

---

# Alternative definitions of informal sector employment in South Africa

HASSAN ESSOP AND DEREK YU

---

Stellenbosch Economic Working Papers: 21/08

KEYWORDS: SOUTH AFRICA, HOUSEHOLD SURVEY, LABOUR MARKET TRENDS, INFORMAL  
SECTOR  
JEL: J00

HASSAN ESSOP  
DEPARTMENT OF ECONOMICS  
UNIVERSITY OF STELLENBOSCH  
PRIVATE BAG X1, 7602  
MATIELAND, SOUTH AFRICA  
E-MAIL: HESSOP@SUN.AC.ZA

DEREK YU  
DEPARTMENT OF ECONOMICS  
UNIVERSITY OF STELLENBOSCH  
PRIVATE BAG X1, 7602  
MATIELAND, SOUTH AFRICA  
E-MAIL: DEREKY@SUN.AC.ZA



A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE  
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

---

# Alternative definitions of informal sector employment in South Africa

HASSAN ESSOP AND DEREK YU<sup>1</sup>

---

## ABSTRACT

---

Before the introduction of the Quarterly Labour Force Survey (QLFS) in 2008, Statistics South Africa (Stats SA) has been using the same methodology to derive the informal sector employment throughout the years, focusing on the enterprise registration status to classify workers (which include both self-employed and employees) as either formal or informal sector workers. Although there are difficulties with attempting to provide any consistent trend data (Yu, 2007 & Essop & Yu, 2008), it is generally accepted that informal sector employment grew relatively more rapidly in the late 1990s, and then stabilized at about 2 million in the early 2000s before it increased (albeit more slowly) again since 2005.

Nonetheless, recent papers by Devey, Skinner & Valodia (2006) as well as Heintz & Posel (2008) argue that the current classifications used by Stats SA hide a significant degree of informality in the formal economy, as some formal jobs are characterized by conditions that are typical of informal work. Therefore, they propose alternative definitions of informal sector employment, focusing on worker characteristics instead of enterprise characteristics. This paper aims to address the reliability or otherwise of these recent approaches, as well as to suggest better ways to define informal sector employment.

Keywords: South Africa, Household survey, Labour market trends, Informal sector  
JEL codes: J00

---

<sup>1</sup> The authors gratefully acknowledge the valuable comments by Servaas van der Berg.

# Alternative definitions of informal sector employment in South Africa

## 1. INTRODUCTION

The latest Labour Force Survey (LFS) shows that the narrow unemployment rate in South Africa has declined from 23.5 per cent in the first quarter of 2008 to 23.1 per cent in the second quarter of 2008. Even though this decline in unemployment is heartening, the fact remains that South Africa still has approximately 4.1 million unemployed working-age individuals. Traditionally, the informal sector in a developing country is seen as a possible alternative when employment in the formal labour market is hard to come by (Fields, 1975 and Mazumdar, 1976)<sup>2</sup>, and given the large number of unemployed, it is typically expected that South Africa should have a relatively large informal sector (Kingdon & Knight, 2004 & 2007)<sup>3</sup>. Consequently, the size and characteristics of the informal sector becomes important to policy makers and researchers alike.

Additionally, Henley, Reza Arabsheibani & Carneiro (2006: 4) provide three additional reasons why policy makers, amongst others, should be concerned about the size of the informal sector. Firstly, the informal sector can facilitate the development of a micro-entrepreneurial sector which, in turn, can enhance economic efficiency; secondly, policy makers need to be aware of the number of workers with little or no employment or other social security, such as medical aid; and finally, policy makers need to understand the informal sector's dimensions in order to achieve the long term goal of shifting informal sector participants to the formal sector, with the ultimate aim being to broaden the tax base.

However, defining the informal sector, and its ensuing measurement, has been problematic, both internationally and domestically. Failure to define and measure the informal sector in an appropriate manner, of course, hampers the ability of policy makers to address the points noted above.

In South Africa, before the introduction of the Quarterly Labour Force Survey (QLFS) in 2008, Statistics South Africa (Stats SA) has been consistent, using the same methodology to derive the informal sector employment in the October Household Surveys (OHSs) as well as the LFSs, focusing on the enterprise registration status to classify workers (which include both self-employed and employees) as either formal or informal sector workers. Although there are difficulties with attempting to provide any consistent trend data (Yu, 2007 and Essop & Yu, 2008), it is generally accepted that informal sector employment grew relatively more rapidly in the late 1990s, and then stabilized at about 2 million in the early 2000s before it increased (albeit more slowly) again since 2005.

Nonetheless, recent papers by Devey, Skinner & Valodia (2006) as well as Heintz & Posel (2008) argue that the current classifications used by Stats SA hide a significant degree of informality in the formal economy, as some formal jobs are characterized by conditions that are typical of informal work. Therefore, the aforementioned authors propose alternative definitions of informal sector employment, focusing on worker characteristics instead of enterprise characteristics. One of the aims of this paper will therefore be to consider the benefits and shortcomings of these recent approaches, and to propose alternative methods to define informal sector employment.

Furthermore, recent work by Kingdon & Knight (2004 & 2007) suggest that South Africa, with low informal sector non-agricultural employment but high unemployment, is an international outlier in the

---

<sup>2</sup> Other interpretations for the existence of the informal sector exist; however, such a discussion falls beyond the scope of this paper. Henley *et al* (2006) provide a brief overview of alternative views, as well as references to authors who have covered these issues in greater depth.

<sup>3</sup> South Africa has a relatively small informal sector, contrary to what is expected of a developing country with large unemployment (see Essop & Yu (2008) for greater detail on the trends and characteristics within the informal sector in South Africa).

size of its informal sector, as indicated by the low ratio of informal sector non-agricultural employment to unemployment. Although they hold the view that such a low ratio is caused by barriers to entry such as crime, lack of access to credit, lack of access to infrastructure and services, etc., they also argue that the low ratio is partly caused by the narrower definition of the informal sector used by Stats SA (Kingdon & Knight, 2007: 824). Thus, in addition to the aim noted above, the real size of the informal sector using the alternative definitions mentioned above will be compared to the current method utilised by Stats SA.

The paper is structured as follows: Section 2 reviews the definition of the informal sector used by Stats SA before 2008, as well as a short discussion on the new definition to be adopted by Stats SA with the introduction of QLFS from 2008, while the Devey *et al.* formal-informal index is analysed in Section 3. The alternative definition suggested by Heintz & Posel is discussed in Section 4, while the revised Devey *et al.* formal-informal index is the focus of Section 5. In Section 6, a comparative analysis of informal sector employment using the various definitions discussed in Sections 3-5 is presented. Section 7 discusses other possible techniques to measure the size of the informal sector, focusing on the currency demand approach. Finally, Section 8 concludes the paper<sup>4</sup>.

---

<sup>4</sup> The OHS and the LFS data were used for the analysis herein. For the remainder of the paper, the OHSs conducted between 1995 and 1999 will be referred to as OHS1995, OHS1996, etc., while the LFSs from 2000 to 2007 will be referred to as LFS2000a (March 2000), LFS2000b (September 2000), LFS2001a, and so forth. In addition, the data from OHS1995 to LFS2000a are weighted using the 1996 census weights, while data from LFS2000b to LFS2007b are weighted using the 2001 census weights.

## 2. INFORMAL SECTOR DEFINITION BY STATS SA

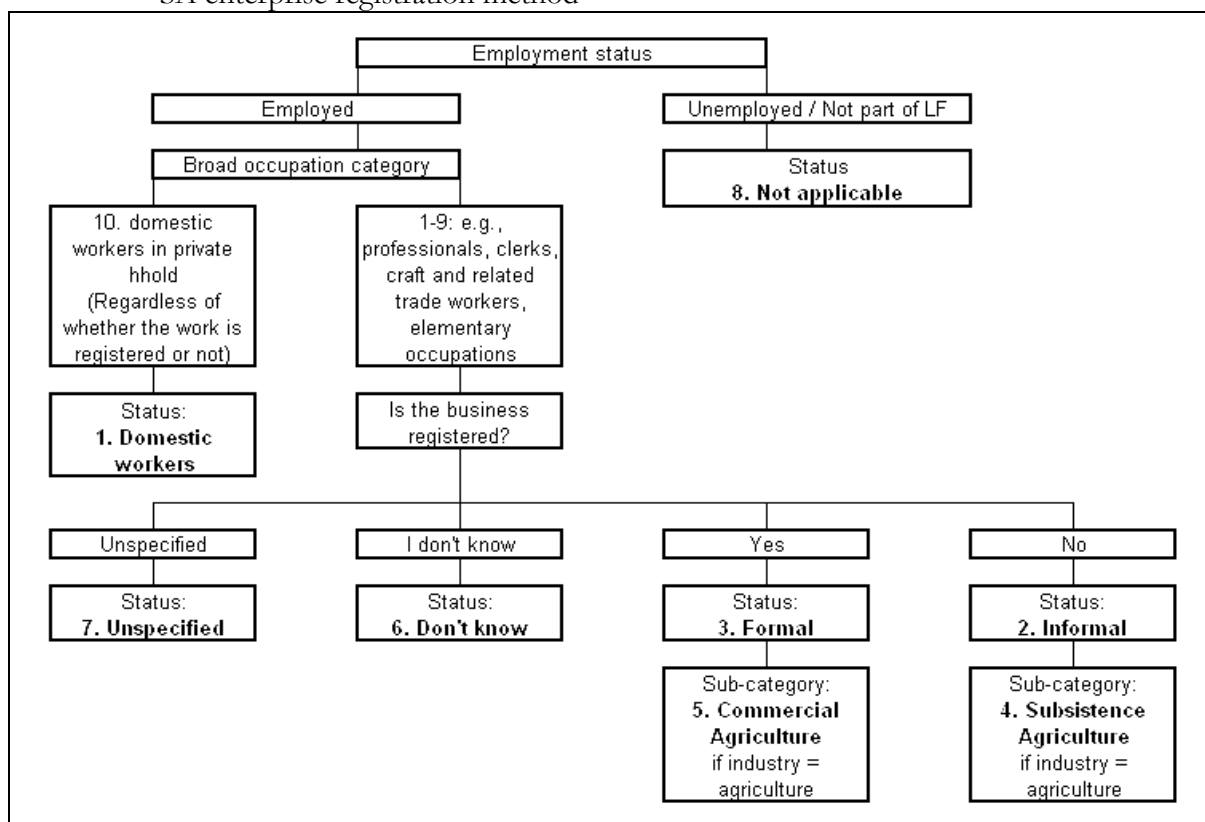
### 2.1 The definition used before 2008

Stats SA has been using the same methodology to measure informal sector employment for the duration of the OHS and the LFS until LFS2007b, focusing on whether an enterprise is registered according to legislation. Further, using a stepwise approach, several questions from the questionnaire are involved to determine the different categories of workers. Firstly, the employment status of the respondent is determined<sup>5</sup>. Next, if the broad occupation category of the employed is ‘domestic workers in the private households’, he/she is grouped under the category ‘domestic workers’, which is an independent category that falls under neither the formal sector nor the informal sector.

The other employed, whose occupation is something other than domestic worker, is classified as either formal or informal sector workers, according to his/her answer on the question concerning the registration of the enterprise. If the respondent does not answer the question, he/she is shifted to the category ‘unspecified’. On the other hand, if the respondent’s answer is ‘I don’t know’<sup>6</sup>, he/she falls under the category ‘don’t know’.

Finally, if the broad industry category of the formal sector worker is agriculture, he/she will be classified as a commercial agriculture worker. On the other hand, if the broad industry category of the informal sector worker is agriculture, he/she will be classified as a subsistence agriculture worker. Figure 1 summarizes the methodology.

Figure 1 Derivation of the different categories of formal and informal sector workers, Stats SA enterprise registration method



<sup>5</sup> The questions on employment as well as the methodology to derive employment status have changed substantially throughout the OHS/LFS surveys. They are explained in the metadata of the surveys as well as in Yu (2007).

<sup>6</sup> The option “don’t know” only became available after LFS2000a.

For the remainder of this paper, unless stated otherwise, ‘informal sector’ means informal sector less subsistence agriculture, ‘formal sector’ stands for formal sector less commercial agriculture, and ‘non-agricultural employment’ means informal sector employment plus formal sector employment (i.e., the categories ‘domestic workers’, ‘subsistence agriculture’, ‘commercial agriculture’, ‘don’t know’ and ‘unspecified’ are excluded).

