



## Measuring Informal Employment in South Africa: The New Quarterly Labour Force Survey

Debbie Budlender<sup>1</sup>

### Introduction

In 2008 Statistics South Africa replaced the bi-annual Labour Force Survey (LFS) with the Quarterly Labour Force Survey (QLFS). The questionnaire was substantially “re-engineered” in this process. This resulted in the omission of many questions, the addition of a few questions, and changes in how some questions were formulated.

WIEGO has previously commissioned analysis of informal employment in South Africa by researchers who used LFS data for this analysis. Heintz and Posel (2008) analyzed data from the surveys of September 2001 and September 2004 to explore reasons for the high unemployment rates in the country. They did so by developing a definition for informal employment based on the definition of informal employment approved by the 17<sup>th</sup> International Conference of Labour Statisticians. Subsequently, Wills used the same definitions to conduct analysis of informal employment using LFS data from 2005, 2006 and 2007 (2010).

This note discusses how the changes introduced in the QLFS (a) will affect how the informal sector and informal employment are or can be defined for analysis purposes, and (b) will affect comparability of estimates based on QLFS data with earlier estimates based on LFS data. Heintz and Posel produced estimates of both the informal sector and informal employment. Wills’ work focused on informal employment rather than on the informal sector. This note discusses both concepts – informal sector and informal employment – in terms of question (a) above, but only informal employment in terms of question (b).

### Informal Sector

The LFS had a direct question that asked whether the employed person – regardless of status in employment (e.g. employee, employer, own-account worker) – saw the business in which they worked as formal or informal. This question provided a simple and direct route for estimating informal sector employment. The question was phrased as follows:

*Is the organization/ business/ enterprise/ branch where ..... works*

*1 = In the formal sector*

*2 = In the informal sector (including domestic work)*

*3 = Don't know*

The question was followed by an italicized explanation as follows:

*Formal sector employment is where the employer (institution, business or private individual) is registered to perform the activity. Informal sector employment is where the employer is not registered.*

Given that the questionnaire was administered face-to-face rather than self-completed, it is not clear whether – and how often – fieldworkers would have read out this explanation to respondents.

The LFS also included questions on whether the business was registered for income tax and/or registered for value-added

---

<sup>1</sup> This note was authored by Debbie Budlender of the Community Agency for Social Enquiry, Cape Town, South Africa.

tax (VAT), and the number of regular workers employed in the business. These questions provided a range of alternative “proxy” ways of defining the informal sector and calculating estimates.

Heintz and Posel (2008) compare estimates of the informal sector derived using the direct question and estimates based on the questions about registration. They find that in the LFS of September 2004, the direct question provides a higher estimate of the informal sector than the questions on registration. Thus the direct question suggests that 69 per cent of employed people are in the formal sector, while the derived approach gives a lower estimate of 61 per cent. Expressed differently, 14 per cent of employed people who are reported to be working in the formal sector are reported to be working in a business that is not registered.

The current QLFS no longer has the direct question about whether the person works in the formal or informal sector. However, this question was included in the questionnaire for the first QLFS rounds. This allows for comparison of QLFS-based estimates using different approaches (see below). The ability to make these comparisons is important because of concerns that the QLFS, with its lesser emphasis on marginal work (see below), may not be picking up as much informal work as the LFS. It is thus useful to be able to compare direct and indirect approaches both for LFS data and for QLFS data.

The current QLFS does have questions about registration for income tax and VAT, as well as a question about the number of employees at the place of work. (The word “regular” has been dropped.) A difference from the LFS is that separate sets of questions are asked of employees on the one hand, and employers, own-account workers and unpaid family workers on the other hand.

Employees are asked whether their employer deducts income tax from their salary or wage. Employers, own-account workers and unpaid family workers are asked if their business (including a household business where relevant) is registered for VAT and whether the business is registered for income tax.

Employees are classified as working in the informal sector if their employer does not deduct income tax and if they work in a business that employs fewer than five employees. The first element of this is inaccurate to the extent that low-paid workers employed in a small establishment, such as a small accounting or medical practice, would be defined inappropriately as working in the informal sector. Employers, own-account workers and unpaid family workers are classified as working in the informal sector if the business is registered for VAT or income tax.

Fortunately, as noted above, the first QLFS questionnaires retained the direct question on the informal sector. We can thus compare the estimates derived from the different approaches to measuring the informal sector. We use the QLFS of the third quarter of 2008 for this purpose.

