



International
Labour
Office
Geneva



Analysis of employment, real wage, and productivity trends in South Africa since 1994

Martin Wittenberg

CONDITIONS OF WORK AND EMPLOYMENT SERIES No. 45

INWORK

Inclusive Labour Markets, Labour Relations
and Working Conditions Branch

***Analysis of employment, real wage,
and productivity trends
in South Africa since 1994***

By Martin Wittenberg

DataFirst and School of Economics
University of Cape Town, South Africa

Copyright © International Labour Organization 2014

Publications of the International Labour Office enjoy copyright under Protocol 2 of the Universal Copyright Convention. Nevertheless, short excerpts from them may be reproduced without authorization, on condition that the source is indicated. For rights of reproduction or translation, application should be made to the Publications Bureau (Rights and Permissions), International Labour Office, CH-1211 Geneva 22, Switzerland. The International Labour Office welcomes such applications.

Libraries, institutions and other users registered in the United Kingdom with the Copyright Licensing Agency, 90 Tottenham Court Road, London W1T 4LP [Fax: (+44) (0)20 7631 5500; email: cla@cla.co.uk], in the United States with the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 [Fax: (+1) (978) 750 4470; email: info@copyright.com] or in other countries with associated Reproduction Rights Organizations, may make photocopies in accordance with the licences issued to them for this purpose.

ILO Cataloguing in Publication Data

Wittenberg, Martin

Analysis of employment, real wage, and productivity trends in South Africa since 1994 / by Martin Wittenberg ; International Labour Office, Inclusive Labour Markets, Labour Relations and Working Conditions Branch. - Geneva: ILO, 2014

Conditions of work and employment series, No.45, ISSN: 2226-8944; 2226-8952 (web pdf)

International Labour Office; Inclusive Labour Markets, Labour Relations and Working Conditions Branch

labour market analysis / labour statistics / data collecting / wages / employment / hours of work / labour productivity / trend / South Africa R

13.01.1

First published 2014

Cover: DTP/Design Unit, ILO

The designations employed in ILO publications, which are in conformity with United Nations practice, and the presentation of material therein do not imply the expression of any opinion whatsoever on the part of the International Labour Office concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its frontiers.

The responsibility for opinions expressed in signed articles, studies and other contributions rests solely with their authors, and publication does not constitute an endorsement by the International Labour Office of the opinions expressed in them.

Reference to names of firms and commercial products and processes does not imply their endorsement by the International Labour Office, and any failure to mention a particular firm, commercial product or process is not a sign of disapproval.

ILO publications can be obtained through major booksellers or ILO local offices in many countries, or direct from ILO Publications, International Labour Office, CH-1211 Geneva 22, Switzerland. Catalogues or lists of new publications are available free of charge from the above address, or by email: pubvente@ilo.org
Visit our website: www.ilo.org/publns

Printed by the International Labour Office, Geneva, Switzerland

Executive Summary

South Africa's transition to democracy in 1994 engendered expectations of rapid changes in racial and gender disparities in wages. Twenty years after the transition there are strong disagreements about whether there have been gains and how these have been distributed. In particular there are debates about whether wage increases have outstripped productivity increases.

Another legacy of apartheid which makes these debates more difficult is that there is little pre-1994 survey evidence that can be used as effective benchmarks for assessing progress. We show in some detail that since 1994 there have been notable changes in the quality of the data and that dealing properly with changes in sampling design, survey instruments and missing values is necessary in order to identify trends.

We show that real wages have increased over the period – however real median wages of employees have remained stagnant. The top tail of the earnings distribution has “fanned out”, with larger gains at the 90th percentile than at the 75th, but both showing good real earnings growth. At the bottom end of the distribution there seems to have been compression – with the 10th percentile making real gains relative to the median. Among the self-employed there is no evidence for systematic shifts in the distribution over the post-apartheid period.

The employment series is bedevilled by a huge break between October 1999 and February 2000, which coincides with the shift from the October Household Surveys (OHSs) to the Labour Force Surveys (LFSs). The LFSs find more marginal and informal activity than the OHSs did. However in the immediate 2000 period they seem to find an excess number of subsistence agricultural workers (around 800 000) which disappear in later surveys.

The much higher level of employment is confirmed in the total hours worked series. Many of the additional workers enumerated in the LFSs (particularly in subsistence agriculture) work lower hours than the norm.

The mean wage reported in the Quarterly Employment Statistics (QES) is almost double the mean wage from the Quarterly Labour Force Survey (QLFS). We examine various hypotheses about why they are so far apart. Differences in coverage, missing information, and various types of measurement error all help to explain parts of the gap, but even assuming some quite implausibly large magnitudes we cannot close the gap entirely. Undoubtedly this is a matter for further research. Given the discrepancies it is clear that statistics from the QES and the QLFS shouldn't be mixed and matched.

With these caveats we use Gross Value Added figures from the national accounts and total employment figures which we cleaned of the excess subsistence workers to construct labour productivity series. Looking at the sectoral figures there seems a clear pattern of divergence: sectors that were relatively low productivity ones in 1994 (Agriculture, Construction, Trade and Services) also show much lower labour productivity growth, while the remaining ones all seem to have achieved productivity gains. On average there is little evidence that wages have outstripped productivity or *vice versa*.

Table of contents

Executive Summary	iii
1. Introduction	1
2. Current labour market statistics: wages, employment, hours, and labour productivity	2
2.1 Earnings estimates and data issues: Literature Review	2
2.2 Surveys	3
2.2.1 The October Household Surveys	3
2.2.2 The Labour Force Surveys	4
2.2.3 The Quarterly Labour Force Surveys	5
2.2.4 The Survey of Employment and Earnings (SEE) and the Quarterly Employment Statistics (QES) survey	5
2.2.5 PALMS – Post-Apartheid Labour Market Series	6
2.3. Data Processing	6
2.3.1 Removal of outliers	6
2.3.2 Dealing with bracket responses and other types of missing values	6
2.3.3 Zero incomes	7
2.3.4 Deflation	8
3. Earnings and wages	8
3.1 Nominal and Real earnings over time	9
3.1.1 Real earnings	9
3.1.2 Nominal earnings	9
3.1.3 Real earnings – Wage Workers	11
3.1.4 Real earnings of the Self-Employed	11
3.2 Average earnings by major groups and sectors	12
3.2.1 Real earnings by Gender	12
3.2.2 Real earnings by Race	14
3.2.3 Real earnings by sector	15
3.3 The earnings distribution	18
3.3.1 Real earnings by selected percentiles of wage workers	18
3.3.2 Real earnings by selected percentiles of the Self-employed	21
4. Employment	23
4.1 Total Employment	23
4.2 Employment by Sector – Wage Workers	24
4.3 Employment by Sector – Self-employed	26
5. Working hours	28
5.1 Total Hours worked	28

5.1.1	Total Hours.....	28
5.1.2	Total Hours by Sector.....	29
5.2	Average hours: overall and by employment status.....	30
5.2.1	Average Hours by Sector – Wage workers	31
5.2.2	Average Hours by Sector – the Self-employed	33
5.3	Earnings per hour	34
5.3.1	Overall trends	34
5.3.2	Earnings per hour among employees by Sector	35
5.3.3	Earnings per hour among the self-employed by Sector.....	37
6.	Reconciling the firm level wage series with the micro household data series	38
6.1	Macro and micro data.....	39
6.2	The measurement used: Medians and means	41
6.3	Coverage differences.....	41
6.3.1	Individuals covered – Employees, Agriculture and Domestic Work	41
6.3.2	Sectoral coverage – a detailed look at the other sectors.....	42
6.3.3	Type of firms	42
6.3.4	Part-time work.....	43
6.4	Possible sources of measurement error	43
6.4.1	Missing high income earners in the QLFS	43
6.4.2	Other missing earners in the QLFS	44
6.4.3	Misclassification of employees in the QLFS.....	44
6.4.4	Missing earnings in the QLFS.....	44
6.4.5	Errors of inclusion in the QES.....	45
6.4.6	Errors of omission in the QES.....	45
6.5	Portioning out the sources of error	45
6.5.1	Coverage and measurement differences	45
6.5.2	Possible errors in the QLFS.....	45
6.5.3	Possible errors in the QES.....	46
6.5.4	Decomposing the difference.....	46
7.	Productivity	47
7.1	Sources used in the measurement of annual GVA	47
7.2	The measurement of annual GVA.....	50
7.3	The measurement of the Cost of Employment	51
7.4	The evolution of labour productivity in South Africa 1994-2011	52
7.4.1	Labour productivity using employment	52
7.4.2	Labour productivity using total hours worked	56
7.4.3	Comparing the productivity measures.....	59
7.5	The relationship between earnings and productivity.....	59

8. Conclusion.....	60
Appendix: Additional Statistics by Gender and Race	62
A.1 Employment by Gender and Race.....	62
A.2 Average Hours by Gender and Race	64
A.3 Earnings per hour by Gender and Race.....	65
References	67
Datasets	68
Conditions of Work and Employment Series.....	69