Table 1 below shows the breakdown of total employment since 1995, using the Stats SA categorization methodology discussed above. Looking at the informal sector employment, it can be seen from Figure 2 that, with the exception of the serious under-estimation in the OHSs (especially OHS1995 and OHS1996)<sup>7</sup> as well as the over-estimation in LFS2001a<sup>8</sup>, the informal sector employment enjoyed an upward trend during the OHSs<sup>9</sup>, and then it stabilized at about 2 million between LFS2000a and LFS2005a. LFS2005b saw a large increase in informal sector employment to nearly 2.46 million, after which informal employment declined to just under 2.1 million in LFS2007b. Furthermore, the informal sector employment’s contribution to total employment has declined over time; with informal sector employment contributing approximately 20% of non-agricultural employment since LFS2002a (see Essop & Yu 2008 for a more in-depth analysis).

Table 1 Breakdown of total employment, 1995 – 2007

	Domestic workers	Informal	Formal	Subsistence agriculture	Commercial agriculture	Don't know	Not specified	Total employed
OHS1995	695 416	521 668	219 213	26 530	49 546	0	7 986 974	9 499 347
OHS1996	766 334	330 100	304 260	24 687	56 296	0	7 484 630	8 966 307
OHS1997	828 254	1 043 347	6 436 017	187 486	525 618	0	72 925	9 093 647
OHS1998	747 281	1 077 141	6 508 097	202 082	725 474	0	110 055	9 370 130
OHS1999	812 465	1 571 646	6 796 008	284 336	798 905	0	92 783	10 356 143
LFS2000a	1 002 719	1 819 556	6 672 951	1 507 625	756 510	86 472	28 576	11 874 409
LFS2000b	941 463	2 026 065	7 077 307	1 074 413	766 917	108 318	229 923	12 224 406
LFS2001a	844 135	2 836 182	6 798 257	742 404	784 712	214 235	40 282	12 260 207
LFS2001b	881 168	1 964 763	7 019 158	382 241	764 521	127 023	28 667	11 167 541
LFS2002a	875 172	1 821 426	7 089 163	862 747	864 576	74 868	15 446	11 603 398
LFS2002b	843 019	1 778 542	7 173 080	550 068	851 897	61 643	25 675	11 283 924
LFS2003a	885 322	1 827 711	7 223 138	443 426	841 440	57 332	19 252	11 297 621
LFS2003b	894 626	1 901 131	7 364 616	365 378	831 526	36 403	17 671	11 411 351
LFS2004a	845 965	1 764 630	7 473 638	340 515	912 831	25 704	14 934	11 378 217
LFS2004b	880 067	1 944 236	7 684 843	425 083	624 358	52 970	18 639	11 630 196
LFS2005a	848 914	2 068 479	7 741 991	513 022	647 448	27 756	46 710	11 894 320
LFS2005b	858 199	2 459 690	7 979 587	337 884	578 059	33 783	40 596	12 287 798
LFS2006a	849 085	2 187 940	8 051 532	702 881	605 795	14 098	26 632	12 437 963
LFS2006b	884 898	2 376 338	8 376 441	472 697	605 129	46 935	24 847	12 787 285
LFS2007a	935 642	2 129 164	8 414 719	459 509	602 942	52 537	40 383	12 634 896
LFS2007b	1 024 039	2 083 855	9 034 135	368 256	666 533	47 251	69 258	13 293 327

Furthermore, Table 2 presents the employment type<sup>10</sup> of informal sector workers from 1997 onwards, and it is evident that the upward trend of informal sector employment during the late OHSs was mainly caused by the improvements in identifying and the collection of self-employment data. Moreover, since the changeover from the OHS to the LFS, the data reveal a trend that initially declines from LFS2000b, before increase from LFS2004a, followed by a decline by the end of LFS2007b. In addition, the LFS

<sup>7</sup> The under-estimation for the OHSs is caused by the fact that self-employment was not well-captured. However, with the improvement of the survey questions, self-employment was eventually captured better throughout the OHS years (Yu, 2007: 17-18). Besides, in OHS of 1995 and 1996, only the self-employed had to answer the question on enterprise registration (Essop & Yu, 2008: 7-8). Thus, it is impossible to determine the sector status of the employees, and subsequently, it is impossible to accurately derive the total number of informal sector workers in 1995 and 1996.

<sup>8</sup> Devey *et al.* (2006: 307-309) and Essop & Yu (2008: 53-54) explain the reasons for over-estimation of informal sector employment in LFS2001a in great detail.

<sup>9</sup> The upward trend is mainly due to the improvements in capturing self-employed informal and low-income employment data, which stemmed directly from the progress Stats SA’s made in the employment-related questions throughout the period concerned (Yu, 2007: 17).

<sup>10</sup> Essop & Yu (2008: 12) explain how the self-employed are distinguished from the employees throughout the years.

series consistently finds that approximately one-third of the informal sector workers are employees, while the remaining two-thirds are self-employed.

Figure 2 Informal sector employment, 1997 – 2007, excluding LFS2001a

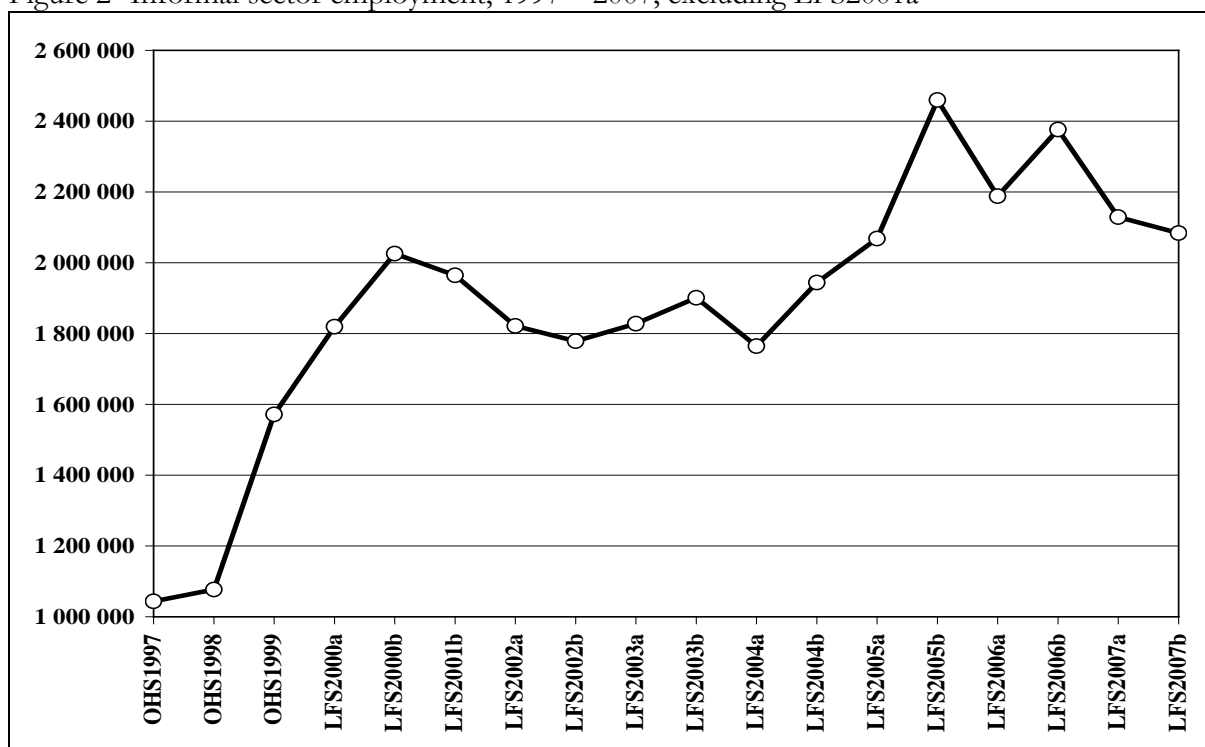


Table 2 Employment type of informal sector workers, 1997 – 2007

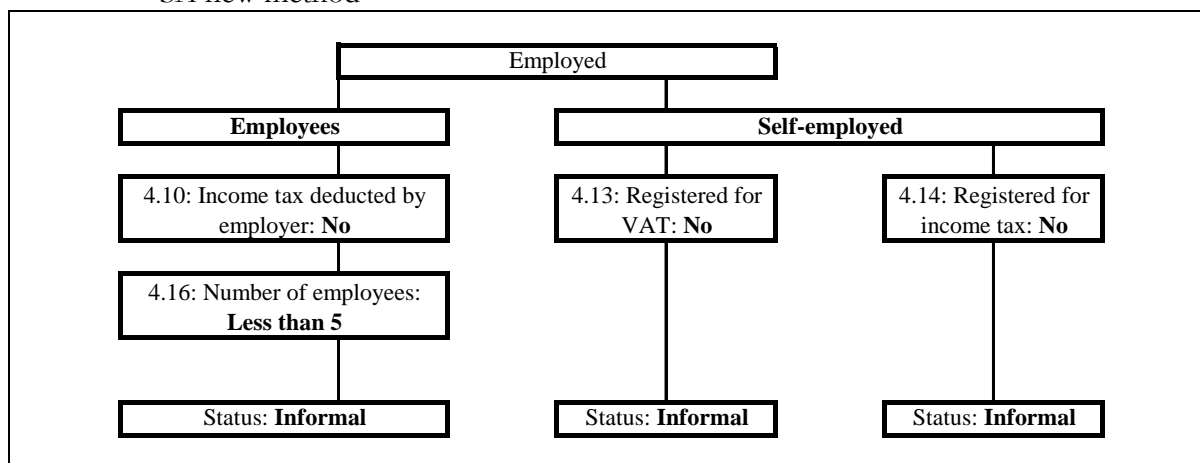
	Employee		Self-employed		Unspecified		Total	
	Number	%	Number	%	Number	%	Number	%
OHS1997	517 761	49.6%	525 586	50.4%	0	0.0%	1 043 347	100.0%
OHS1998	486 185	45.1%	590 956	54.9%	0	0.0%	1 077 141	100.0%
OHS1999	684 908	43.6%	886 738	56.4%	0	0.0%	1 571 646	100.0%
LFS2000a	607 441	33.4%	1 211 650	66.6%	465	0.0%	1 819 556	100.0%
LFS2000b	740 677	36.6%	1 284 252	63.4%	1 136	0.1%	2 026 065	100.0%
LFS2001a	776 680	27.4%	2 058 695	72.6%	807	0.0%	2 836 182	100.0%
LFS2001b	633 205	32.2%	1 330 568	67.7%	990	0.1%	1 964 763	100.0%
LFS2002a	585 946	32.2%	1 235 480	67.8%	0	0.0%	1 821 426	100.0%
LFS2002b	553 441	31.1%	1 225 101	68.9%	0	0.0%	1 778 542	100.0%
LFS2003a	619 645	33.9%	1 207 748	66.1%	318	0.0%	1 827 711	100.0%
LFS2003b	625 345	32.9%	1 275 786	67.1%	0	0.0%	1 901 131	100.0%
LFS2004a	576 490	32.7%	1 188 140	67.3%	0	0.0%	1 764 630	100.0%
LFS2004b	619 352	31.9%	1 324 532	68.1%	352	0.0%	1 944 236	100.0%
LFS2005a	757 388	36.6%	1 311 091	63.4%	0	0.0%	2 068 479	100.0%
LFS2005b	870 047	35.4%	1 589 643	64.6%	0	0.0%	2 459 690	100.0%
LFS2006a	712 459	32.6%	1 475 481	67.4%	0	0.0%	2 187 940	100.0%
LFS2006b	794 486	33.4%	1 581 852	66.6%	0	0.0%	2 376 338	100.0%
LFS2007a	753 548	35.4%	1 375 616	64.6%	0	0.0%	2 129 164	100.0%
LFS2007b	667 811	32.1%	1 416 044	67.9%	0	0.0%	2 083 855	100.0%

## 2.2 New definition introduced in 2008

With the inception of the QLFS from 2008, Stats SA also decided to make several changes to the questionnaire itself. One of these changes includes a new definition of informal sector employment, which is summarized in Figure 3.

When the self-employed are considered, the new Stats SA method defines them as informal sector workers if their businesses are not registered for either income tax **or** value-added tax. On the other hand, the employees are classified as informal sector workers if they are not registered for income tax **and** work in establishments that employ fewer than 5 employees.

Figure 3 Derivation of the different categories of formal and informal sector workers, Stats SA new method



Note: The question number refers to the QLFS2008 questionnaire.

It is difficult to use this newly adopted 2008 method to derive the informal sector employment in 1995-2007, due to the following reasons:

- Before 2008, the firm size question was asked as ‘How many regular workers has the organization/business/enterprise where ... works, including him/herself’, so the self-employed could also be included. But in 2008, the question clearly states that only the employees are counted.
- The VAT registration question was only asked since LFS2001a.
- The income tax registration question was only asked in LFS2005b, LFS2006b – LFS2007b. Besides, there was only one question asked to both self-employed and employees, ‘Is the organization/business/enterprise/branch where ...works registered for income tax?’ However, Figure 3 above clearly shows that in the quarterly survey, there are two questions on income tax registration, one to employees (‘Does your employer deduct income tax (PAYE/SITE) from your salary/wage?’) and one to self-employed (‘Is the business or household business where you work registered for income tax?’).