Table 1 reveals that the estimates derived using the two methods are fairly similar. The direct question finds that the informal sector accounts for 14 per cent of employment, while the derived estimate is slightly higher, at 17 per cent of employment. The direction of the difference between the direct and derived approaches is thus the same as that found by Heintz and Posel (2008), but the size of the difference is smaller. Part of the difference in the estimates for 2008 is accounted for by the 1 per cent of employed persons for whom the answer to the direct question is “Don’t know.” Slightly more divergence is found when we disaggregate by gender. For example, there is now a four percentage point difference in the estimate of the size of the informal sector for males.

**Table 1: Gender Comparison of Estimates Using Different Methods**

Sector	Male	Female	Total
<b>Direct question</b>			
Formal sector	81%	69%	76%
Informal sector	13%	14%	14%
Private households	4%	15%	9%
Do not know	1%	1%	1%
Total	100%	100%	100%
<b>Derived</b>			
Formal sector	80%	66%	74%
Informal sector	17%	17%	17%
Private households	4%	17%	9%
<b>Total</b>	100%	100%	100%

In addition to the gender breakdown, we can also compare the estimates from the viewpoint of status in employment. Table 2 again finds that the two approaches provide similar pictures. The largest relative disparities are among own-account workers, where the direct question finds that 86 per cent are informal while the derived method puts the estimate at 79 per cent. The derived method gives a lower percentage of informal across all types of status in employment.

**Table 2: Status in Employment Comparison of Estimates Using Different Methods**

Sector	Employee	Employer	Own-account	Unpaid family	Total
<b>DIRECT QUESTION</b>					
Formal	81%	69%	12%	40%	74%
Informal	8%	30%	86%	59%	17%
Private household	11%	0%	2%	1%	9%
<b>Total</b>	100%	100%	100%	100%	100%
<b>DERIVED</b>					
Formal	84%	68%	15%	40%	76%
Informal	5%	28%	79%	58%	14%
Private household	10%	3%	5%	2%	9%
Do not know	1%	1%	1%	1%	1%
<b>Total</b>	100%	100%	100%	100%	100%

Similar overall estimates do not necessarily mean that particular individuals are classified similarly by the two approaches. Thus a given number of people who are classified as informal by one method and formal by another may be cancelled out in the overall estimates by a similar number of people who are classified as formal by the first method and informal by the other. Table 2 cross-tabulates the two classifications to investigate the extent to which this occurs. The shaded cells – which sum to 82 per cent of all employed – represent those who are similarly classified by the two methods. The largest group of non-matches – accounting for 5 per cent of employed people – is made up of those who are classified as formal through the direct question and informal through the derived method.

**Table 3: Cross-tabulation of Estimates from Different Methods**

	Direct				
Derived	Formal	Informal	Private	Don't know	Total
Formal	71%	2%	0%	1%	74%
Informal	5%	11%	1%	0%	17%
Private	0%	1%	8%	0%	9%
<b>Total</b>	76%	14%	9%	1%	100%

## Informal Employment

Statistics South Africa did not itself define informal employment in terms of LFS data, although several researchers (including Heintz and Posel; and Wills) derived estimates of informal employment using these data-sets.

With the introduction of the QLFS, Statistics South Africa introduced a definition of informal employment. Under this definition, all employers and own-account workers who are in the informal sector (see above) are considered to be in informal employment. All other employers and own-account workers are given code 8, for “other,” rather than the code for formal employment.

In addition, employees are considered to be in informal employment if their employer does not contribute to medical insurance and does not contribute to a pension fund, and they also do not have a written contract of employment. (“Don’t know” responses to these questions are assumed to indicate “no.” This could introduce a small bias in estimates of informal employment. The bias is not important as the number of “don’t know” responses is very small.)

The official *Guide to the Quarterly Labour Force Survey* (Statistics South Africa, 2008) states that this definition is applied to all employees. In reality, it is not applied to employees in the household sector, who are assumed to be in informal employment. This approach automatically defines all domestic workers as informal. Also not stated explicitly in the *Guide* is that all unpaid family workers are regarded as informal. Analysis of the data from the QLFS of February 2010 reveals that 19 per cent of all those employed in the private household sector and 22 per cent of domestic workers would not be classified as informal if the approach described in the *Guide* was followed.

Statistics South Africa is uncomfortable with its own definition and plans to reconsider it. It will thus not be reporting on this variable until a new definition has been decided. One source of discomfort within Statistics South Africa is the assumption that all domestic workers are informal. Another source of discomfort is the assumption that a written contract makes the employment formal.

Heintz & Posel and Wills, using LFS data, define employees as formally employed if (a) they have a written contract or (b) receive paid leave and also have pension contributions made on their behalf by their employers. Employers and own-account workers are classified as formal if their business is registered as a company or close corporation, or if they are registered for VAT. The approach is thus similar to that used by Statistics South Africa, but not identical. We refer to the approach used by Heintz & Posel and Wills as the “WIEGO approach”.