List of Figures

Figure 1.	The importance of dealing with outliers and brackets	8
Figure 2.	The impact of the weights on mean wages	9
Figure 3.	Nominal Earnings of Wage Workers	10
Figure 4.	Nominal Earnings of the Self-Employed	10
Figure 5.	Earnings of wage workers over time.....	11
Figure 6.	Earnings of Employers/Self-employed	12
Figure 7.	Real Earnings of Wage Workers by Gender	13
Figure 8.	Real Earnings among the Self-Employed by Gender.....	13
Figure 9.	Real earnings of Wage Workers by Race.....	14
Figure 10.	Real Earnings of the Self-Employed by Race	15
Figure 11.	Real Earnings of Wage Workers by Sector - Part 1	16
Figure 12.	Real Earnings of Wage Workers by Sector - Part 2	16
Figure 13.	Real Earnings of the Self-employed by Sector - Part 1	17
Figure 14.	Real Earnings of the Self-employed by Sector - Part 2.....	18
Figure 15.	Real Earnings of Wage workers over time at different percentiles	19
Figure 16.	Change in position of different percentiles relative to the median.....	20
Figure 17.	The gap between median and mean earnings among wage workers.....	20
Figure 18.	Real Earnings of the Self-employed at different positions of the earnings	21
Figure 19.	Change in position of different percentiles of the Self-employed earnings distribution relative to the median	22
Figure 20.	The ratio of median to mean earnings among the Self-employed.....	22
Figure 21.	Total employment.....	23
Figure 22.	Annualised change in employment - the big shift in February 2000.....	24
Figure 23.	Employment among Wage Workers, by Sector - Part 1.....	25
Figure 24.	Employment among Wage Workers, by Sector - Part 2.....	25
Figure 25.	Changes of employment among Wage Workers for selected sectors.....	26
Figure 26.	Numbers of self-employed by Sector - Part 1	27
Figure 27.	Numbers of self-employed by sector - Part 2.....	27

Figure 28.	Another take on employment in South Africa - total hours worked.....	28
Figure 29.	Total hours worked per week by sector - Part 1	29
Figure 30.	Total hours worked per week by sector - Part 2	30
Figure 31.	Hours worked - by employment status	31
Figure 32.	Average hours worked by wage workers, selected sectors - Part 1	32
Figure 33.	Average hours worked by wage workers, selected sectors - Part 2.....	32
Figure 34.	Average hours worked among the Self-employed, by Sectors - Part 1	33
Figure 35.	Average hours worked among the Self-employed, by Sector - Part 2.....	34
Figure 36.	Real earnings per hour, 1994-2011	35
Figure 37.	Real earnings per hour by sector among employees, 1994-2011 - Part 1	36
Figure 38.	Real earnings per hour by sector among employees 1994-2011 - Part 2	36
Figure 39.	Real earnings per hour by sector among the self-employed 1994-2011 - Part 1.....	37
Figure 40.	Real earnings per hour by sector among the self-employed 1994-2011 - Part 2.....	38
Figure 41.	Comparing the firm surveys to the household surveys.....	39
Figure 42.	Nominal earnings over time in the firm surveys and household surveys	40
Figure 43.	The National Accounts CoE series and the Gross Earnings Series from	52
Figure 44.	Annual Employment numbers	53
Figure 45.	Labour productivity: 1994 to 2012.....	54
Figure 46.	Labour Productivity by Sector 1994-2012 - Part 1.....	55
Figure 47.	Labour Productivity by Sector 1994-2012 - Part 2.....	55
Figure 48.	Smoothing over the spike in 2000	56
Figure 49.	Labour productivity measured as output per hour worked.....	57
Figure 50.	Labour productivity measured in output per hour worked by sector, Part 1	58
Figure 51.	Labour productivity measured in output per hour worked by sector, Part 2	58
Figure 52.	Comparing labour productivity trends when using hours worked or employment	59
Figure 53.	The evolution of productivity and real earnings, 1994 to 2011.....	60
Figure 54.	The shift up with the LFS is particularly noticeable among women.....	62
Figure 55.	The shift up is most dramatic for Africans (left axis) but less noticeable among other groups (right axis)	63

Figure 56.	Self-employment levels by Race.....	63
Figure 57.	Average hours worked by gender.....	64
Figure 58.	Average hours worked, by Population Group.....	65
Figure 59.	Real earnings per hour 1994-2011, by gender.....	65
Figure 60.	Real earnings per hour 1994-2011, by Race.....	66

1. Introduction

The relationship between changes in wages, productivity levels and inequality has again moved centre-stage in many countries. The big picture is that inequality has increased in many places in the last few decades with the income gains concentrated disproportionately at the top. South Africa, as we show below, follows this global pattern. Earnings have improved rapidly for high incomes, while they have been stagnant at the median, leading to a growing gap between the top and the middle of the income distribution. Nevertheless South Africa is also different given the fact that there were strong expectations of more redistributive outcomes following the collapse of Apartheid in 1994. Unlike many other countries the policy stance was not hostile to unions.

Indeed in 1994 the African National Congress came to power in an explicit alliance with the strongest union movement, the Congress of South African Trade Unions, promising to undo a legacy of policies which had been explicitly anti-black unions. Whereas collapsing unionisation rates have been blamed for increasing inequality in some OECD countries, this cannot be the mechanism in South Africa.

Of course the Apartheid period had also been marked by explicit racial discrimination, so the new government committed itself to a series of affirmative action policies. Consequently it was also expected that the entrenched racial and gender disparities in the work place would be gradually eliminated. The evidence shown below suggests that the changes have not been as rapid as might have been anticipated.

Along with the persistence of racial and gender differentials, the democratic government has found it difficult to markedly reduce poverty. Consequently there have been renewed calls for a much more aggressive minimum wage policy, with suggestions that the minimum wage should be tied to a specific percentage of the average wage as reflected in the Quarterly Employment Statistics. This approach has been questioned by others on the basis that wage growth has outstripped productivity growth. Questions of the relationship between average wages, wage growth and productivity have therefore moved centre-stage. In this South Africa is not all that different from other countries where questions about the distribution of benefits from growth have arisen.

The South African case, however, offers another perspective on these questions. In order to talk about the long-run trends one has to pay careful attention to the quality of the data and to measurement issues. These are thrown into particularly sharp relief in South Africa given that there is little credible pre-1994 information and the quality of the survey evidence has changed over the post-Apartheid period. If South Africa teaches one new lesson to the global debate, it is that good quality measurement is absolutely vital if one wishes to even establish what the trends are. As we show below, changes in sampling design, in the measurement instrument and in the way missing values are dealt with can all have enormous impacts on the picture that is presented.

The plan of this paper is as follows: in section 2 we will review some of the literature and discuss many of the data issues. Section 3 then looks at the evidence of the evolution of earnings and wages since 1994. Sections 4 and 5 deal with employment and working hours respectively, while section 6 discusses some major discrepancies between the micro and the macro data. Section 7 then looks at the evolution of productivity and section 8 concludes.

2. Current labour market statistics: wages, employment, hours, and labour productivity

The Apartheid system was not only designed to politically subjugate the majority of the South African population and to keep them economically marginal and subservient, it also inhibited the systematic collection of information. Population censuses ceased to cover the entire population in 1980, with separate “censuses” being delegated to different homeland administrations. Furthermore there were no nationally representative surveys, nor were the microdata released until the World Bank funded “Project for Statistics on Living Standards and Development” (PSLSD) in 1993. The first nationally representative survey run by the Central Statistical Services (renamed Statistics South Africa a short while later) was in 1994. One of the problems encountered by these early data gathering exercises was that there was no agreed population frame from which sampling could proceed until the 1996 census. Furthermore the national statistical agency still had to develop the systems to implement and run these surveys. So although there are data from nationally representative surveys run by Statistics South Africa since 1994 there are discontinuities in the series which need to be taken into consideration when analysing the data.

2.1 Earnings estimates and data issues: Literature Review

While there is no shortage of literature that uses the earnings information in the October Household Surveys (OHSs) and Labour Force Surveys (LFSs), e.g. to estimate union wage effects or returns to education, there are not that many articles in the academic literature that attempt to look at the earnings series and assess the longer run trends. Many of the papers that discuss longer-run trends (e.g. Mazumdar and van Seventer 2002, Klein 2012) use the wage series emanating from the firm surveys (Survey of Employment and Earnings and subsequently the Quarterly Employment Statistics). We will discuss the relationship between the microdata and the macro series in the data section. For the moment we should note that mixing employment data (from the QLFS) and earnings data (from the QES) is probably not desirable, given that the two surveys cover different populations.

The most comprehensive review of the measurement of earnings in the post-Apartheid period is by Wittenberg and Pirouz (2013), which serves as the basis for this discussion¹. Among papers that have looked at the earnings information in the OHS/LFS, Casale (2004) remarked on a precipitous decline in real earnings over the period 1995 to 2001. She notes that better enumeration of low wage work may explain some of the decline, but argues that

“the fall in informal self-employment earnings between 1995 and 2001 is unlikely to be the result of improved data collection alone, as more and more people crowding into already low income-generating informal activities would be expected to depress average earnings even further “. (p.264)

Burger and Yu (2007) provide the most detailed review beforehand. They argue that the bulk of the decline in earnings is driven by the discontinuity between the October Household Surveys (OHSs) and the Labour Force Surveys (LFSs). They also note that outliers contaminate the series, a problem which they argue is particularly noticeable in the September 2000 LFS. In other regards the LFS series is less prone to extreme incomes. Indeed they argue that people earning more than R1 million a year affect average earnings in all years prior to 2000. They suggest that this was related to changes in

¹ It has additional technical details about some of the measurement changes over the 1994 to 2012 period and about the construction of the earnings series that are used in this paper.

the earnings brackets available to the self-employed. If the “millionaires” are removed from the series it becomes a lot smoother over time.

