Necessity*, Available at: http://www.econstor.eu/bitstream/10419/35884/1/618732640.pdf.

Caliński, T. & Harabasz, J., 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1), pp.1–27.

Clark, A.E., 2003. Unemployment as a social norm: Psychological evidence from panel data. *Journal of labor economics*, 21(2), pp.323–351.

Coetzee, M., 2013. Finding the Benefits: Estimating the Impact of The South African Child Support Grant. *South African Journal of Economics*, 81(3), pp.427–450. Available at: http://dx.doi.org/10.1111/j.1813-6982.2012.01338.x.

Dickens, W. & Lang, K., 1988. The reemergence of segmented labor market theory. *American Economic Review*, 78(2), pp.129–134. Available at: http://ideas.repec.org/a/aea/aecrev/v78y1988i2p129-34.html.

Duflo, E., 2003. Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review*, 17(1), pp.1–25. Available at: http://wber.oupjournals.org/cgi/doi/10.1093/wber/lhg013.

Fields, G.S., 1990. Labour Market Modelling and the Urban Informal Sector : Theory and Evidence, in David Turnham, Bernard Salomé, and Antoine Schwarz, eds., The Informal Sector Revisited. (Paris: Development Centre of the Organisation for Economic Co-Operation and Development).

Gower, J.., 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, 27(4), pp.857–871. Available at: http://www.jstor.org/stable/2528823.

Grimm, M., Knorringa, P. & Lay, J., 2012. Constrained Gazelles: High Potentials in West Africa's Informal Economy. *World Development*, 40(7), pp.1352–1368. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0305750X12000460 [Accessed November 23, 2014].

Günther, I. & Launov, A., 2012. Informal employment in developing countries Opportunity or last resort? *Journal of Development Economics*, 97, pp.88–98. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0304387811000046 [Accessed November 12, 2014].

Heintz, J. & Posel, D., 2008. Revisiting informal employment and segmentation in the South African labour market. *South African Journal of Economics*, 76(1), pp.26–44. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1813-6982.2008.00153.x/full.

Hussmanns, R., 2004. Statistical definition of informal employment : Guidelines endorsed by the Seventeenth International Conference of Labour Statisticians (2003) By. , (February), pp.2–4.

Johnson, R.A. & Wichern, D.W., 2007. *Applied multivariate statistical analysis*, Upper Saddle River, NJ: Pearson/Prentice Hall.

Kingdon, G.G. & Knight, J., 2004. Unemployment in South Africa: The Nature of the Beast. *World Development*, 32(3), pp.391–408. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0305750X03002407 [Accessed November 17, 2014].

Klasen, S. & Woolard, I., 2008. Surviving Unemployment Without State Support: Unemployment and Household Formation in South Africa. *Journal of African Economies*, 18(1), pp.1–51. Available at: http://jae.oxfordjournals.org/cgi/doi/10.1093/jae/ejn007 [Accessed September 15, 2015].

Layard, P.R.G., 2005. *Happiness : lessons from a new science*, New York: Penguin Press.

Leibbrandt, M. et al., 2010. *Trends in South African Income Distribution and Poverty since the Fall of Apartheid*, Available at: http://dx.doi.org/10.1787/5kmms0t7p1ms-en.

Lucas, R.E. et al., 2004. Unemployment Alters the Set Point for Life Satisfaction. *Psychological Science*, 15(1), pp.8–13. Available at: http://pss.sagepub.com/lookup/doi/10.1111/j.0963-7214.2004.01501002.x [Accessed October 13, 2015].

Lund, F., 1998. Women street traders in urban South Africa : A synthesis of selected research findings. , (15), pp.1–50. Available at: http://wiego.org/sites/wiego.org/files/publications/files/Lund-women-street-traders.pdf.

Manning, C., 1993. INTERNATIONAL EXPERIENCES OF INFORMAL SECTOR ACTIVITY AND THE LESSONS FOR SOUTH AFRICA. *Transformation*, 22, pp.51–69. Available at: http://www.kznhass-history.net/ojs/index.php/transformation/article/viewFile/630/446.

Mkhize, S., Dube, G. & Skinner, C., 2013. *Informal Economy Monitoring Study: Street Vendors in Durban, South Africa*, Manchester, UK. Available at: http://wiego.org/sites/wiego.org/files/publications/files/IEMS-Durban-Street-Vendors-City-Report-English.pdf.

Moser, C., 1978. Informal Sector or Petty Commodity Production : Dualism or Dependence in. *World Development*, 6(9/10), pp.1041–1064. Available at: http://www.sciencedirect.com/science/article/pii/0305750X78900621.

Nattrass, N. & Walker, R., 2005. Unemployment and reservation wages in working-class Cape Town. *South African Journal of Economics*, 73:3(September), pp.498–509. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1813-6982.2005.00034.x/abstract.

Posel, D., Fairburn, J. a. & Lund, F., 2006. Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa. *Economic Modelling*, 23(5), pp.836–853. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0264999305000970 [Accessed October 26, 2015].

Reich, M., Gordon, D.M. & Edwards, R.C., 1973. Dual Labor Markets : A Theory of Labor Market Segmentation. *American Economic Review*, 63(2), pp.359–365. Available at: http://www.jstor.org/stable/1817097.

Rogerson, C.M., 2000. Emerging from Apartheid's Shadow: South Africa's Informal Economy. *Journal of International Affairs*, 53(2), pp.673–697.