Nonetheless, it should be noted that the focus of Stats SA’s method remains on the registration status of the firm, whilst adding the criterion on “smallness” for all firms in the informal sector.

For the remainder of the paper, the enterprise registration method adopted until the end of 2007 will be compared with numerous recently proposed alternative methods, to be discussed in Sections 3-5.



### **3. THE DEVEY, SKINNER & VALODIA FORMAL-INFORMAL INDEX**

Instead of focusing on the enterprise registration status as in the Stats SA method, Devey *et al.* (2006: 315 – 316) focus on the worker characteristics and use 13 indicators as shown in Table 3 to derive a formal-informal index<sup>11</sup>. The indicators used for the index are not weighted. In other words, the most formal worker would achieve a score of 13 for the index while the most informal worker would attain a score of zero. The aim is to find out the proportion of informal sector workers displaying formal-sector characteristics, as well as the proportion of formal sector workers with informal-sector characteristics

Table 3 The indicators used to derive the Devey *et al.* formal-informal index

<b>Question number***</b>	<b>Index = 1</b>	<b>Index = 0</b>
4.4: Number of employers	(1): One employer (2): More than one employer	???
4.6: Permanence of work	(1): Permanent	(2): Fixed period contract (3): Temporary (4): Casual (5): Seasonal
4.8: Written contract with employer	(1): Yes	(2): No
4.10: Who pays wage	(1): Employer (2): Labour broker (3): Contractor or agency	(4): Other
4.11: Employer contributes to pension of retirement fund	(1): Yes	(2): No
4.12: Paid leave	(1): Yes	(2): No
4.13: Membership of trade union	(1): Yes	(2): No
4.16 Number of regular workers in enterprise	(6): 50 or more	(1): 1 (2): 2 – 4 (3): 5 – 9 (4): 10 – 19 (5): 20 – 49
4.17: Working for a registered company or close corporation	(1): Yes	(2): No
4.18: Employer makes UIF deductions	(1): Yes	(2): No
4.19: Employer makes medical aid or health insurance payments	(1): Yes, for himself only (2): Yes, for himself and his dependents (3): Yes, but he is not using it	(4): No, because he is covered by someone else's medical aid (5): No medical aid benefits provided
4.20: Enterprise is registered to pay VAT	(1): Yes	(2): No
4.23: Location of work	(3): Inside a formal business premises (4): At a service outlet	(1): In the owner's home (2): In someone else's home (5): At a market (6): On a footpath or street (7): No fixed location (8): Other

\*\*\* The question number refers to the LFS2007b questionnaire.

Table 4 Formal-informal index for formal and informal workers by Devey *et al.*, LFS2004a

	<b>Status (Using the Stats SA enterprise registration methodology)</b>	
<b>Index</b>	<b>Formal sector employees</b>	<b>Informal sector employees</b>

<sup>11</sup> The decision to focus on worker characteristics is in line with the latest view of the ILO taken at the 17th International Conference of Labour Statisticians (see Essop & Yu, 2008 and Devey *et al.*, 2006).

score	Number of people	Percentage	Cumulative percentage	Number of people	Percentage	Cumulative percentage
0	574 626	7.3%	7.3%	398	0.0%	0.0%
1	1 205 941	15.4%	22.7%	5 126	0.3%	0.3%
2	1 333 428	17.0%	39.8%	7 714	0.4%	0.7%
3	1 341 682	17.1%	56.9%	7 561	0.4%	1.1%
4	939 984	12.0%	68.9%	12 491	0.7%	1.8%
5	589 071	7.5%	76.5%	8 250	0.4%	2.3%
6	404 610	5.2%	81.6%	15 689	0.9%	3.1%
7	373 774	4.8%	86.4%	23 055	1.3%	4.4%
8	383 909	4.9%	91.3%	46 482	2.5%	6.9%
9	251 509	3.2%	94.5%	67 655	3.7%	10.6%
10	226 719	2.9%	97.4%	160 172	8.7%	19.3%
11	133 597	1.7%	99.1%	265 126	14.5%	33.8%
12	41 353	0.5%	99.7%	106 194	5.8%	39.6%
13	27 048	0.3%	100.0%	1 107 701	60.4%	100.0%
Total <sup>12</sup>	7 827 251	100.0%		1 833 614	100.0%	

Source: Devey *et al.*, 2006: 316

Using this methodology, Devey *et al.* derived the formal-informal index of the formal and informal sector workers using LFS2004a. The results are presented in Table 4 above. Although these results impart a new dimension to the analysis of the informal sector in South Africa, a careful review reveals the following problems:

- (1) In each of the 13 questions, the respondent's answer could be 'I don't know' or 'unspecified', but Devey *et al.* did not provide any explanation on whether the respondent is assigned a mark of 0 or 1 for each question. For example, if the respondent's answer on the paid leave question (4.12) is '(9): Unspecified', is he/she assigned a mark of 0 or 1?
- (2) In question 4.4 (Number of employers), there are only 2 options available: '(1): One employer' or '(2): More than one employer'. Devey *et al.* allocated a mark of 1 to both options, but this seems to imply that as long as the employed gave a definite answer, they are assured to get 1 mark (99.82% of employees specified their answer in LFS2004a). This also implies that only those (a mere 0.18% of employees) who did not specify their answers on this question will be given 1 mark. Devey *et al.* (2006: 321) only mentioned that if the respondent's answer is 'other' in 4.4, the mark is 0, but no explanation is given on what 'other' stands for.
- (3) If the respondent is self-employed, he/she is not asked to answer the first 7 questions on Table 3 and his/her answers for all these questions are coded as '(8): No applicable'. However, Devey *et al.* did not provide any explanation on whether the self-employed is given 0 or 1 mark in each of these indicators.
- (4) Another problem with this method is the comparability of scores. Two individuals who obtain the same scores may have selected different answers. However, does that mean that someone with a pension fund, medical aid and a contract can have the same sector status as someone with a permanent job, who works with several other workers and has paid leave? Also, the question as to which characteristics are more important is not considered.

Considering the third problem in greater detail, if Devey *et al.* assume 0 mark in each of these 7 indicators (which is more likely, since the self-employed are not required to answer these questions), then even if a self-employed worker gets 1 mark on each of the remaining 6 questions, his total index score could only be 6. Consequently, one could mistakenly think that the self-employed individual

<sup>12</sup> The employment figures by Devey *et al.* (formal sector employment: 7 827 251; informal sector employment: 1 833 614) are different from the figures in this paper (formal sector employment: 7 473 638; informal sector employment: 1 764 630), because the LFS2000b – LFS2005a data were re-weighted using the Census 2001 weight only after the Devey *et al.* article was released. Consequently, it is likely that the LFS2004a data used by Devey *et al.* were still weighted using Census 1996 weights.

displays strong informal-sector characteristics due to the low overall index score. Therefore, it seems that if all 13 indicators are used to derive the index, the index would only be useful to distinguish informal sector **employees**, since only the latter are able to answer all 13 questions. However, Devey *et al.* definitely included **both** self-employed and employees in Table 4 above<sup>13</sup>.

Table 5 presents the Devey *et al.* formal-informal index for employees only, using the LFS2004a data weighted using the Census 2001 weights. Additionally, if the respondent's answer on the question concerned is 'I don't know' or 'unspecified', a 0 mark is assumed for each indicator. In other words, the Devey *et al.* index is revised after taking the first and third problems mentioned above into consideration. Also, depending on the index score one selects to define informal sector participation, the number of employees in the informal sector can vary, as can be seen in the last column of Table 5.

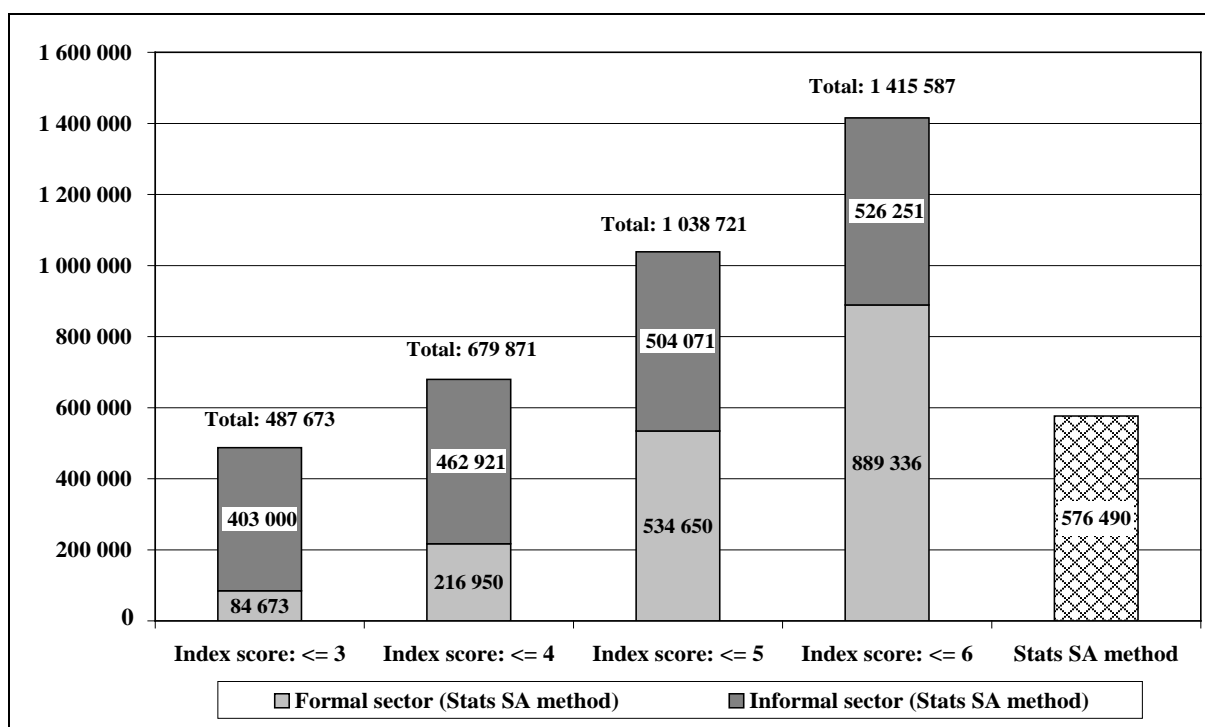
Table 5 Formal-informal index for formal and informal sector employees (after taking the first and third problems into consideration), using the Devey *et al.* methodology, LFS2004a

Index score	Status (Using the Stats SA enterprise registration methodology)						Total number of informal sector employees for all index values
	Formal sector employees			Informal sector employees			
	Number of people	%	Cumulative %	Number of people	%	Cumulative %	
0	163	0.0%	0.0%	1 463	0.3%	0.3%	1 626
1	2 990	0.0%	0.0%	4 450	0.8%	1.0%	9 066
2	17 591	0.3%	0.3%	258 030	44.8%	45.8%	284 687
3	63 929	0.9%	1.2%	139 057	24.1%	69.9%	487 673
4	132 277	1.9%	3.1%	59 921	10.4%	80.3%	679 871
5	317 700	4.6%	7.7%	41 150	7.1%	87.4%	1 038 721
6	354 686	5.1%	12.8%	22 180	3.8%	91.3%	1 415 587
7	394 000	5.7%	18.5%	12 489	2.2%	93.5%	1 822 076
8	578 860	8.3%	26.9%	8 996	1.6%	95.0%	2 409 932
9	937 298	13.5%	40.4%	11 301	2.0%	97.0%	3 358 531
10	1 261 047	18.2%	58.6%	6 267	1.1%	98.1%	4 625 845
11	1 244 494	17.9%	76.5%	6 519	1.1%	99.2%	5 876 858
12	1 092 388	15.8%	92.3%	4 364	0.8%	99.9%	6 973 610
13	536 504	7.7%	100.0%	303	0.1%	100.0%	7 510 417
Total	6 933 927	100.0%		576 490	100.0%		

As an example, if one makes an assumption that the employee with overall index value equal to or smaller than 5 is classified as an informal sector worker, then 7.7% of employees (534 650 people in total) defined as formal sector workers under the Stats SA enterprise registration methodology are better regarded as informal sector workers. Similarly, 87.4% of employees (504 071 people in total) defined as informal sector workers under the Stats SA methodology are still regarded as informal sector workers under the formal-informal index methodology. Therefore, the total number of informal sector employees in LFS2004a under the formal-informal index methodology would be 1 038 721 (compared with 576 490, under the Stats SA methodology), as shown in Figure 4.