The question as to how respondents who gave an earnings interval only are treated in these analyses is not transparent at all, although it matters considerably for the results. We will discuss this issue in more detail below.

A number of contributions in the literature have commented on changes in the instrument. Daniels (2013) investigates changes in the earnings question addressed to employees in the 1997-1999 October Household Surveys and the 2000-2003 Labour Force Surveys. He focuses particularly on how missing information is dealt with. Unfortunately some rather large changes in the instrument precede this period (as shown below) and one of the biggest breaks between the October Household Surveys (OHSs) and the Labour Force Survey (LFSs) is the collapse of two separate questions (in the OHSs) to just one question in the LFSs.

Yu (2007) provides a detailed discussion of many variables measured in both the OHSs and LFSs, including earnings. He shows that there was an increase in the reporting of zero incomes and unspecified responses with the change to the LFSs. Indeed the February 2000 survey seems particularly anomalous with regard to both (Yu 2007, Table 24, p.27). His companion piece Yu (2009) discusses the relationship between the LFSs and the Quarterly Labour Force Surveys (QLFSs) introduced in 2008. He does not consider the question of earnings, because in the period he reviewed there was no earnings information available for the QLFSs.

Anthea Heap (2008) analyses earnings inequality over the period 1995 to 2006. She ends up discarding the 1995 data because it looked anomalous. In particular there does not appear to be a male earnings premium in that year. Indeed other research has also suggested that the 1995 OHS may be anomalous. Wittenberg (forthcoming) notes that employment looks higher and household services look better in 1995.

2.2 Surveys

The data on employment, wages and hours worked come from the October Household Surveys, Labour Force Surveys and the Quarterly Labour Force Surveys. These datasets have been assembled and harmonised as “PALMS” (PostApartheid Labour Market Series), see Kerr *et al* (2013). Information in the “macro” wage series (reported, for instance, in the South African Reserve Bank Quarterly Bulletin) comes from the firm surveys – Survey of Employment and Earnings (SEE) and the Quarterly Employment Statistics (QES) survey. We will therefore discuss those surveys also.

2.2.1 The October Household Surveys

The October Household Surveys were designed as multi-purpose instruments. They typically covered around 30 000 households (with exceptions noted below). They had a labour module but also asked questions about household services and household composition. The nature of the questionnaire and the sample design show considerable changes over the period. Some of the noteworthy idiosyncrasies are:

- OHS 1994
Documentation for this survey is almost non-existent. It is clear that the earnings information in this survey is heavily imputed. Unfortunately the imputations are not very sensible.

Undoing them takes considerable effort (see Wittenberg 2008), but is possible. Unlike all subsequent surveys, the income information given is **net income** (i.e. take home pay).

- OHS 1995
The earnings brackets used in this survey are extraordinary, to say the least. The bottom bracket extends from R1 up to R999, which covers a huge chunk of the wage distribution. It is possible that poor technique in dealing with the bracket information may be to blame for some of the anomalous findings about earnings.
- OHS 1996
This survey was a much smaller survey (around 16 000), given that it ran immediately after the census. Indeed it was linked to the Post-Enumeration Survey. The instrument was also much simpler and earnings were only recorded in intervals.
- OHS 1997 and 1998
The 1998 survey was again smaller than the average (around 20000 households). Both these surveys (and indeed all the surveys prior to 1999) asked for **gross turnover** in the self-employment question and then asked for any deductions against that turnover.
- OHS 1999
This was the first survey to be run under the new “Master sample”. We discuss some of the coverage changes that resulted below.

Coverage issues in the OHSs

Kerr and Wittenberg (2013a) discuss several fieldwork practices in the early OHSs which led to a systematic underenumeration of individuals living in small households. Chief among these was the practice of only interviewing one household if there were several living at a particular “visiting point” (dwelling). Households were selected proportional to their size and there is no traceable “upweighting” of the smaller households to compensate. On top of this it is understood that the early OHSs did not sample hostel residents properly. The combination of these factors means that there is a very sharp rise in the number of small households (in particular of one person households) in the period October 1997 to February 2000. The missing individuals were more likely to be working (or unemployed) than not economically active. Consequently the same period also sees a very sharp rise in participation and employment. Unfortunately it is not possible to correct properly for these changes. The appropriate method for dealing with it is a matter of ongoing research.

2.2.2 The Labour Force Surveys

The Labour Force Surveys replaced the OHSs from 2000. Instead of annual surveys these were run every six months (initially February and September, then March and September). One of the big changes in relation to earnings information was that two separate earnings questions in the OHS (one for employees and one for employers/the self-employed) were collapsed into a single question. An additional change is that more care was taken to screen for any possible form of economic activity. Consequently the enumeration of informal sector workers is better in the LFSs.

There are also some idiosyncracies that need to be noted. The February 2001 LFS was run in conjunction with the “Survey of Employers and the Self-employed”. Fieldworkers were paid extra for completing the latter, so had extra incentive to find self-employed individuals. The result was a once-off spike up in the size of the measured informal sector which has not been matched since.

2.2.3 The Quarterly Labour Force Surveys

Since 2008 the Labour Force Surveys have been replaced by the Quarterly Labour Force Surveys. The earnings question was initially dropped from the survey instrument, although it was reinstated in the third quarter of 2009. The earnings information was, however, not released with the QLFS data itself. Instead it was released, considerably later, as the “Labour Market Dynamics Study”.

The earnings question was changed once more. The question was again asked separately of wage workers and employers/self-employed workers. Unlike in the OHSs, however, it is not possible for a respondent to answer both.

Unlike the OHSs and LFSs, the publicly released data imputes for people who only responded in brackets, as well as those who refuse or report that they do not know their earnings. It is impossible to retrospectively separate them. The imputation routine is not discussed but it is likely to be some form of “hot deck”², given that there appear to be repeat values, particularly in the upper tail.

2.2.4 The Survey of Employment and Earnings (SEE) and the Quarterly Employment Statistics (QES) survey

Firm surveys have traditionally been the source of the macroeconomic remuneration series. The surveys are done in a stratified way, so ensuring almost complete inclusion of large firms (which make up a big proportion of total employment and remuneration). Kerr and Wittenberg (2013b) suggest that between 45 percent and 55 percent of the total number of formal, non-agricultural, private sector workers are directly captured by the firms included in the QES samples between 2005 and 2011. There is also almost 100% inclusion of public sector employment in the **sample** (i.e. before any weighting adjustments are made). Because the sample by itself accounts for such a large proportion of the formal sector the QES figures should potentially provide a very accurate picture of total remuneration and employment patterns.

However firm surveys have their own coverage issues. The frame from which the original Survey of Employment and Earnings sample was drawn became progressively unrepresentative of the South African economy. Consequently Statistics South Africa spent considerable resources on constructing a new business register which provided the frame from where the new QES survey was extracted. Furthermore a bigger attempt was made to include small businesses into the frame for the QES.

Nevertheless coverage in the QES and QLFS is undoubtedly different, with the QES reporting around 8.4 million workers in the first quarter of 2012 against 13.9 million people classified as “employed” in the QLFS. The agricultural sector is excluded from the QES and only firms with turnover of more than 1 million rand are included. Thus the difference will be made up of informal sector workers and agricultural workers and probably some casual workers. This has obvious implications for the earnings reported in the two surveys. These are quite far apart: Statistics South Africa’s “Labour Market Dynamics in South Africa in 2011” only reports median earnings. These were around R3000 per month (Table 4-38, p4-51). By contrast the average wage reported in the QES for the same period was in excess of R12 000 per month. Below we will explore some reasons why these figures may be so far apart.

² In a “hot deck” imputation missing values are replaced with draws from the nonmissing values, where observations are matched according to similarity on some covariates (e.g. same earnings bracket, gender, race etc.)

2.2.5 PALMS – Post-Apartheid Labour Market Series

This is actually just the entire labour market survey sequence (OHSs, LFSs, QLFSs) stacked together with some attention paid to harmonising the coding and definition of variables. One additional “value added” variable is included – a weighting variable that doesn’t embed inconsistencies within it. The Statistics South Africa datasets are released with weights calibrated to the prevailing demographic model. As the baseline information changes (particularly with new census information), the demographic model changes likewise. This means that later surveys are calibrated to totals that are no longer comparable with those from earlier surveys. This should not affect means and proportions to a large extent, but it will affect totals. Branson and Wittenberg (forthcoming) discuss some of the inconsistencies that can arise from using weights that are not strictly comparable. In order to deal with this, the original Statistics South Africa released weights have been recalibrated to the ASSA 2008 demographic model using a cross-entropy (CE) approach. As will be discussed later, the weights selected do not have a major impact on the means (see Figure 2 in section 3.1.1).