Unfortunately, it is not possible to replicate the WIEGO definition using QLFS data as there is no longer a question relating to registration of a company or close corporation. The employee part of the definition would, however, be able to be applied as there are questions on contract, paid leave and pension contributions in the QLFS.

Table 4 reveals that 94 per cent of employees would be similarly classified using the WIEGO approach and Statistics South Africa’s definition. The difference between the two methods is accounted for by employees who are classified as formal by WIEGO but informal by Statistics South Africa.

**Table 4: Cross-tabulation of Employee Estimates Using Different Methods**

	Statistics South Africa		
WIEGO approach	Formal	Informal	Total
Formal	73%	5%	78%
Informal	0%	21%	22%
Total	74%	26%	100%

One unusual aspect of Wills’ approach is her definition of “non-agricultural.” Usually this would be defined on the basis of industry. Wills adds occupation, considering both those defined as skilled agricultural workers and those working in agriculture as “agricultural.” Using the QLFS of the second quarter of 2010, using occupation adds about 5 per cent to the agricultural total.

Heintz and Posel do not disaggregate employment into agricultural and non-agricultural. However, Wills’ approach of considering other variables beyond industry in distinguishing agricultural from non-agricultural work has also been used in other WIEGO-commissioned analysis (Heintz, personal communication, 8 October 2010). The concern lies, in particular, with self-employed workers as information on these workers tends to be less reliable than that obtained about wage employment. WIEGO has used different variables across countries in identifying agriculture in addition to the industry variable, with the choice of variables informed by the survey instrument and examination of the data.

## Identifying Urban Employment

WIEGO is especially concerned with urban informal employment. Heintz and Posel do not disaggregate their estimates in this respect. Wills, in contrast, provides a simple distinction between metropolitan (metro) and non-metro areas.

For the LFS, Statistics South Africa sampled in a way that, in theory at least, would allow for separate estimates for each of the six metropolitan areas. The dataset contained a variable that indicated the district council, and each metropolitan area was categorized as a separate district council.

Statistics South Africa’s sampling approach changed with the introduction of the QLFS. Sampling is now done by province and, within each province, sampling is done separately for metropolitan and non-metro areas. This allows

estimates for a particular metro area to be produced where that metro is the only such area in a province, as is the case for Cape Town, eThekweni and Nelson Mandela Metropolitan area. The new sampling approach does not, however, allow for separate (reliable) estimates for each of the three metropolitan areas in Gauteng, namely Johannesburg, Tshwane and Ekurhuleni.

A further constraint is that, to date, Statistics South Africa has not included the variable indicating the metro distinction in the standard datasets that can be downloaded from its website. Those wishing to conduct analysis of this sort have thus had to request preparation of a special dataset. There are reportedly plans to make the variable concerned available to the general public in the standard dataset.

Metropolitan areas do not encompass all urban areas. Thus “metropole” is a political and administrative construct, rather than one defined on the basis, for example, of population density. The metro political-administrative areas have high population density overall, although they might encompass some less densely-settled areas. Conversely, there are densely-settled urban areas that do not form part of the metros. Unfortunately, Statistics South Africa has reportedly been developing an official definition of “urban” for about a decade, but has not yet finalized this task.

## **Other Differences Between LFS and QLFS**

An important change from the LFS to the QLFS is that non-market production activities are not considered to constitute employment in the latter unless the person is also engaged in market production. The Statistics South Africa publications define non-market production activities as those carried out solely for the consumption/use of the household. Further, those identified as engaged in market production activities do not answer the questions about whether or not they were engaged in non-market production activities during the reference week. This gives an under-count of engagement in non-market production.

Unfortunately, we cannot directly measure the difference that the change in approach makes to overall employment as we do not have a single questionnaire that includes all the necessary questions. To get a sense of the size of the difference, we compare employment as recorded in the LFS of March 2007 and the QLFS for the second quarter of 2008. We find that the QLFS gives a slightly higher employment rate (42%) than the LFS (40%). The employment rate is two percentage points higher in the QLFS than the LFS for males (50% and 48% respectively) and three percentage points higher in the QLFS than the LFS for females (35% and 32% respectively). This comparison suggests that the omission of non-market production does not affect employment estimates in any substantial way in South Africa.

The QLFS also no longer provides for the two rates of unemployment which were (sometimes) reported for the LFS, i.e. the “official” rate and the “expanded” rate. Instead the QLFS introduces the category of “discouraged” workers who are placed within the “not economically active” category. This category includes those who give, as the main reason why they did not try to find work or start a business in the reference period, that there were no jobs available in the area, that they were unable to find work requiring their skills, or that they had lost hope of finding any kind of work. While the category at first glance seems similar to those who were broadly but not narrowly unemployed, Yu’s analysis shows a sharp fall in numbers in this category between the LFS and QLFS. Yu concludes that “using the new methodology might not be comparable with the LFS methodology” (Yu 2008).