Figure 4 Number of informal sector employees, using the Devey *et al.* formal-informal index, LFS2004a

<sup>13</sup> In Table 15.1 of Devey *et al.* (2006: 304), the total formal and informal sector employment figures for LFS2004a are exactly the same as those in Table 4 in this paper. From this it can be concluded that Devey *et al.* did not exclude some of the employed for the formal-informal sector index.



In addition, Table 6 compares the number of informal sector employees using the Stats SA definition and the Devey *et al.* methodology since LFS2001b, excluding LFS2004b<sup>14</sup>.

Table 6 Number of informal sector employees, Stats SA method vs. Devey *et al.* method 2001 – 2007

	Stats SA method	Devey <i>et al.</i> method			
		Informal sector if index value: ≤ 3	Informal sector if index value: ≤ 4	Informal sector if index value: ≤ 5	Informal sector if index value: ≤ 6
LFS2001b	633 205	547 509	787 598	1 192 045	1 644 722
LFS2002a	585 946	526 278	730 781	1 083 369	1 468 833
LFS2002b	553 441	486 071	692 058	1 020 960	1 391 866
LFS2003a	619 645	516 864	744 879	1 091 022	1 475 623
LFS2003b	625 345	495 921	666 530	1 032 132	1 395 978
LFS2004a	576 490	487 673	679 871	1 038 721	1 415 587
LFS2005a	757 388	584 770	813 437	1 174 306	1 616 530
LFS2005b	870 047	696 910	965 023	1 424 220	1 616 352
LFS2006a	712 459	575 414	824 317	1 241 439	1 862 828
LFS2006b	794 486	675 105	918 664	1 364 107	1 710 958
LFS2007a	753 548	582 463	853 469	1 328 726	1 836 649
LFS2007b	667 811	518 838	775 185	1 223 554	1 845 049

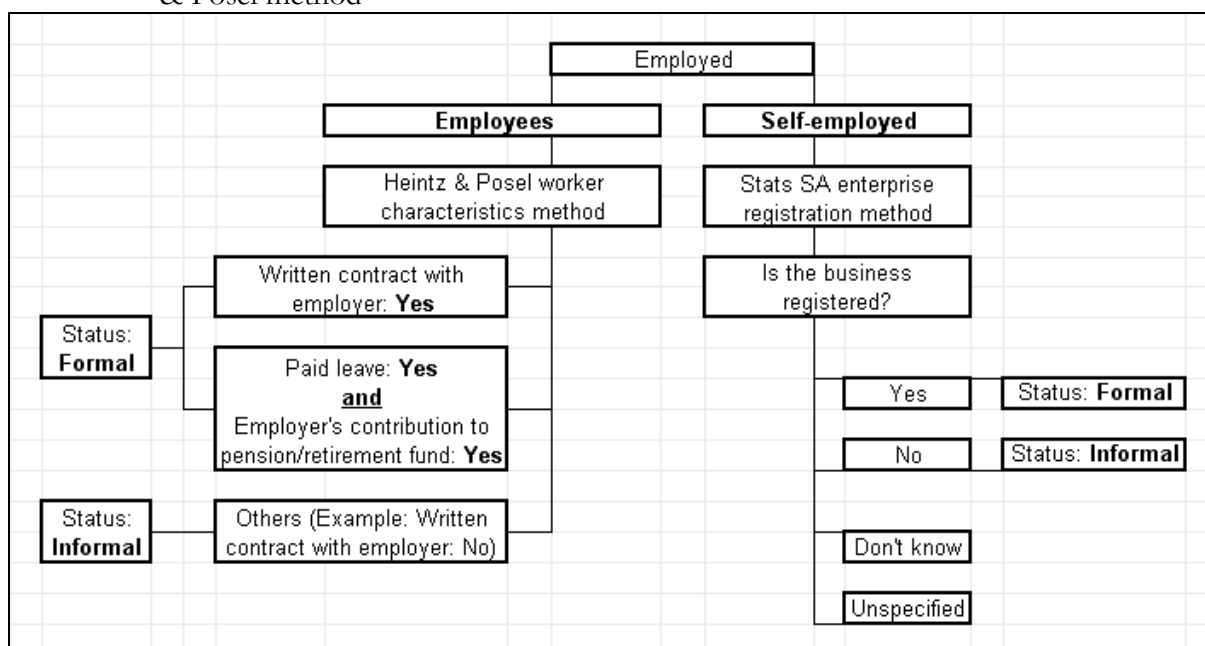
<sup>14</sup> The 13 questions for the index were asked together for the first time in LFS2001a. However, since the categorization of the options in the location of work question in LFS2001a was significantly different from the other surveys, it was decided to exclude LFS2001a from Table 6. Further, due to the coding error problem in the question on the number of regular workers in the enterprise in LFS2004b (Yu, 2007: 23), the LFS2004b result was also excluded from Table 6.

## 4. THE HEINTZ & POSEL PROPOSED DEFINITION

Heintz & Posel (2008) argue that the Stats SA enterprise registration methodology cannot estimate the number of informal sector employees correctly since the Stats SA methodology “fails to capture adequately the number of individuals working in informal jobs – that is, in forms of employment that lack legal or social protection. Informal employment occurs outside of the informal sector when individuals are employed by households (e.g., domestic workers) or when individuals are employed in unprotected jobs in formal enterprise” (Heintz & Posel, 2008: 30). However, they state that the enterprise registration methodology could still be applied to distinguish the formal self-employed from the informal self-employed, because “registration subjects formal self-employment to greater regulatory oversight, one example being stronger enforcement of tax collection” (Heintz & Posel, 2008: 32).

As a result, Heintz & Posel (2008: 32) suggest an alternative definition of informal sector employment that is based on the definition proposed in 2002 by the International Labour Organization (ILO) at the 17th International Conference of Labour Statisticians (ICLS)<sup>15</sup>, and is presented in Figure 5 below. Firstly, the Stats SA enterprise registration method is still applied to classify the self-employed as either formal or informal workers. On the other hand, an employee is considered as a formal sector worker if he/she has an employment contract or receives both paid leave and pension/retirement fund contributions from the employer. The remaining employees are then classified as informal sector workers. Note that the Heintz & Posel method could only be applied from OHS1999 onwards since all three important questions mentioned above (written contract with employer, paid leave, and pension/retirement fund contributions from employer) were only asked together since OHS1999.

Figure 5 Derivation of the different categories of formal and informal sector workers, Heintz & Posel method



This alternative definition of informality in employment produces a significantly larger measure of informal employment than generated using Stats SA’s enterprise definition as well as the Devey *et al.* formal-informal sector index, if one compares Tables 6 and 7. The number of informal sector employees using the Heintz & Posel method ranges between 1.5 and 1.7 million since LFS2002a (during OHS1999-LFS2001b, this figure stayed close to 2 million), while the Devey *et al.* method estimates similar figures only under the assumption that the employee is an informal sector worker if his formal-

<sup>15</sup> The ILO proposed that the informal economy should be seen as comprised of informal employment, without secure contracts, worker benefits or social protection), both inside and outside informal enterprises (ILO, 2002b).

informal index is equal to or smaller than 6. In contrast, the Stats SA enterprise registration method only estimates the number of informal sector employees between 0.6 and 0.8 million for the same period.

Table 7 Number of informal sector employees, Stats SA method vs. Devey *et al.* method vs. Heintz & Posel method, 1999 – 2007

	Stats SA method	Devey <i>et al.</i> method			Heintz & Posel method
		Informal sector if index value: ≤ 4	Informal sector if index value: ≤ 5	Informal sector if index value: ≤ 6	
OHS1999	684 908	n/a <sup>#</sup>			2 146 155
LFS2000a	607 441				1 921 650
LFS2000b	740 677				2 037 210
LFS2001a	776 680	n/a <sup>##</sup>			1 816 914
LFS2001b	633 205	787 598	1 192 045	1 644 722	1 905 627
LFS2002a	585 946	730 781	1 083 369	1 468 833	1 623 791
LFS2002b	553 441	692 058	1 020 960	1 391 866	1 543 593
LFS2003a	619 645	744 879	1 091 022	1 475 623	1 549 921
LFS2003b	625 345	666 530	1 032 132	1 395 978	1 413 217
LFS2004a	576 490	679 871	1 038 721	1 415 587	1 332 506
LFS2004b	619 352	n/a <sup>###</sup>			1 448 260
LFS2005a	757 388	813 437	1 174 306	1 616 530	1 505 770
LFS2005b	870 047	965 023	1 424 220	1 616 352	1 729 334
LFS2006a	712 459	824 317	1 241 439	1 862 828	1 598 062
LFS2006b	794 486	918 664	1 364 107	1 710 958	1 677 776
LFS2007a	753 548	853 469	1 328 726	1 836 649	1 715 578
LFS2007b	667 811	775 185	1 223 554	1 845 049	1 582 327

# All thirteen indicators for the Devey *et al.* method were only asked altogether since LFS2001a.

## Categorization problem in the location of work question in LFS2001a.

### Coding error in the number of regular workers question in LFS2004b.

From the above, it is evident that the Heintz & Posel method only considers three out of the thirteen worker characteristics questions used in the Devey *et al.* method. This method therefore appears to overlook the proposed definition of the 15th ICLS that includes other indicators and ‘private unincorporated enterprises (excluding quasi corporations), which produce at least some of their goods or services for sale or barter, have less than five paid employees, are not registered, and are engaged in non-agricultural activities (including professional or technical activities)’ (ILO: 2000a: 5)<sup>16</sup>. An example to illustrate this problem would be that of a car guard with a short term contract, with no additional benefits. It would be difficult to justify the classification of such an individual as a formal sector employee.

In addition, it is not clear why Heintz & Posel use paid leave and pension fund contributions as the alternative characteristics, and not, for instance, medical aid contributions or permanence of work. If someone has paid leave and medical aid, he/she can be classified as being an informal sector worker; it is not clear how this varies – in a significant manner – from someone with paid leave and a pension, who will automatically be termed a formal sector worker, if the Heintz & Posel method is applied. The selection of these characteristics appears somewhat random, and it is also not entirely clear why it is limited to only three out of the thirteen criteria noted by Devey *et al.* Some of these concerns, of course, are also relevant to the Devey *et al.* method, but their greater range of criteria allows for a more nuanced view of the informal sector, and removes the random selection of criteria problem. Given these views, the Devey *et al.* method provides a more suitable definition of the informal sector. However, the second problem noted earlier still requires adjustment. Consequently, additional improvements to the Devey *et al.* methodology must be considered, and this will be discussed in greater detail in Section 5.

<sup>16</sup> Households that pay domestic maids are excluded.

## 5. THE REVISED DEVEY, SKINNER & VALODIA FORMAL- INFORMAL INDEX

### 5.1 The revised Devey et al. index

At the outset, it is noted that two of the main errors of the Devey *et al.* method have already been addressed. In this section, the problem with regard to the manner in which the question on the number of employers (Question 4.4 – See Table 3) was scored is addressed. The most obvious solution would be to reduce the criteria from 13 to 12, and exclude the number of employers' question. However, this reduces the number of worker characteristics used, as well as impacting on the comparability of the revised index of 12 characteristics with the old index of 13 characteristics. Consequently, the question "4.4: Number of employers" was replaced with the question "4.26: Flexibility in work hours". Other questions were also considered, such as "4.7: Ownership of equipment" and "4.27: Willingness to work longer hours" but it is not clear that these variables will provide significantly different results for the informal sector as compared to the formal sector. Question "4.26: Flexibility in work hours" seems to be the best alternative indicator, as a much higher proportion of formal sector employees under the Stats SA method (more than 90% for the LFSs) state that work hours are fixed by employers, but this proportion is only slightly above 70% for informal sector employees.

Table 8 shows the 13 indicators used for the revised Devey *et al.* formal-informal index. It is evident that most of the problems in the original Devey *et al.* index mentioned in Section 3 have been solved. In addition, Table 9 compares the number of informal sector employees using the Stats SA definition and the revised Devey *et al.* methodology since LFS2001b, excluding LFS2004b. The results show that the revised Devey *et al.* methodology estimated a greater number of informal sector employees, compared with the Stats SA method and even the original Devey *et al.* method (Table 6).