2.3 Data Processing

In order to produce the analyses reported below several “data cleaning” steps were necessary, in particular the removal of outliers and dealing with the “bracket” responses.

2.3.1 Removal of outliers

Burger and Yu (2007) drew attention to the role of extreme values in contaminating the “average wage” series over time. Simply taking out “millionaires” (in constant 2000 prices) seems arbitrary and will potentially understate inequality (by removing genuine millionaires). In order to deal with this in a non-arbitrary way we estimated a simple Mincerian type wage regression on **all** the monthly (log) real earnings figures across all surveys. We then created studentised residuals (i.e. residuals normalised against their residual standard deviation, but calculated from a regression in which that observation is left out). Observations that had an absolute studentised residual in excess of 5 were marked as outliers. If the underlying distribution was normal we would expect to see such extreme values by chance less than once in our sample. The observations thus flagged (around 400 out of around 600 000) had earnings that were far too high, given their education, age and occupation, or implausibly low.

2.3.2 Dealing with bracket responses and other types of missing values

Individuals who refused to give a Rand amount but indicated an earnings bracket needed to be brought into the analysis. Three approaches were used:

a) Mean imputation within bracket

The idea in this case is to assign all the “bracket responders” the mean of the values recorded by other respondents who fell into the same bracket but gave actual Rand amounts. This was done separately by survey. In the case of the 1996 October Household Survey (which only had brackets) they were assigned the bracket means of the 1997 OHS real earnings values. The main advantage of this technique is that it is quick and easy. The main disadvantage is that it creates “spikes” in the data which affect distributional measures other than the mean. Furthermore if within each bracket there are systematic differences between certain groups of people (e.g. men and women), then the overall relationship may become obscured. Consequently we will generally **not** use this approach except as a quick check on calculations done with the other two approaches. We will, however, use both of the

next two approaches, so that we have two conceptually different ways of “correcting” the data. This will give us some confidence that the trends that we are seeing are not an artefact of a particular measurement approach.

b) *Reweight the people giving actual values to represent the bracket responses*

This is also fairly straightforward. It has the virtue that estimation always happens only on actual values. Like the mean imputation response it maintains the mean within a bracket. It also has the virtue that it doesn’t create any artificial spikes in the data. Furthermore it doesn’t create combinations of variables (e.g. gender and income) that do not exist in the “raw” data already. It has one disadvantage in that it cannot deal with cases where there are no Rand values to be reweighted. This means that the OHS 1996 cannot be put into the series.

c) *Multiple imputation: missing values*

The most complicated procedure adopted was to impute for the missing values by Stata’s “multiple imputation” routines³. The sequence adopted was to first impute bracket categories to those who refused to even indicate a bracket (the “don’t know”, “refused” and “unspecified” categories) as well as the outliers (since their responses were classified as bad data). This was done by an ordered logit using province, gender, education, population group, a quadratic in age and occupation as explanatory variables and was done separately for each survey. Instead of simply predicting the bracket, the draw happens from the “posterior distribution”⁴. In this way the uncertainties in the modelling are directly taken into consideration.

In the second step of the imputation each individual missing a Rand amount but that (now) has a bracket was assigned a value by “predictive mean matching”. This is a variant of a hot deck. Instead of directly matching on characteristics (bracket, education, gender and type of employer), individuals are matched on their predicted draws (again taken from the posterior distribution). Despite the random draws, it is very unlikely that individuals would ever be matched with anyone outside their bracket. This process was also done within each survey separately.

The 1996 responses were dealt with by “predictive mean matching” of the 1996 bracket respondents with respondents providing Rand amounts in the 1997 OHS. The real earnings figures (i.e. deflated values) were matched, so that inflation between the periods is controlled for to some extent.

Ten “replicates” for the entire PALMS wage series were drawn in this way. In the analysis all statistics are calculated over all replicates. One advantage of this procedure is that it is possible to explicitly correct for individuals missing incomes other than those providing only earnings bracket information.

2.3.3 Zero incomes

Vermaak (2012) notes that imputing values for individuals providing zero incomes makes an appreciable difference to the earnings distribution. Since the pattern of people reporting zero incomes varies considerably over the surveys it is unlikely that they are providing good data. On the other hand

³ The theory is given by Rubin (1987). A more accessible introduction is given by the Stata multiple imputation manual (StataCorp 2013).

⁴ We can sketch the idea out as follows. If we try to predict from an ordinary linear regression, this is normally taking the regression coefficients as being “the truth”, e.g. $\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k$. When we draw from the “posterior” distribution we allow for the fact that the coefficients could be wrong – indeed we have some sense how far off they could be through our estimated standard errors. In effect we take a draw for the coefficients from a normal distribution centred on our estimates allowing for this error. Doing this multiple times gives us a range of plausible values for our prediction.

it seems dubious to assume that someone who explicitly records a zero should really have recorded a positive amount. It seems more likely that this response reflects that the data comes from a different “data generating process” – e.g. unpaid family workers, individuals whose attachment to the job is tenuous (maybe on a waiting list). The bulk of these are, in fact, subsistence agricultural workers (see Wittenberg and Pirouz 2013). Within the constraints of the current exercise we thought it was more prudent to exclude them.

2.3.4 Deflation

We chose to deflate earnings to constant 2000 prices using the Consumer Price Index. Using an “across the board” adjustment may not be correct if there are systematic differences in inflation, e.g. between urban and rural areas, but using one national rate per survey has the virtue of not distorting results within a particular survey. Furthermore there are no consistently measured regional sub-indices for the entire period.

A picture of what the different adjustments do is given in Figure 1. The original series (indicated with dots) shows some really strange spikes in the early 2000 period. All these spikes disappear, however, once the outliers are removed. It is also evident that it is really important to correct for brackets. The three “corrections” all yield very similar series, while the series which takes out the outliers but doesn’t deal with the brackets is around R800 per month below the adjusted values.

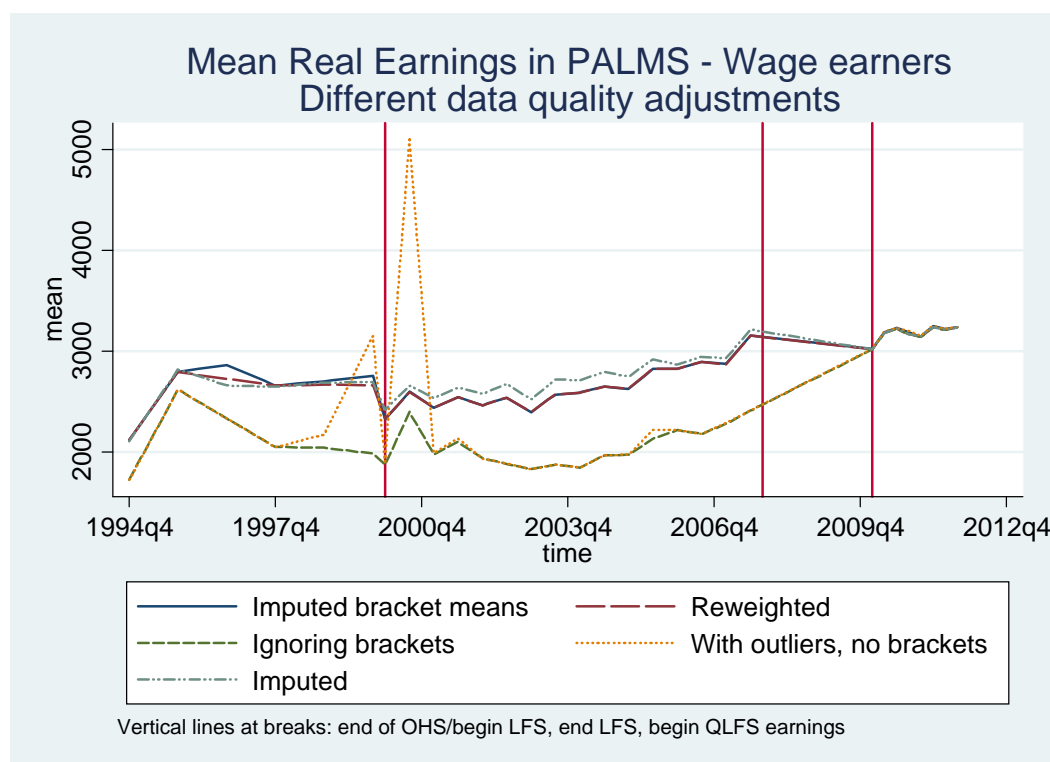


Figure 1. The importance of dealing with outliers and brackets

3. Earnings and wages

In the analyses below we present the information largely graphically. Spreadsheets with the data series can be obtained on request. In the graphs we typically label the one series “reweighted” and the other “imputed”. It should be stressed that **both** series apply the recalibrated sampling weights, i.e.

they both are “representative” of the corresponding South African population. However they deal with the bracket responses differently as discussed above – the “reweighted” series makes an additional correction to the sample weights while the “imputed” series presents the average from multiple imputations.