Santarelli, E. & Vivarelli, M., 2006. Entrepreneurship and the Process of Firms ' Entry , Survival and Growth. , (2475). Available at: http://icc.oxfordjournals.org/content/16/3/455.abstract.

Statistics South Africa, 2001. Labour Force Survey Metadata. , (September), pp.1–35.

Tryon, R.C., 1939. *Cluster analysis: correlation profile and orthometric (factor) analysis for the isolation of unities in mind and personality*, Edwards brother, Incorporated, lithoprinters and publishers.

Van der Berg, S. & Bredenkamp, C., 2002. Devising social security interventions for maximum poverty impact. *Social Dynamics*, 28(2), pp.39–68. Available at: http://dx.doi.org/10.1080/02533950208458732.

Van der Berg, S., Siebrits, K. & Lekezwa, B., 2010. *Efficiency and equity effects of social grants in South Africa*, Available at: http://ssrn.com/abstract=1727643.

Weeks, J., 1975. Policies for Expanding Employment in the Informal Urban Sector of Developing Economies. *International Labor Review*, 111(1), pp.1–13.

Appendix

Table A1: Regressions of women who were out of work in the previous six months

	Logit		Logit	
	Survivalist	**(FE)**	**Growth oriented**	**(FE)**
HH income (-1)	-0.160	-0.022	0.124	1.380
	(4.24)***	(0.13)	(1.33)	(1.65)*
Primary	0.041		-0.009	
	(2.60)***		(0.17)	
Secondary	-0.022		0.303	
	(0.78)		(4.35)***	
Tertiary	-0.427		0.236	
	(1.28)		(1.04)	
Age	0.174	0.306	0.195	0.195
	(8.60)***	(2.26)**	(3.54)***	(0.27)
Age squared	-0.002	-0.003	-0.002	-0.002
	(8.71)***	(2.15)**	(3.55)***	(0.16)
Coloured	-1.111		-0.605	
	(5.36)***		(1.39)	
Indian	-1.687		-0.078	
	(3.65)***		(0.16)	
White	-1.323		-1.377	
	(4.14)***		(2.69)***	
Head	0.801		0.476	
	(7.87)***		(1.48)	
Married	0.472	-0.075	0.556	0.377
	(5.68)***	(0.13)	(2.49)**	(0.24)
HH size	-0.040	-0.087	-0.019	0.402
	(2.88)***	(1.24)	(0.51)	(1.13)
Constant	-5.842		-10.028	
	(11.15)***		(7.46)***	
N	21,412	412	20,691	43

* $p<0.1$; ** $p<0.05$; *** $p<0.01$
Base category: Black men who are out of work
Logged per capita household income in 2012 prices
Primary, secondary, tertiary: splines of years completed schooling
Coefficients for province, rural, and period are not shown

Table A2: Regressions of men who were out of work in the previous six months

	Logit		Logit	
	Survivalist	**(FE)**	**Growth oriented**	**(FE)**
HH income (-1)	-0.131	-0.042	-0.013	0.357
	(2.58)***	(0.14)	(0.20)	(0.92)
Primary	0.049		0.066	
	(2.12)**		(2.08)**	
Secondary	-0.098		-0.118	
	(2.73)***		(2.52)**	
Tertiary	-0.505		-0.197	
	(1.06)		(0.51)	
Age	0.124	0.363	0.172	0.041
	(4.90)***	(1.83)*	(5.04)***	(0.17)
Age squared	-0.002	-0.004	-0.003	-0.002
	(6.14)***	(1.86)*	(6.20)***	(0.54)
Coloured	-0.264		0.144	
	(1.24)		(0.56)	
Indian	-0.209		0.845	
	(0.47)		(2.15)**	
White	-1.363		0.816	
	(2.29)**		(2.46)**	
Head	0.703		0.669	
	(4.36)***		(3.26)***	
Married	0.777	1.025	0.955	17.149
	(5.33)***	(0.89)	(5.14)***	(0.01)
HH size	-0.047	-0.206	-0.056	0.121
	(2.54)**	(1.50)	(2.21)**	(0.51)
Constant	-4.773		-6.583	
	(7.39)***		(7.76)***	
N	12,086	189	11,886	97

* $p<0.1$; ** $p<0.05$; *** $p<0.01$
Base category: Black men who are out of work
Logged per capita household income in 2012 prices
Primary, secondary, tertiary: splines of years completed schooling
Coefficients for province, rural, and period are not shown

The **Research Project on Employment, Income Distribution and Inclusive Growth (REDI3x3)** is a multi-year collaborative national research initiative. The project seeks to address South Africa's unemployment, inequality and poverty challenges.

It is aimed at deepening understanding of the dynamics of employment, incomes and economic growth trends, in particular by focusing on the interconnections between these three areas.

The project is designed to promote dialogue across disciplines and paradigms and to forge a stronger engagement between research and policy making. By generating an independent, rich and nuanced knowledge base and expert network, it intends to contribute to integrated and consistent policies and development strategies that will address these three critical problem areas effectively.

Collaboration with researchers at universities and research entities and fostering engagement between researchers and policymakers are key objectives of the initiative.

The project is based at SALDRU at the University of Cape Town and supported by the National Treasury.

Consult the website for further information.

Tel: (021) 650-5715

www.REDI3x3.org

REDI3X3

Research Project on Employment, Income Distribution & Inclusive Growth

Towards an integrated response to unemployment, inequality and poverty