Table 8 The indicators used to derive the revised Devey *et al.* formal-informal sector index

<b>Participants: Employees only</b>		
<b>Question number***</b>	<b>Index = 1</b>	<b>Index = 0</b>
4.6: Permanence of work	(1): Permanent	(2): Fixed period contract (3): Temporary (4): Casual (5): Seasonal (6): Don't know (9): Unspecified
4.8: Written contract with employer	(1): Yes	(2): No (3): Don't know (9): Unspecified
4.10: Who pays wage	(1): Employer (2): Labour broker (3): Contractor or agency	(4): Other (5): Don't know (9): Unspecified
4.11: Employer contributes to pension of retirement fund	(1): Yes	(2): No (3): Don't know (9): Unspecified
4.12: Paid leave	(1): Yes	(2): No (3): Don't know (9): Unspecified
4.13: Membership of trade union	(1): Yes	(2): No (3): Don't know (9): Unspecified

\*\*\* The question number refers to the LFS2007b questionnaire

Table 8 Continued

Participants: Employees only		
Question number	Index = 1	Index = 0
4.16 Number of regular workers in enterprise	(6): 50 or more	(1): 1 (2): 2 – 4 (3): 5 – 9 (4): 10 – 19 (5): 20 – 49 (7): Don't know (9): Unspecified
4.17: Working for a registered company or close corporation	(1): Yes	(2): No (3) Don't know (9) Unspecified
4.18: Employer makes UIF deductions	(1): Yes	(2): No (3): Don't know (9): Unspecified
4.19: Employer makes medical aid or health insurance payments	(1): Yes, for himself only (2): Yes, for himself and his dependents (3): Yes, but he is not using it	(4): No, because he is covered by someone else's medical aid (5): No medical aid benefits provided (6): Don't know (9): Unspecified
4.20: Enterprise is registered to pay VAT	(1): Yes	(2): No (3): Don't know (9): Unspecified
4.23: Location of work	(3): Inside a formal business premises (4): At a service outlet	(1): In the owner's home (2): In someone else's home (5): At a market (6): On a footpath or street (7): No fixed location (8): Other (9): Unspecified
4.26: Flexibility in work hours	(3): Work hours fixed by employer	(1): Can decide fully for himself (2): Can decide, but within a limited range (4): Don't know (9): Unspecified

Table 9 Number of informal sector employees, Stats SA method vs. revised Devey *et al.* method 2001 – 2007

	Stats SA method	Revised Devey <i>et al.</i> method			
		Informal sector if index value: $[\leq 3]$	Informal sector if index value: $[\leq 4]$	Informal sector if index value: $[\leq 5]$	Informal sector if index value: $[\leq 6]$
LFS2001b	633 205	586 123	832 767	1 230 811	1 687 180
LFS2002a	585 946	554 832	766 354	1 112 372	1 507 348
LFS2002b	553 441	509 497	715 447	1 047 973	1 415 142
LFS2003a	619 645	539 317	759 568	1 116 111	1 491 011
LFS2003b	625 345	511 770	711 907	1 068 805	1 437 479
LFS2004a	576 490	503 639	706 993	1 069 133	1 444 771
LFS2005a	757 388	602 827	827 509	1 205 564	1 650 328
LFS2005b	870 047	724 576	993 659	1 446 944	1 900 697
LFS2006a	712 459	599 867	855 525	1 271 399	1 749 121
LFS2006b	794 486	693 607	938 990	1 379 025	1 845 961
LFS2007a	753 548	619 391	884 481	1 344 338	1 831 536
LFS2007b	667 811	538 125	786 404	1 239 980	1 862 225



## 5.2 Revised Devey et al. index using principal-components analysis (PCA)

The problem of the comparability of scores and the weighting of the criteria can be resolved, to some extent, by conducting a principal-components analysis (PCA). Instead of simply adding up the 13 indicators (or dummies) from Table 8, a formal-informal index is created using PCA. The rationale is simply that a greater weight should be attached to a variable if fewer people possess that characteristic. This reduces the comparability of the scores problem to some extent (although not satisfactorily) and removes the randomness when selecting the most important characteristics as criteria for the definition of the informal sector.

Table 10 Scoring coefficients of each dummy variable

Dummy variable	Scoring coefficient
Permanence of work: Permanent	0.3742
Written contract with employer: Yes	0.3508
Who pays wage: Employer, labour broker, contractor or agency	0.0519
Employer contributes to pension of retirement fund: Yes	0.3956
Paid leave: Yes	0.4006
Membership of trade union: Yes	0.2661
Number of regular workers in enterprise: 50 or more	0.2227
Working for a registered company or close corporation: Yes	0.1495
Employer makes UIF deductions: Yes	0.2909
Employer makes medical aid or health insurance payments: Yes	0.2786
Enterprise is registered to pay VAT: Yes	0.1453
Location of work: Inside a formal business premises or at a service outlet	0.2901
Flexibility in work hours: Fixed by employer	0.0836

The scoring coefficients of the various characteristics as shown in Table 10, with paid leave, pension fund contribution and permanence of work scoring the highest. Table 11 shows the findings from the revised Devey *et al.* method, using PCA and LFS2004a data. After the index is divided into deciles, a few remarkable results are found. The first or lowest decile when arranged by formal sector score – i.e., the decile of employees showing the strongest informal sector characteristics – includes nearly 50% of people classified as formal sector employees under the Stats SA method. In other words, nearly half of employees in this decile who are in formal sector employment according to the Stats SA definition, show very strong informal sector characteristics.

Table 11 The revised Devey *et al.* method using PCA, LFS2004a

Decile	Status (Using the Stats SA enterprise registration methodology)					
	Formal sector employees	Informal sector employees	Total	Formal sector employees	Informal sector employees	Total
1	49.1%	50.9%	100.0%	6.1%	76.5%	10.0%
2	88.8%	11.2%	100.0%	8.8%	13.5%	10.0%
3	96.6%	3.4%	100.0%	9.7%	4.2%	10.0%
4	98.7%	1.3%	100.0%	10.7%	1.7%	10.0%
5	99.1%	0.9%	100.0%	10.7%	1.2%	10.0%
6	99.5%	0.5%	100.0%	10.9%	0.7%	10.0%
7	99.3%	0.7%	100.0%	12.3%	1.1%	10.0%
8	99.7%	0.4%	100.0%	9.8%	0.4%	10.0%
9	99.4%	0.6%	100.0%	12.5%	0.9%	10.0%
10	99.9%	0.1%	100.0%	8.3%	0.1%	10.0%
Total	92.3%	7.7%	100.0%	100.0%	100.0%	100.0%

### 5.3 Informal sector definition summary

At this juncture, several methods to define the informal sector have been presented. The methods used, in the main, can either have an enterprise base or employment characteristic base. However, this can be confusing and at times the methods employed utilise both enterprise and employment criteria, using one method for the self-employed, and another for the employees.

Table 12 summarizes these possibilities for the four methodologies evaluated here. Firstly, enterprises can either be grouped as formal or informal. Secondly, employment can either be grouped as formal or informal (this is necessary to allow for informal employment within formal sector firms). This matrix presents four possibilities, labelled from [1] to [4] in Table 12, where [1] represents formal employment in a formal enterprise, [2] represents informal employment in a formal enterprise, [3] represents formal employment in an informal enterprise, and [4] represents informal employment in an informal enterprise.

Consequently, it is possible to determine whether the methodology used to define the informal sector is enterprise based (noted as A) or employment based (noted as B). If the methodology to define the informal sector utilises [3] and [4], it can be defined as an enterprise based definition<sup>17</sup>. On the other hand, if the methodology to define the informal sector utilises [2] and [4], it can be defined as an employment based definition.

Furthermore, the methodologies reviewed earlier used different definitions to measure the informal sector for the self-employed and the employees. Heintz and Posel, for instance, used an enterprise definition to classify the informal sector for the self-employed, whilst using a three-indicator employment definition to classify informal sector employees.

Table 12 Summary of the four informal sector definition methods

<b>Enterprise-based vs. Employment-based definition of informal sector</b>		
<b>Production units</b>	<b>Type of employment</b>	
	<b>Formal</b>	<b>Informal</b>
<b>Formal enterprises</b>	[1]	[2]
<b>Informal enterprises</b>	[3]	[4]
(A) Enterprise-based definition of informal sector: [3] + [4]		
(B) Employment-based definition of informal sector: [2] + [4]		
<b>Definition used in each method</b>		
<b>Methodology</b>	<b>Definition used to define informal sector</b>	
	<b>Self-employed</b>	<b>Employees</b>
Statistics South Africa	(A)	(A)
Devey <i>et al.</i> index	(B) – but with problems <sup>#</sup> Rather use (A) instead?	(B) – using 13 indicators
Heintz & Posel index	(A)	(B) – using 3 indicators
Revised Devey <i>et al.</i> index	(A)	(B) – using 13 indicators, with one of them being different from one used in the Devey <i>et al.</i> index

<sup>#</sup> Self-employed are only asked to answer questions on 6 out of 13 indicators.

<sup>17</sup> Note that, according to this definition, a formal employee in an informal enterprise will be regarded as an informal sector worker.

## **6. COMPARATIVE ANALYSIS OF INFORMAL SECTOR EMPLOYMENT STATISTICS USING THE VARIOUS DEFINITIONS**

The matter which remains to be completed is a comparison of the results obtained using the different methodologies. The main results are summarised in this section, with Table 13 below showing the following summary statistics:

- The number of informal sector employees, derived using the various methodologies.
- The number of informal sector workers, including both employees and self-employed.

An important point should be highlighted here. Apart from Devey *et al.*, the other definitions discussed in this paper use the Stats SA methodology to derive the self-employed informal sector workers (see Table 12)<sup>18</sup>. Consequently, the total number of informal sector workers (employees + self-employed) is derived by adding the number of informal sector employees from each method to the number of self-employed informal sector workers (i.e., 4th column of Table 2). For example, the total number of employed in informal sector using the Heintz & Posel method in LFS2001b is equal to 1 905 627 (employees) + 1 330 568 (self-employed).

---

<sup>18</sup> The problems encountered with the Devey *et al* method for determining the self-employed was noted earlier.

Table 13 The number of informal sector workers using various definitions

	Stats SA	Heintz & Posel	Devey <i>et al.</i> index ( $\leq 3$ )	Devey <i>et al.</i> index ( $\leq 4$ )	Devey <i>et al.</i> index ( $\leq 5$ )	Revised Devey <i>et al.</i> index ( $\leq 3$ )	Revised Devey <i>et al.</i> index ( $\leq 4$ )	Revised Devey <i>et al.</i> index ( $\leq 5$ )
<b>Number of informal sector employees</b>								
LFS2001 b	633 205	1 905 627	547 509	787 598	1 192 045	586 123	832 767	1 230 811
LFS2002 a	585 946	1 623 791	526 278	730 781	1 083 369	554 832	766 354	1 112 372
LFS2002 b	553 441	1 543 593	486 071	692 058	1 020 960	509 497	715 447	1 047 973
LFS2003 a	619 645	1 549 921	516 864	744 879	1 091 022	539 317	759 568	1 116 111
LFS2003 b	625 345	1 413 217	495 921	666 530	1 032 132	511 770	711 907	1 068 805
LFS2004 a	576 490	1 332 506	487 673	679 871	1 038 721	503 639	706 993	1 069 133
LFS2005 a	757 388	1 505 770	584 770	813 437	1 174 306	602 827	827 509	1 205 564
LFS2005 b	870 047	1 729 334	696 910	965 023	1 424 220	724 576	993 659	1 446 944
LFS2006 a	712 459	1 598 062	575 414	824 317	1 241 439	599 867	855 525	1 271 399
LFS2006 b	794 486	1 677 776	675 105	918 664	1 364 107	693 607	938 990	1 379 025
LFS2007 a	753 548	1 715 578	582 463	853 469	1 328 726	619 391	884 481	1 344 338
LFS2007 b	667 811	1 582 327	518 838	775 185	1 223 554	538 125	786 404	1 239 980
<b>Number of informal sector workers (employees + self-employed)</b>								
LFS2001 b	1 963 773	3 236 195	1 878 077	2 118 166	2 522 613	1 916 691	2 163 335	2 561 379
LFS2002 a	1 821 426	2 859 271	1 761 758	1 966 261	2 318 849	1 790 312	2 001 834	2 347 852
LFS2002 b	1 778 542	2 768 694	1 711 172	1 917 159	2 246 061	1 734 598	1 940 548	2 273 074
LFS2003 a	1 827 393	2 757 669	1 724 612	1 952 627	2 298 770	1 747 065	1 967 316	2 323 859
LFS2003 b	1 901 131	2 689 003	1 771 707	1 942 316	2 307 918	1 787 556	1 987 693	2 344 591
LFS2004 a	1 764 630	2 520 646	1 675 813	1 868 011	2 226 861	1 691 779	1 895 133	2 257 273
LFS2005 a	2 068 479	2 816 861	1 895 861	2 124 528	2 485 397	1 913 918	2 138 600	2 516 655
LFS2005 b	2 459 690	3 318 977	2 286 553	2 554 666	3 013 863	2 314 219	2 583 302	3 036 587
LFS2006 a	2 187 940	3 073 543	2 050 895	2 299 798	2 716 920	2 075 348	2 331 006	2 746 880
LFS2006 b	2 376 338	3 259 628	2 256 957	2 500 516	2 945 959	2 275 459	2 520 842	2 960 877
LFS2007 a	2 129 164	3 091 194	1 958 079	2 229 085	2 704 342	1 995 007	2 260 097	2 719 954
LFS2007 b	2 083 855	2 998 371	1 934 882	2 191 229	2 639 598	1 954 169	2 202 448	2 656 024

The results between the Devey *et al.* and revised Devey *et al.* method vary somewhat, with the informal sector being larger than the results obtained by the Stats SA method if the index score is smaller and equal to 4. However, the Heintz & Posel method delivers the largest informal sector, even when compared to a Devey *et al.* index score of smaller and equal to 5.