3.1 Nominal and Real earnings over time

3.1.1 Real earnings

The overall pattern of real earnings is shown in Figure 2. The graph shows the evolution of three different series: the overall mean, the mean among wage workers and the mean among employers/self-employed. The graph also shows the difference between the recalibrated weights (labelled “CE”) and the original ones. It is evident that these weights do not drive any of the major changes. A cursory glance shows that the series for the self-employed is by far the most volatile. Indeed self-employment earnings seem to collapse with the shift from the OHSs to the LFS. The main reason for this is almost definitely changes in coverage, particularly the increase in the enumeration of low paid informal sector workers. The OHS part of the “self-employed” series is therefore probably an over-statement of earnings among the self-employed. It probably reflects mainly the “formal” part of the self-employment spectrum.

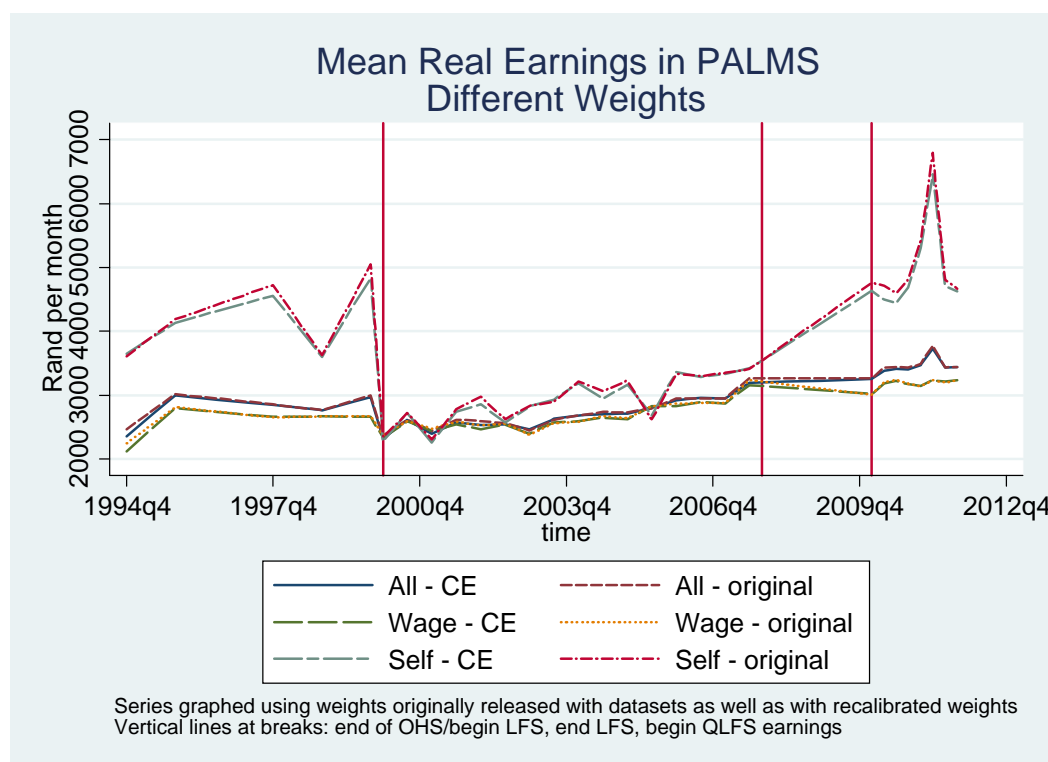


Figure 2. The impact of the weights on mean wages

3.1.2 Nominal earnings

The nominal earnings series are given in Figure 3 and Figure 4. The series for wage workers looks unremarkable, except for a pronounced dip in February 2000. The series for the self-employed, by contrast, shows that the break between the October Household Surveys and Labour Force Surveys

implies a huge drop in nominal earnings. As noted above, the simplest explanation for the drop is the big change in the coverage of low earning workers in the informal sector.

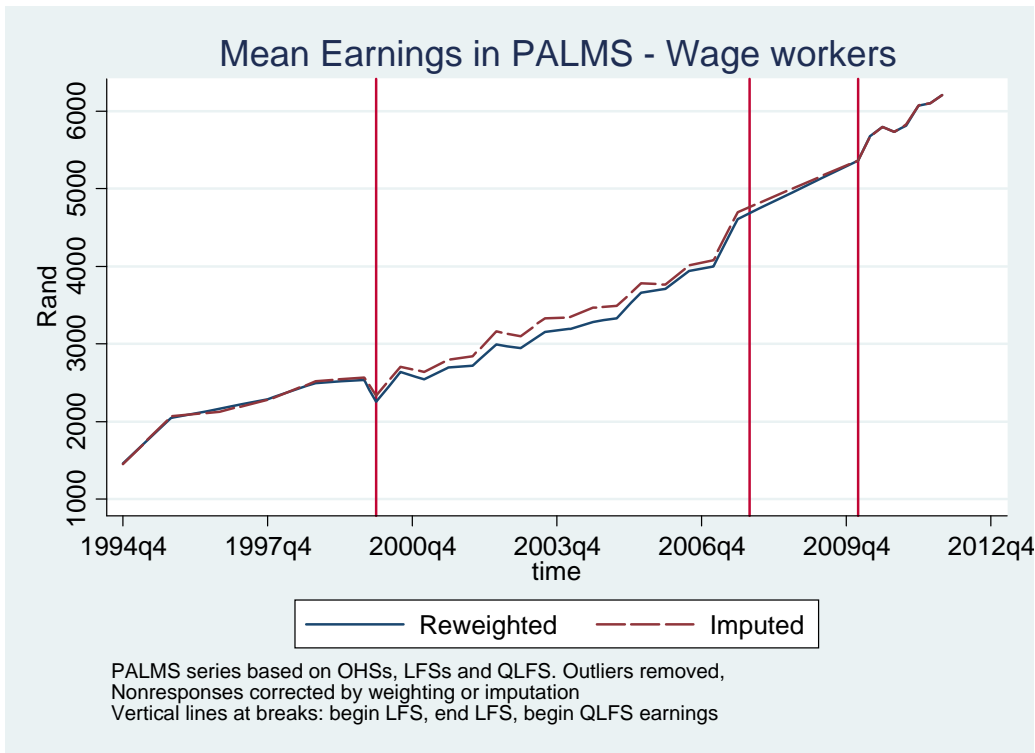


Figure 3. Nominal Earnings of Wage Workers

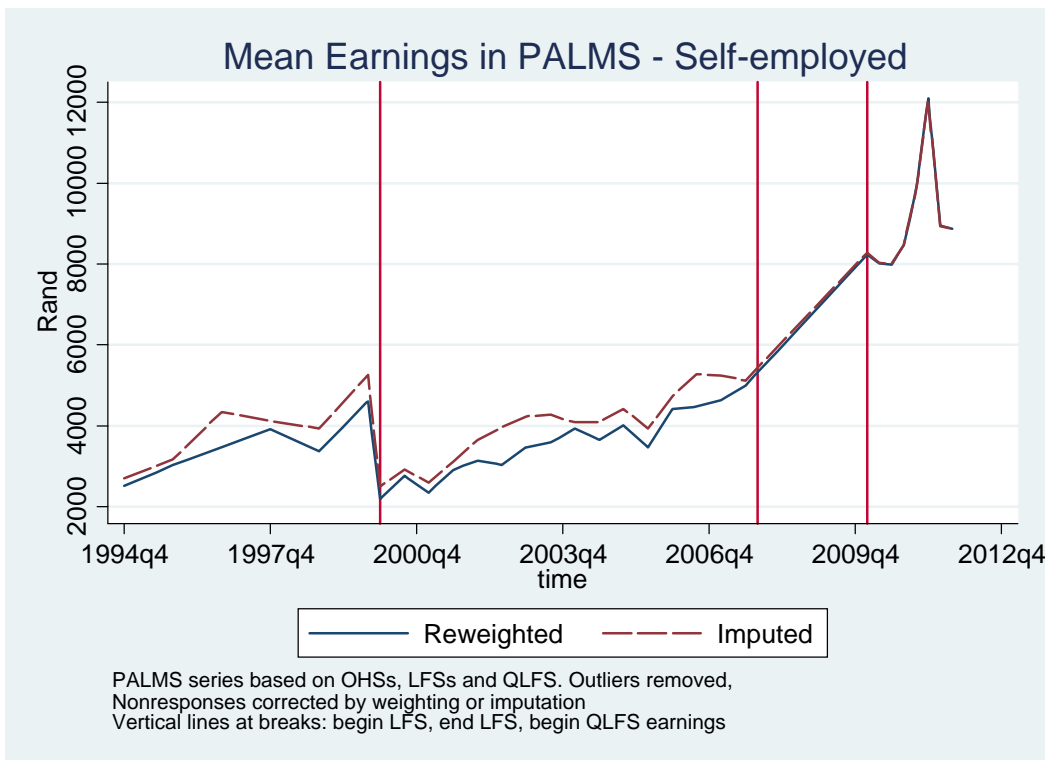


Figure 4. Nominal Earnings of the Self-Employed

3.1.3 Real earnings – Wage Workers

A more detailed look at the “wage worker” part of the series is shown in Figure 5. Several points deserve comment. Firstly the 1994 starting point is too low – for that year only “net income” values are available whereas all subsequent years ask about income before deductions.