Our own comparison using the LFS of March 2007 and the QLFS for the second quarter of 2008 finds that both record 13 per cent of the population aged 15 years and above as being unemployed using the narrower definition. (If we include a decimal place, a small difference emerges, with 13.4 per cent unemployed in the LFS and 12.5 per cent in the QLFS). However, as suggested by Yu, there is a much bigger difference between the number recorded in the LFS as unemployed using the expanded definition and the combined total of those who are classified as unemployed and as discouraged workers in the QLFS. The LFS records 24 per cent of the population aged 15 years and above as unemployed using the expanded definition, compared to 16 per cent in the QLFS. The disparity between the two estimates is larger for women (16% versus 27%) than for men (15% versus 21%).

## **Recommendations for Improvement of QLFS Data**

Realistically, we cannot expect Statistics South Africa to revert to the LFS questionnaire in respect to issues such as the questions that define employment and unemployment (both official and expanded). Further, even if such reversion were feasible, it would have the drawback of creating another “break” in the series. This would hamper trend analysis just as other changes in instrument and methodology created breaks in 2001 with the introduction of the LFS.

Statistics South Africa has, however, already accepted the need to bring back several aspects that were omitted when the QLFS was introduced. These are:

- the earnings question and variable. Statistics South Africa apparently reintroduced the question in 2009, but has not yet released the related variable in the publicly available dataset.
- the geographic variables. Statistics South Africa reportedly has plans to make, at the least, a variable distinguishing metro areas available.

Data relating to these two reversions should be made available as soon as possible to avoid the existing interruption in the trends relating to these variables being extended.

The wording of the earnings question in the LFS was not ideal. In particular, the wording implied wage or salary earnings, although the question was addressed to all workers, both employees and self-employed. This weakness is serious enough that the benefits of improved wording might outweigh the disadvantages of a break in the series, especially as there will already be a break. Statistics should thus consider rephrasing the question, or asking separately for employees and others.

Statistics South Africa should re-introduce the question about location of work. This is a single question that provides very useful information for understanding informal work in particular.

Statistics South Africa should also re-introduce the question about trade union membership. This will, among other things, allow analysis of the extent to which union membership is associated with formal employment.

As discussed above, Statistics South Africa's derivation of informal employment is not in line with the description to the official *Guide*. Further, the agency plans to review the way the variable is derived. Again, this should be done as speedily as possible, but the approach should include provision for consultation with analysts who have experience and knowledge in this area. Once the new approach is agreed on, the *Guide* should be amended so that the dataset and *Guide* are aligned.

In terms of disaggregation, we cannot realistically expect Statistics South Africa to redesign its sampling approach for the QLFS, or to increase the sample size. It is likely we must accept that the QLFS will not be able to provide metropolitan data beyond the current level of disaggregation – i.e. metro versus non-metro for each province. We can, however, urge that:

- Statistics South Africa make the metropolitan variable available more or less immediately in the public datasets; and
- Statistics South Africa move speedily to agreeing on a definition of urban and non-urban, so that this variable can be included in future datasets. The planned 2011 population census seems an ideal opportunity to finalize a definition.

## References

Heintz, James and Dorrit Posel. March 2008. "Revisiting Informal Employment and Segmentation in the South African Labour Market" in *South African Journal of Economics* Vol. 76, No. 1: 26-44.

Statistics South Africa. August 2008. *Guide to the Quarterly Labour Force Survey*. Pretoria.

Wills, Gabrielle. 2010. *South Africa's Informal Economy: A Statistical Profile*. WIEGO Urban Policies Research Report, No. 7.

Yu, Derek. 2008. *The Comparability of Labour Force Survey (LFS) and Quarterly Labour Force Survey (QLFS)*. Stellenbosch Economic Working Papers: 08/09. University of Stellenbosch.

**About WIEGO:** Women in Informal Employment: Globalizing and Organizing is a global research-policy-action network that seeks to improve the status of the working poor, especially women, in the informal economy. WIEGO draws its membership from membership-based organizations of informal workers, researchers and statisticians working on the informal economy. For more information see [www.wiego.org](http://www.wiego.org).

**About Inclusive Cities:** Launched in 2008, the Inclusive Cities project aims to strengthen membership-based organizations (MBOs) of the working poor in the areas of organizing, policy analysis and advocacy, in order to ensure that urban informal workers have the tools necessary to make themselves heard within urban planning processes. Inclusive Cities is a collaboration between MBOs of the working poor, international alliances of MBOs and those supporting the work of MBOs. For more information see [www.inclusivecities.org](http://www.inclusivecities.org).