Figure 6 highlights the trends for the number of informal sector employees for the various methods from LFS2001b to LFS2007b (with LFS2004b excluded as noted earlier). It is apparent that all the methods discern the same trend, but that the overall number of informal sector employees changes significantly, depending on the method (and index score) used.

Figure 6 Number of informal sector employees for the various methods

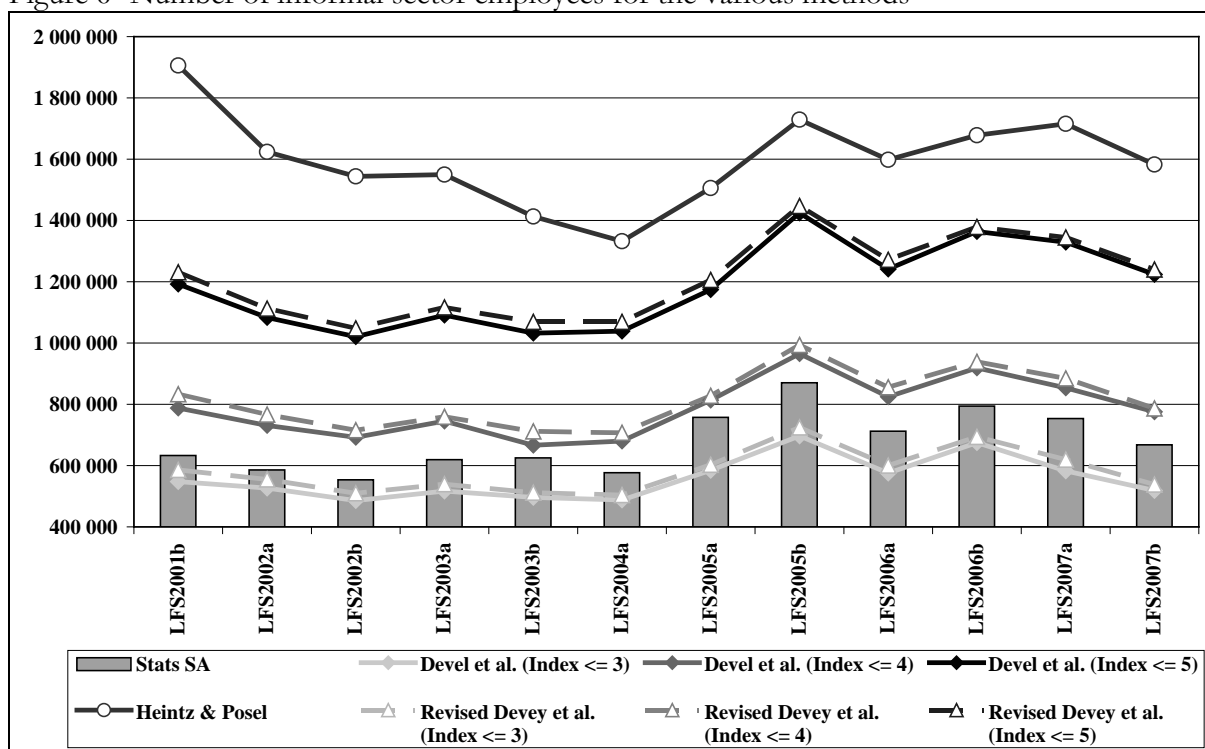


Figure 7, in turn, provides a similar trend for the total number of workers (both employees as well as the self-employed) for the period noted above, when using the various methods. The range, again, is large, with the total number of workers ranging between 1.93 million (Devey *et al.* method, index score of smaller or equal to 3) to nearly 3 million (Heintz & Posel) in LFS2007b.

In addition, Figure 8 shows the informal sector employment (employees + self-employed) as % of non-agricultural employment using various methods. As mentioned in Section 2, non-agricultural employment equals the sum of the third and fourth columns of Table 1. It can be seen from Figure 8 that this proportion hovers around 20% under the Stats SA method, but increases to about 28% using both the Devey *et al.* and revised Devey *et al.* indices (on condition the index value is equal to or smaller than 5), and about 30% under the Heintz & Posel method.

Figure 7 Number of informal sector workers (employees + self-employed) under various methods

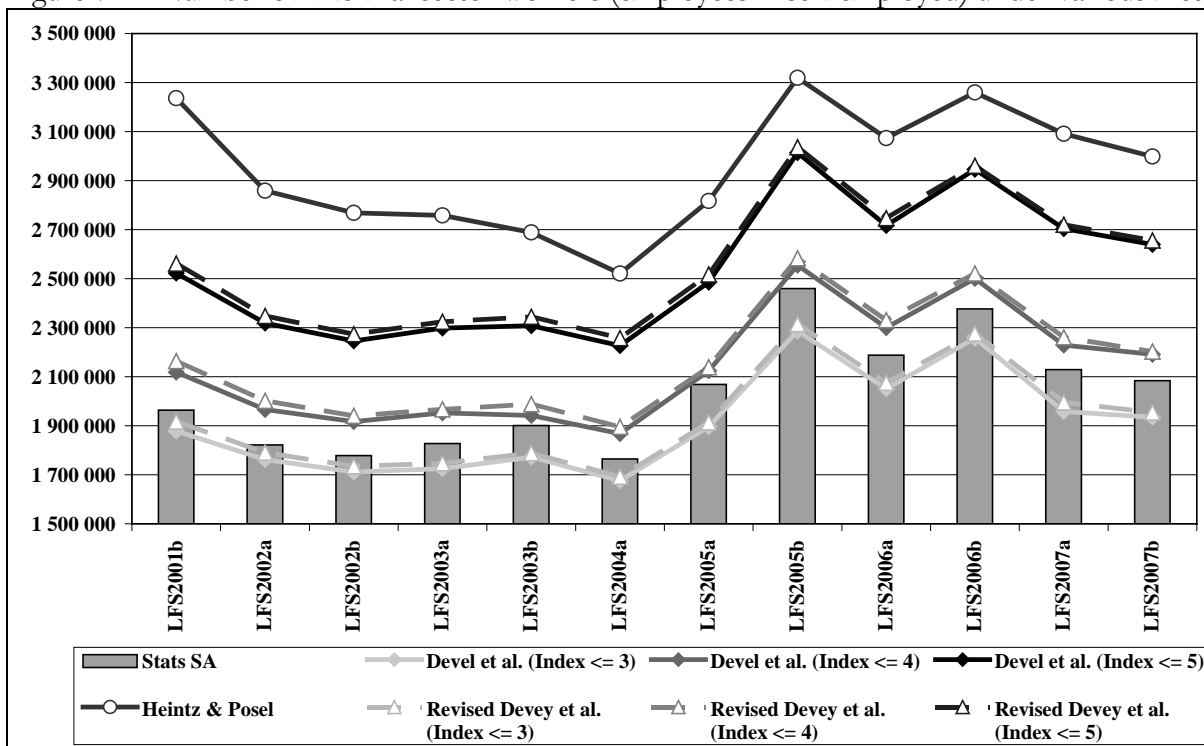
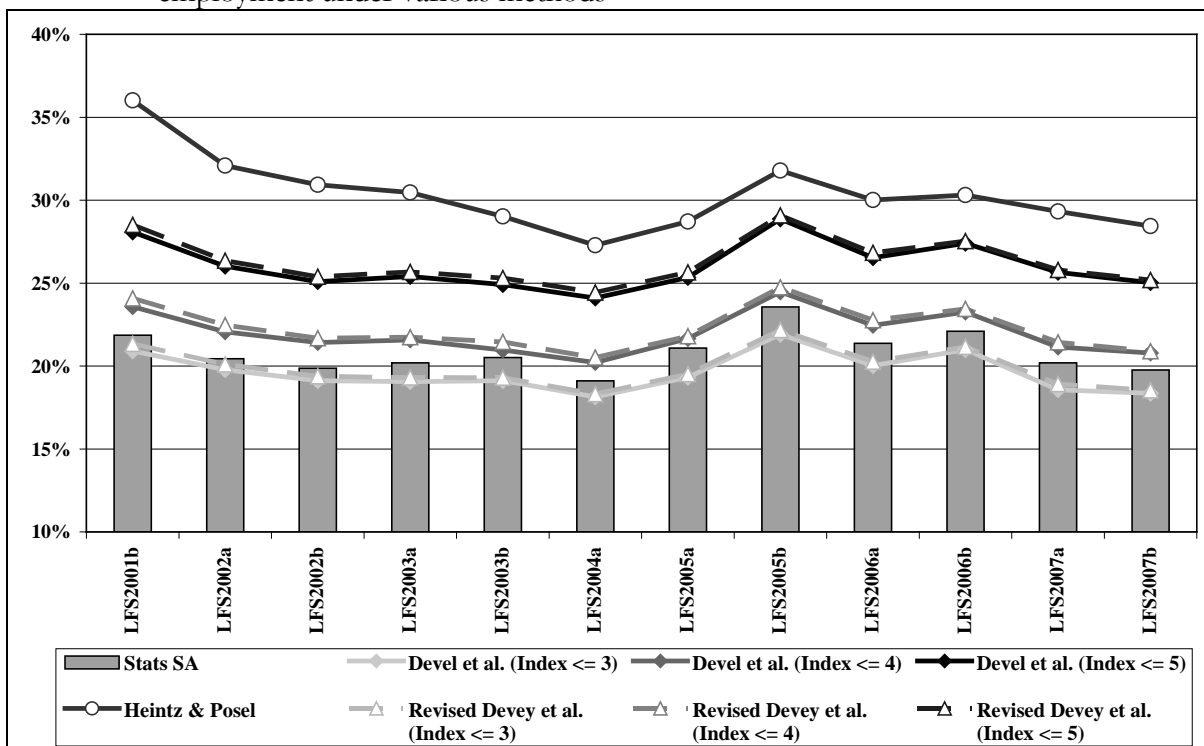


Figure 8 Informal sector workers (employees + self-employed) as % of all non-agricultural employment under various methods



Figures 9 and 10 illustrate the ratio of informal sector employment as percentage of the overall unemployment rate in South Africa, using both the narrow and broad definitions of unemployment. It can be seen that the ratio shows a slight upward shift since LFS2004a, regardless of which methodology is used to derived informal sector employees. As is expected, this ratio is highest if the Heintz & Posel method is applied. It is conceivable that this upward shift is likely caused by the decrease in unemployment, rather than an increase in informal sector employment.