Secondly the figures for October 1995 seem too high. If we focus on the change between that peak and the Labour Force Surveys (Casale 2004 uses September 2001) then there does seem to be a dramatic reduction in real wages. Looking at the entire series, however, it appears that the “collapse” is an artefact of an implausible starting point. Even ignoring 1995 the OHS levels are probably a bit too high which, given the discussion of sample coverage above, is not altogether surprising.

Thirdly, it is apparent that dealing with missing values other than the brackets (through multiple imputation) pushes up the average wage a little, particularly in the LFS surveys. This reduces the gap between the OHSs and the LFSs somewhat.

Fourthly in the QLFS part of the series there is hardly any imputation necessary (or possible), given that Statistics South Africa performed the imputation. Consequently the two “imputed” series are basically identical.

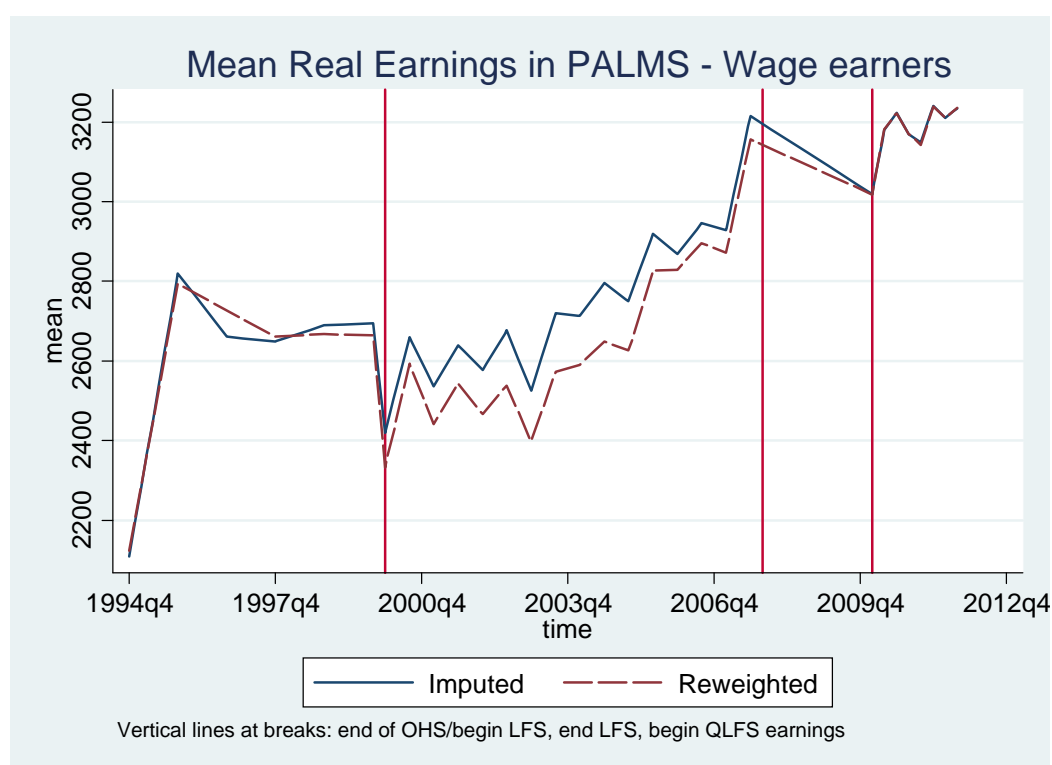


Figure 5. Earnings of wage workers over time

3.1.4 Real earnings of the Self-Employed

The detailed pattern for self-employed earners is given in Figure 6. The “collapse” in earnings between the OHSs and the LFSs is much more dramatic in this case. Given the fact that this is concentrated in the five months between October 1999 and February 2000 it is simply not plausible that this is due to “crowding” of the unemployed in the informal sector. It is a result of much higher

enumeration of informal activities, particularly of a marginal (and thus low-paid) kind in the Labour Force Surveys.

It is unclear whether the higher levels of self-employment earnings in the QLFSs is simply due to strong earnings growth between the end of 2007 (the last LFS figure) and early 2010 (the first QLFS data point), given that there is no information for the intervening period. The anomalous spike in the second quarter of 2011 is produced by a cluster of “millionaires” that wasn’t removed by the outlier detection routine, i.e. these did not look anomalous given their characteristics. Of course the threshold for flagging strange data was set at quite a high level. It is also possible that some of the high values were “hot decked”. This would, of course, make it harder for the outlier detection algorithm to isolate them, since the routine assumes that each data point provides independent information. The volatility of this series is one reason why Statistics South Africa only publishes medians on the QLFS earnings series. As we will show later, however, this practice obscures some important changes over time.

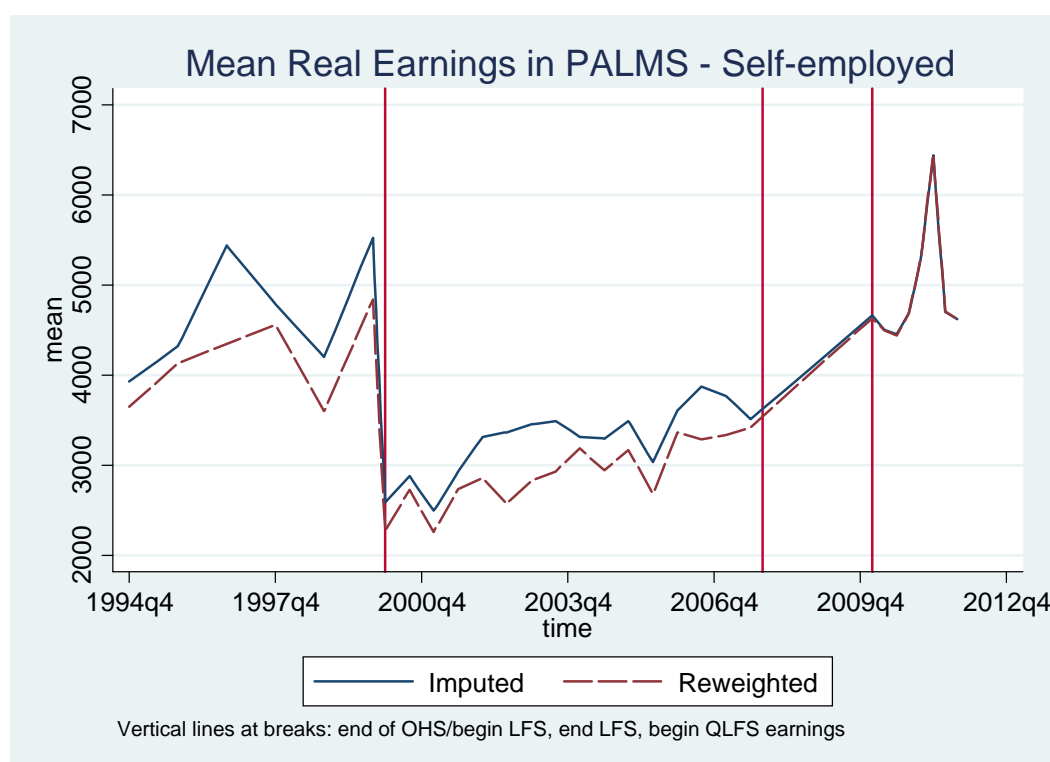


Figure 6. Earnings of Employers/Self-employed

3.2 Average earnings by major groups and sectors

3.2.1 Real earnings by Gender

The real earnings series for men and women are given in Figure 7 and Figure 8. Two points are noteworthy. Firstly, we cannot reproduce Heap’s findings that there was no gender gap in the 1995 October Household Survey wage series. The reason for this difference is probably due to different data cleaning protocols. The 1995 and 1994 data are particularly tricky because there were in effect different brackets for different reporting periods. Furthermore the earnings from self-employment are gross and the costs have to be deducted out. Reassuringly Casale (2004) also finds a sizable gender gap in 1995. In order to check the robustness of our results to different bracket imputation decisions we also show the bracket mean imputations in Figures 7 and 8.

Instead of a missing gender gap, we find a strong gender gap in **every** year in the sequence. Indeed the gap among the self-employed in 1994 and 1995 is extraordinary to say the least. What might explain this anomalous behaviour? Although there are some differences in the way the earnings question was asked it is clear that the difference didn't elicit different responses across the board – the early surveys seem to have found more high earning self-employed **males** or perhaps fewer low earning ones.

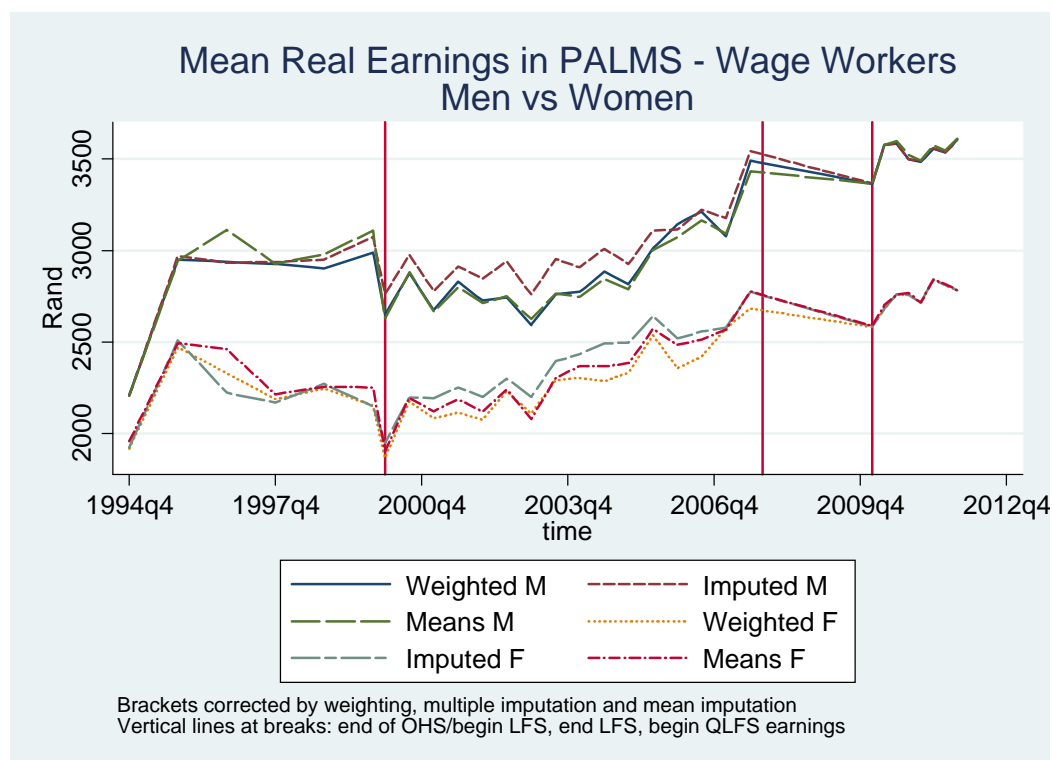


Figure 7. Real Earnings of Wage Workers by Gender

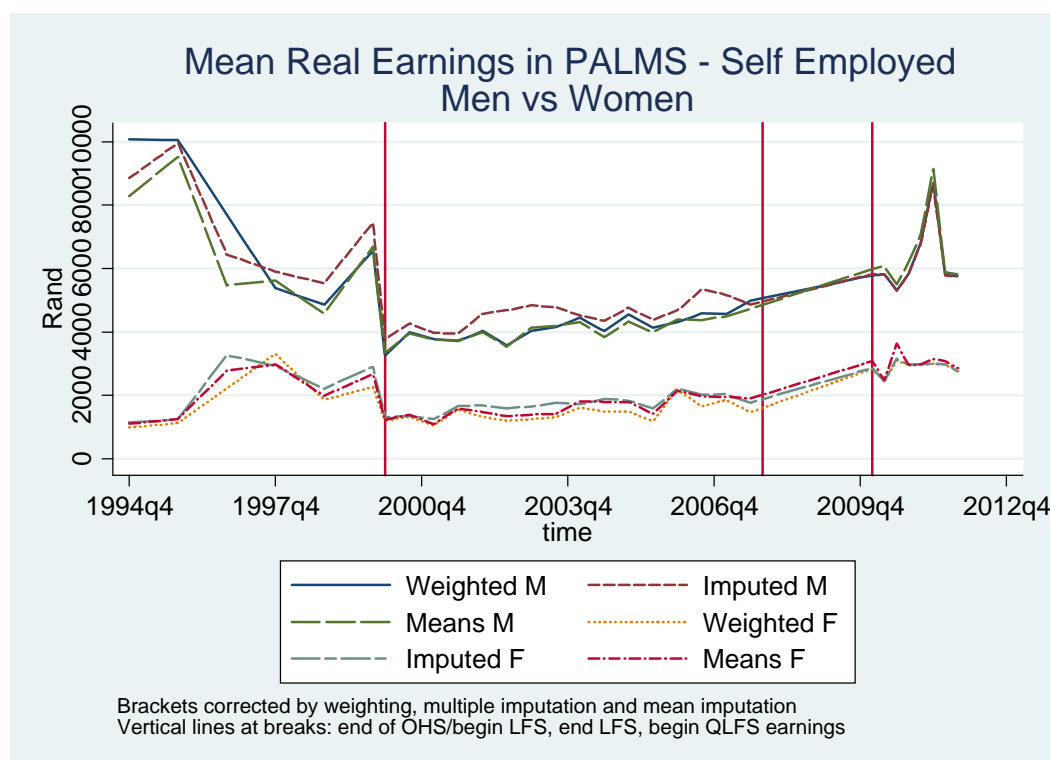


Figure 8. Real Earnings among the Self-Employed by Gender

3.2.2 Real earnings by Race

In 1994 South Africa was characterised by sharp racial disparities. As Figure 9 and Figure 10 show this hadn't changed all that much by the end of the period. The "racial order" of earnings – with Whites at the top, then Indians, Coloureds and Africans at the bottom was still preserved. Of course these averages conceal individual movement, but these movements have not been strong enough to erode the between group differences materially. The series for the "Indians" seems to show some catch-up with the Whites, but the series is also quite volatile, due to the rather small subsample on which these estimates are based.

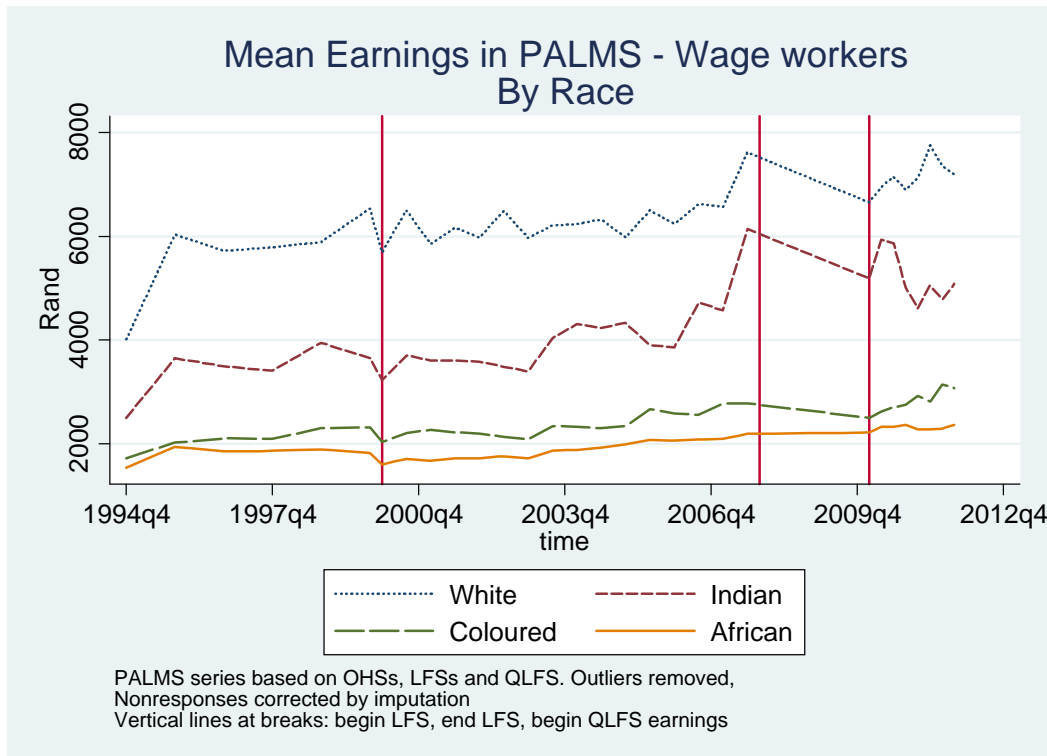


Figure 9. Real earnings of Wage Workers by Race

Looking at the real earnings series for the self-employed (Figure 10) it is clear that some of the strange patterns of self-employment earnings in the October Household Surveys are due mainly to a volatile pattern among Whites. It seems that there are more high earning White self-employed people in the 1994 and 1995 samples than is the case in later years. Cross-checking with the previous section, these are likely to be disproportionately men. It is not clear what might have produced this particular pattern. It is also clear that the cluster of "millionaires" in the second quarter of 2011 is largely White.

Interestingly some shift downwards in the distribution between October 1999 and February 2000 can be seen in every group. It suggests that the LFS did not only pick up more marginal work among the African majority, but also found more low-paying forms of self-employment in the other groups.