Figure 9 Ratio of informal sector employment as % of non-agricultural employment to the narrow unemployment rate

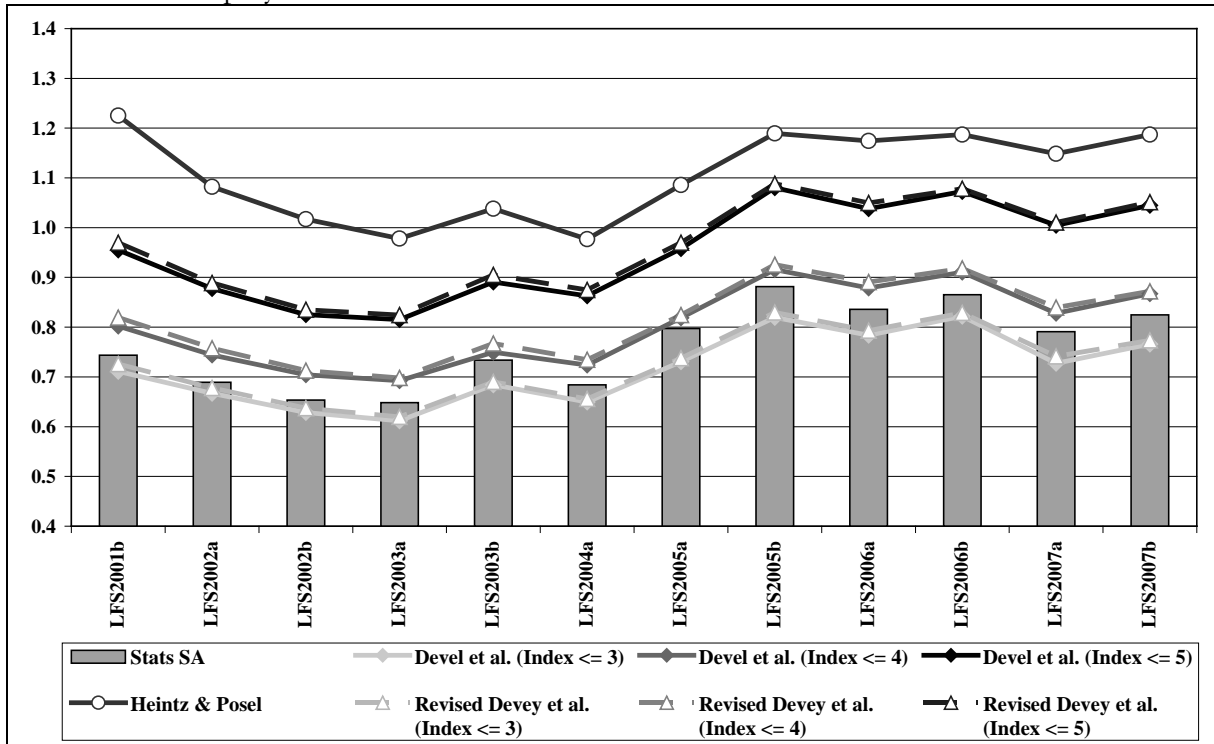
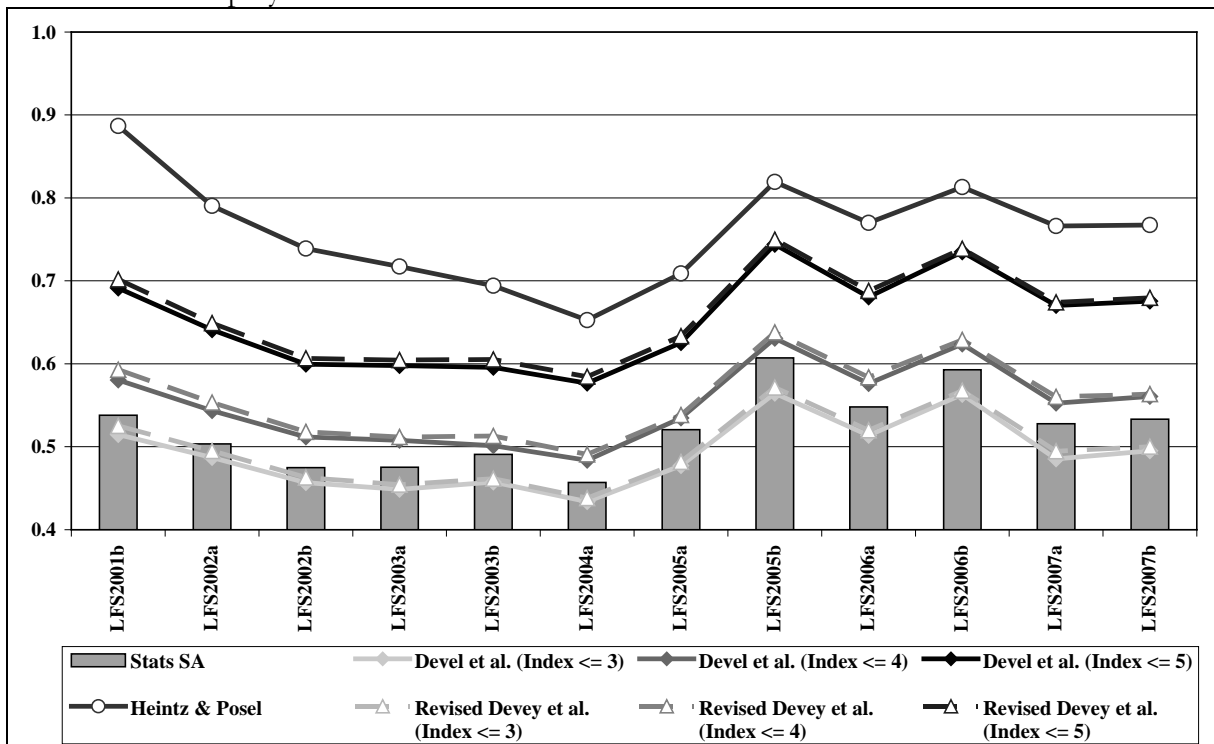


Figure 10 Ratio of informal sector employment as % of non-agricultural employment to the broad unemployment rate ratio



## **7. THE CURRENCY DEMAND APPROACH**

### **7.1 Introduction**

The Stats SA, Devey *et al.*, revised Devey *et al.* and Heintz & Posel approaches to measure the size of the informal sector as explained in Sections 3-6 use the survey data (i.e., OHSs and LFSs). This is defined by Brambila Macias (2008: 4) as direct methods. However, Brambila Macias argues that it is also possible to measure the size of the informal sector using alternative data sources and other methods, and they are as follows:

- Indirect methods: Indirect methods were introduced by authors such as Tanzi (1983), and make use of differences in official data to determine the size of the informal sector. Such methods are used in some instances where survey data quality is undermined, or where the data do not exist. Some examples of the differences in official data are variations between national income and consumption, discrepancies between official and actual labour force data, etc. The currency demand approach, to be discussed in Section 7.2, falls under this category.
- MIMIC / Model approach: This method uses structural equations to model causal relationships amongst the unobserved variables. The Model approach should be classified with the indirect approach, but Brambila Macias argues that the former differs from the latter as it is able to connect observed and unobserved variables.

### **7.2 The currency demand approach**

The currency demand approach has been widely used in both developed countries (e.g., Tanzi (1983)) and developing countries (e.g., Brambila Macias (2008) applied this technique to Mexico). The main assumption behind the currency demand approach is based on the idea that transactions in the informal sector are mainly based on cash. This allows participants to make observation by authorities more difficult, and prosecution can thus be avoided. The currency demand data is used to determine what is required by the formal sector of the economy, whilst the remaining consumption is assumed to be used by the informal sector<sup>19</sup>.

However, Hanousek and Palda (2004) note several shortcomings with this key assumption. They argue that the deductions made based on these assumptions only hold if the parameters of the money demand equation are accurate and stable, or if the changes of these parameters are known and adjusted for. If parameters, such as velocity, are “too variable”, then indirect methods as a whole would not be useful in determining the size of the informal sector.

Specifically, Hanousek and Palda (2004: 3-4) note that – for transition economies – intensive financial innovation and the increased number of financial products that this process enables, affect the rate at which currency demand grows at a “greater and more variable pace than they do in more mature western economies”. In this regard, South Africa’s banking sector can be regarded to be more “mature” when compared to other developing countries.

The model used is as follows (Brambila Macias, 2008: 5-10):

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 TAX_t + \beta_3 R_t + \beta_4 REM_t + v_t, \text{ where}$$

C = Natural logarithm of currency in circulation outside the banks over GDP deflator

Y = Natural logarithm of real GDP

TAX = Total tax revenues over GDP

R = Natural logarithm of the average of time deposit interest rates

REM = Natural logarithm of the amount of remittances received normalized by GDP

The specification above estimates the long run relationships between the explanatory variables and the currency demand. Additionally, annual data from the South African Reserve Bank (SARB) from 1990 to



2007 was used (Table 14), with the sole change being the exclusion of the remittances variable as such data is difficult to obtain. Besides, Brambila Macias does not specify clearly whether the remittances stand for the amounts received domestically or from abroad. In other words, the model becomes:

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 TAX_t + \beta_3 R_t + v_t.$$

Table 14 Variables used in the South African currency demand model

Variable	Description	Code in the SARB Quarterly Bulletin
C	Natural logarithm of (M3/GDP deflator#)	M3: 1374
Y	Natural logarithm of real GDP (2000 prices)	GDP: 6006
TAX	Total tax revenues over GDP	Tax revenue: 4582
R	Natural logarithm of prime rate (Average of the 12 monthly figures)	Prime rate: 1403

Note: GDP deflator = (Nominal GDP / Real GDP) × 100

From theory, it is expected that GDP and taxes should have a positive impact on currency demand. Tanzi (1983) deduces the latter result as follows: if taxation increases, tax evasion is encouraged, which leads to greater use of cash to circumvent detection and records of transactions. Consequently, the use of currency increases. In addition, interest rates are expected to have a negative effect, as economic agents would want to reduce their currency holdings in favour of other financial instruments (Brambila Macias, 2008: 7).

As is the norm for currency demand models, the series were tested for the presence of unit roots and cointegration. In this instance, the Augmented Dickey-Fuller unit root test was conducted, and it was found that C, Y, TAX and R contain one unit root. Given the non-stationarity of the series and the presence of a commonly stochastic trend, the equation needs to be estimated using a vector-correction model (VECM).

Using the equation above to estimate the VECM, the regression is then used to derive  $\hat{C}$ . Next, the tax variable is set to 0, after which the regression is re-run to derive  $\tilde{C}$ , which stands for the currency demand at a zero tax level. The difference between  $\hat{C}$  and  $\tilde{C}$  gives the amount of extra currency in the economy (EC). In turn, the velocity of circulation of money ( $v$ ) is obtained by:  $v = \frac{Y(= GDP)}{M3 - EC}$ . Finally, the size of the informal sector can be obtained by:  $EC * v = Y_{\text{informal}}$ , and then the size of the informal sector in formal GDP terms can be inferred ( $Y_{\text{informal}} + Y_{\text{formal}} = Y$ ).

The results are presented in Table 15 and Figure 11, with a comparison between the informal sector size as defined by the official StatsSA methodology added. Firstly, there seems to be some variation in the pre-2002 results. However, this could be due to informal sector earnings being over-estimated in OHSs (Essop & Yu, 2008: 24), and in LFS2001a the number of informal sector workers were also over-estimated (Essop & Yu, 2008: 10). This might help explain why the two graphs do not have the same patterns from 1997 to 2001.

Table 15 Long-run cointegrating equation, currency demand approach

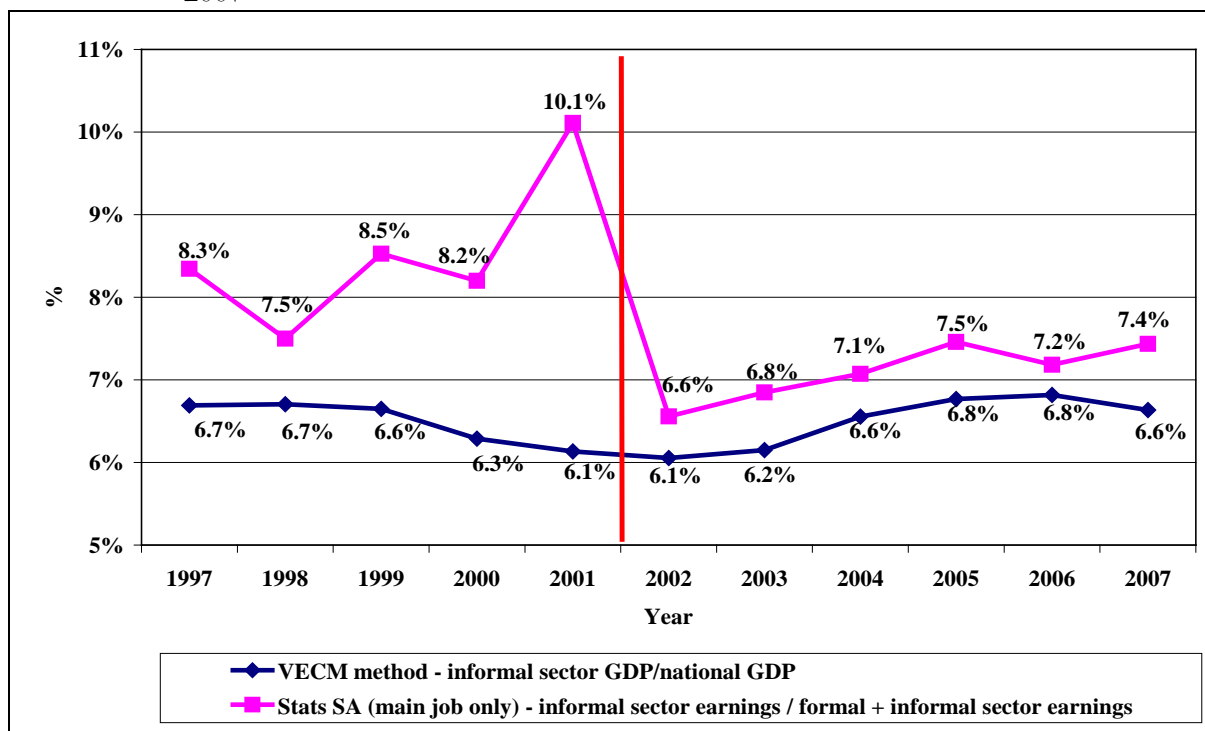
	Cointegrating coefficients	
	Coefficient	Standard error
$C_{t-1}$	1.0000	
$Y_{t-1}$	-2.2718***	0.0574
$TAX_{t-1}$	-9.3832***	0.8028
$R_{t-1}$	-0.0083***	0.0018
Constant	23.5101	
Log likelihood	156.2581	

Note: All series used in the model contain one unit root [i.e., I(1)]. Besides, the model assumes one cointegrating equation and was estimated using one lag.

\*\*\* indicates significance at the 1% level.

Secondly, even when considering the slight differences in estimations for the full series when compared to the OHS and LFS data, the South African informal sector does not exceed 7% of GDP with the VECM, reflecting a much smaller informal sector when compared to Mexico, where Brambila Macias (2008) estimates that the informal sector is approximately 20-30% of GDP from the early 1990s until the mid 2000s.

Figure 11 Size of the informal sector: Stats SA method vs. Currency demand approach (VECM), 1997 – 2007



Note: Taking the average of the March and September values in each LFS since 2000.

Note: In the Stats SA graph, only earnings from main job are included. Also, people earning more than 1 million in 2000 prices per annum were regarded as outliers and excluded.

Furthermore, it appears that the difference between the VECM approach and the Stats SA approach stabilizes at +0.5 percentage points from 2002 onwards. This may indicate the VECM model is appropriate to use within the South African context, with the concerns of Hanousek and Palda (2004) not being as relevant here as in transition economies. However, the VECM approach undoubtedly presents an underestimation of the informal sector, as the StatsSA measure has been shown in this paper<sup>20</sup> to provide smaller estimates for the informal sector compared to other methodologies (see Figures 9 and 10 above); yet, in Figure 11, the StatsSA method provides a marginally larger estimate of

<sup>20</sup> Also see Essop & Yu (2008).

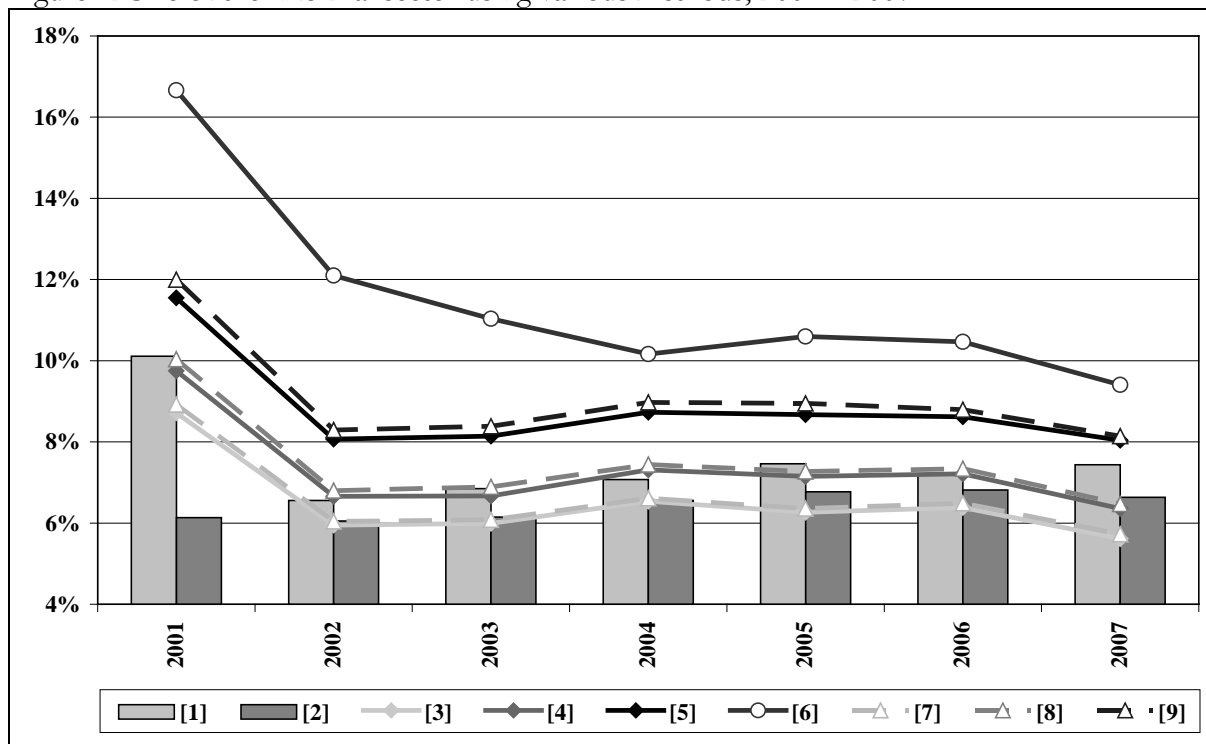
the informal sector as compared to the VECM approach, albeit with somewhat different variables used.

When comparing the size of the informal sector as measured by the VECM approach and all the other methods discussed in this paper, a few points can be discerned as presented in Figure 12. (See Table 15 for a description of the method used for both employed and self-employed). Firstly, it seems that the informal sector size derived by the VECM method is similar to the results using the original Devey *et al.* and the revised Devey *et al.* methods with an index score of less than four (methods [3] and [7] in Table 15). Secondly, if the Stats SA method is applied on self-employed and the Heinz & Posel method is applied on employees (i.e., method [6] in Table 15), the results again show that the Heinz & Posel method provide a much larger estimate of the informal sector compared to the VECM approach. Overall, however, the VECM approach appears to provide the same trend (since 2001) in the informal sector as found with other methods, albeit with an apparent under-estimation of the overall size of the informal sector.

Table 15 Methods employed to measure and compared the size of the informal sector

Method	Method used to distinguish informal sector workers from formal sector workers – self-employed	Method used to distinguish informal sector workers from formal sector workers – employees	Meaning of the ratio in Figure 12
[1]	Stats SA method	Stats SA method	Informal sector earnings / (Formal + informal sector earnings)
[2]	VECM method		Informal sector GDP / national GDP
[3]	Stats SA method	Devey <i>et al.</i> (Index $\leq 3$ )	Informal sector earnings / (Formal + informal sector earnings)
[4]	Stats SA method	Devey <i>et al.</i> (Index $\leq 4$ )	Informal sector earnings / (Formal + informal sector earnings)
[5]	Stats SA method	Devey <i>et al.</i> (Index $\leq 5$ )	Informal sector earnings / (Formal + informal sector earnings)
[6]	Stats SA method	Heintz & Posel	Informal sector earnings / (Formal + informal sector earnings)
[7]	Stats SA method	Revised Devey <i>et al.</i> (Index $\leq 3$ )	Informal sector earnings / (Formal + informal sector earnings)
[8]	Stats SA method	Revised Devey <i>et al.</i> (Index $\leq 4$ )	Informal sector earnings / (Formal + informal sector earnings)
[9]	Stats SA method	Revised Devey <i>et al.</i> (Index $\leq 5$ )	Informal sector earnings / (Formal + informal sector earnings)

Figure 12 Size of the informal sector using various methods, 2001 – 2007



Note: Taking the average of the March and September values in each LFS since 2000.

Note: Only earnings from main job are included. Also, people earning more than 1 million in 2000 prices per annum were regarded as outliers and excluded.

## **8. CONCLUSION**

As mentioned earlier, there are several reasons why it is imperative for policy makers, amongst others, to be concerned about the size of the informal sector. Consequently, using an appropriate definition and measurement of the informal sector becomes important. However, defining the informal sector, and its ensuing measurement, has been problematic, both internationally and domestically. Failure to define and measure the informal sector in an appropriate manner, of course, hampers the ability of policy makers to implement appropriate measures to address problems in the informal labour market. This paper considered several definitions and methods currently used in South Africa to measure the size of the informal sector, these methods being the Stats SA pre-2008 enterprise methodology, and newly adopted 2008 methodology, Devey *et al.* formal-informal index, Heintz & Posel methodology, as well as the currency demand approach. In addition, due to some concerns with the Devey *et al.* formal-informal index, a new, revised Devey *et al.* method was formulated and included in the analysis.

Overall, it appears that the Stats SA enterprise methodology could have resulted in an under-estimation of the informal sector employees in South Africa. However, if an employee characteristic method as promoted by the ICLS is used, it appears that South Africa has a larger informal sector as compared to the official statistics. This result appears consistent for all methods use, bar the currency demand approach, which provides a smaller estimate of the informal sector in South Africa when compared to the official Stats SA method. Furthermore, of all the methods evaluated, the Heintz & Posel measurement provides the largest informal sector, and may potentially be an overestimation of the informal sector as noted earlier.

As mentioned in Section 1, a question that still needs to be addressed is whether South Africa would still be an international outlier in the size of its informal sector if the alternative definitions of informal sector analysed in this paper are adopted. Unfortunately, it is difficult to raise enough evidence to answer the question raised, due to the following reasons: these alternative methods need to be applied on the other countries' labour force survey data before cross-comparisons could be made; no judgement was made on the appropriate index score for Devey *et al.* or the revised Devey *et al.* formal-informal indices; and finally, the Heintz & Posel method could potentially be an overestimation.

## **REFERENCES**

Brambila Marcias, B., 2008. *The dynamics of parallel economies. Measuring the informal sector in Mexico*. Munich Personal RePEc Archive Paper No. 8400. Munich: University Library of Munich.

Devey, R., Skinner, C. & Valodia, I., 2003. *Informal economy employment data in South Africa: A Critical Analysis*. Paper presented at the TIPS & DPRU Forum held at Johannesburg, 8-10 September.

Devey, R, Skinner, C and Valodia, I., 2006. Definitions, Data and the Informal Economy in South Africa: A Critical Analysis. In: Padayachee, V (ed.), *The development decade? Economic and social change in South Africa, 1994 – 2004*. Cape Town: Human Sciences Research Council Press.

Essop, H. & Yu, D., 2008. *The South African informal sector (1997 – 2006)*. Stellenbosch Economic Working Papers: 03/08. Stellenbosch: Stellenbosch University.

Fields, G.S., 1975. *Rural-urban migration, urban unemployment and underemployment, and job search activity in LDC's*. Journal of Development Economics, 2: 165-187.

Hanousek, J. & Palda, F., 2004. *Mission implausible III: Measuring the informal sector in a transition economy using macro methods*. William Davidson Institute Working Paper Number 683. Michigan: The William Davidson Institute, University of Michigan Business School.

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

Henley, A., Reza Arabsheibani, G. & Carneiro, F.G., 2006. *On defining and measuring the informal sector*. World Bank Policy Research Paper 3866. Washington: World Bank.

International Labour Organization, 2002a. *International informal sector statistics*. Available: <http://www.ilo.org/public/english/bureau/stat/papers/comp.htm>

International Labour Organization, 2002b. *Women and men in the informal economy, a statistical picture*. International Labour Organisation: Geneva.

Kingdon, G.G. & Knight, J., 2004. Unemployment in South Africa: The nature of the beast. *World Development*: 32(3): 391 – 408.

Kingdon, G.G. & Knight, J., 2007. Unemployment in South Africa, 1995 – 2003: Causes, problems and policies. *Journal of African Economies*: 16(5): 813 – 848.

Muller, C., 2003. *Measuring South Africa's informal sector: an analysis of national household surveys*. DPRU Working Paper 03/71. Cape Town: Development Policy Research Unit.

Muller, C. & Posel, D., 2004. Concerns with measuring informal sector employment: an analysis of national household surveys in South Africa, 1993 – 2001. *Studies in Economics and Econometrics*, 28(1): 1 – 19.

Muzamdar, D., 1976. *The urban informal sector*. World Development, 4(8): 655-679.

Statistics South Africa, 2008. *Guide to the Quarterly Labour Force Survey*. Available: <http://www.statssa.gov.za/Publications/Report-02-11-01/Report-02-11-01August2008.pdf>

Tanzi, V., 1983. *The underground economy in the United States: Annual estimates, 1930-1980*. IMF Staff Papers, 33: 283-305.

Yu, D., 2007. *The Comparability of the Statistics South Africa October Household Surveys and Labour Force Surveys*. Stellenbosch Economic Working Papers: 17/07. Stellenbosch: Stellenbosch University.