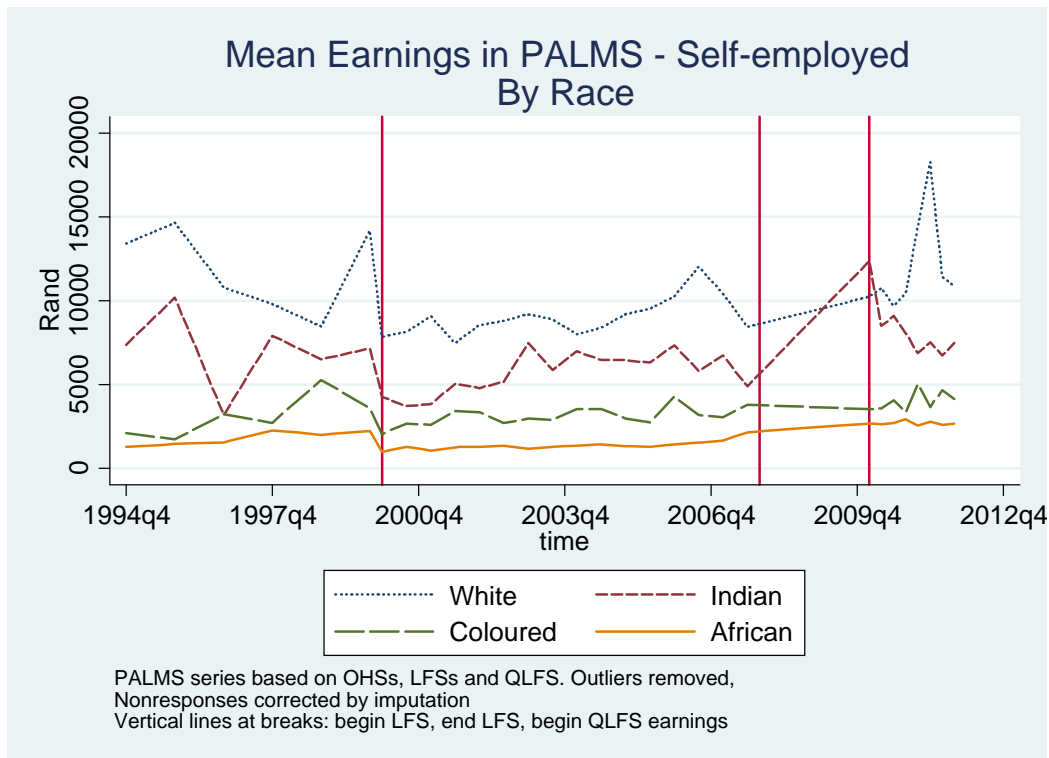


Figure 10. Real Earnings of the Self-Employed by Race

3.2.3 Real earnings by sector

The sectoral real earnings figures for wage workers are given in Figure 11 and Figure 12. No time series was calculated for “Utilities”, since the subsample size is too small in each survey. The overall picture is mixed. Real earnings seem to have risen in agriculture, mining, trade and services, while the patterns in the other sectors seem complicated. The wages for domestics look odd in the October Household Surveys. The very high mean wage recorded in 1994 is due to a very small sub-sample. The values for the rest of the OHSs are also likely to be badly measured, because of the coverage problems referred to earlier. Live-in domestics in particular are likely to have been underenumerated. The drop in construction wages between the OHSs and the LFSs may be due to better enumeration of casual workers.

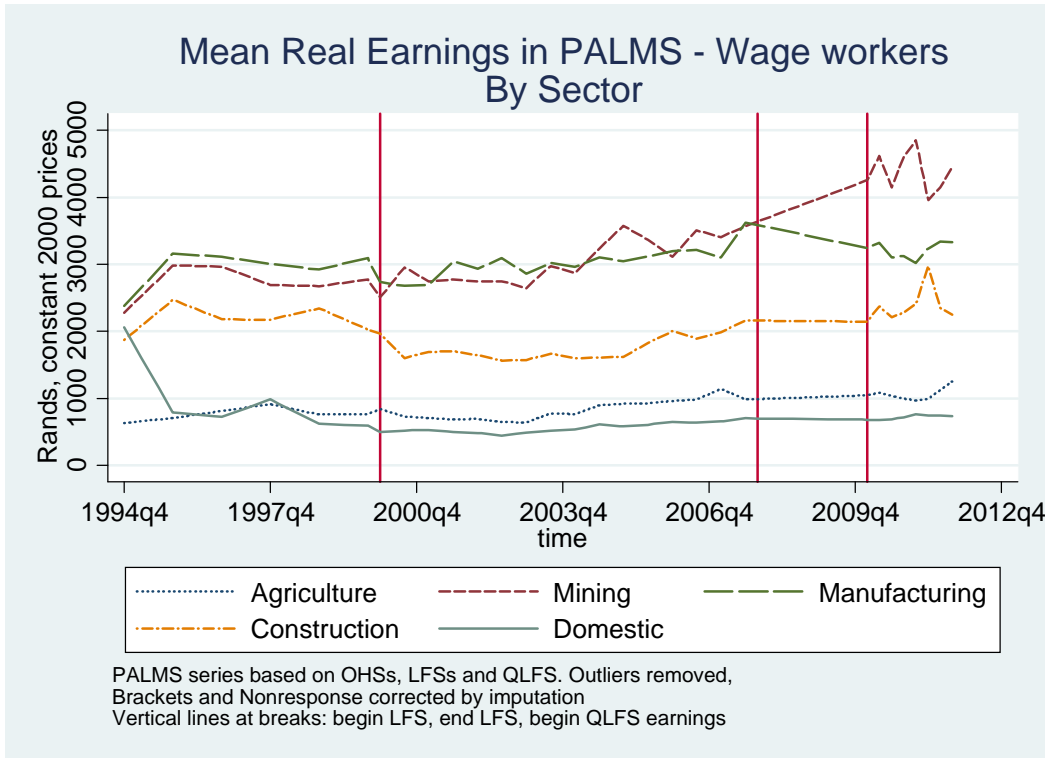


Figure 11. Real Earnings of Wage Workers by Sector - Part 1

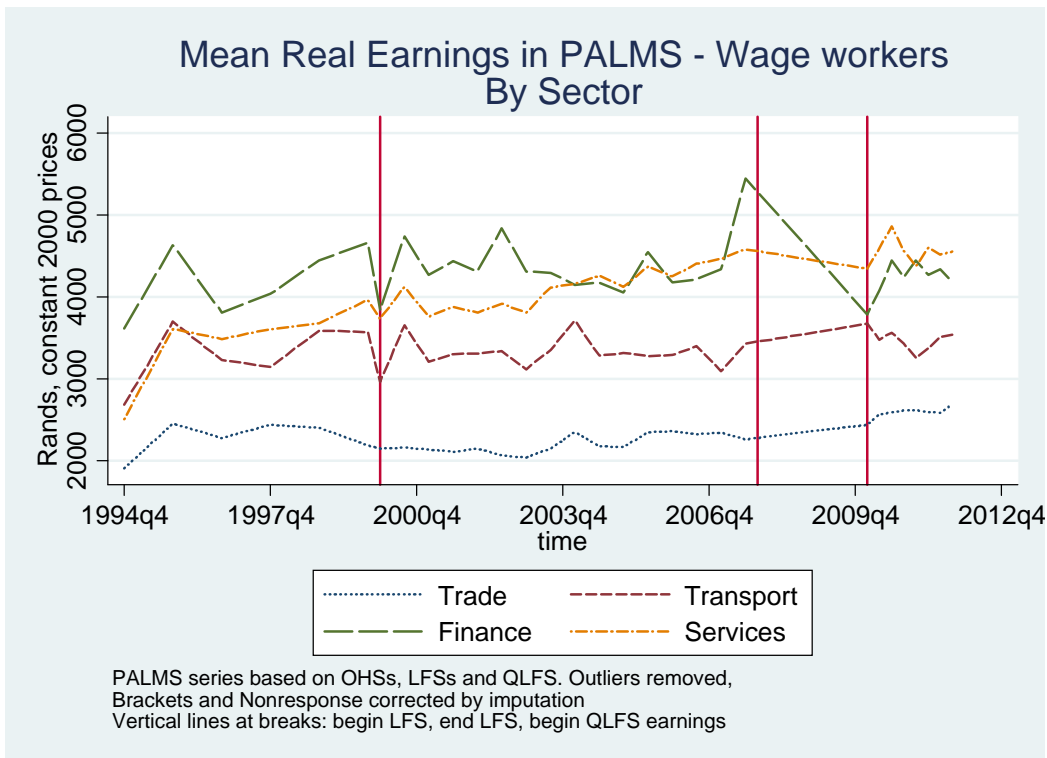


Figure 12. Real Earnings of Wage Workers by Sector - Part 2

The corresponding set of series for the self-employed is affected by small subsample sizes in several sectors. The pattern for sectors where there were at least a hundred cases in most years is shown in Figure 13 and Figure 14. The first point to note is that the agriculture series yields very high values in 1994 and 1995. Wittenberg (2004) noted that the OHS 1995 seems to find too many agricultural workers relative to the 1996 census and subsequent surveys. This suggests that the **formal** farm sector may be overrepresented in these early surveys. The relative overrepresentation of White farmers might explain some of the other odd patterns in these years. At one level it would not be surprising if these early surveys got the geographical distribution of the South African population wrong, given that the sampling had to occur off inadequate sampling frames.

The volatility of the financial sector earnings are not all that surprising, given that the distribution tends to be stretched – so one or two very high earnings individuals are likely to have a disproportionate influence on the series as a whole. Interestingly 1994 and 1995 again turn out to look on the high side for earnings in this sector.

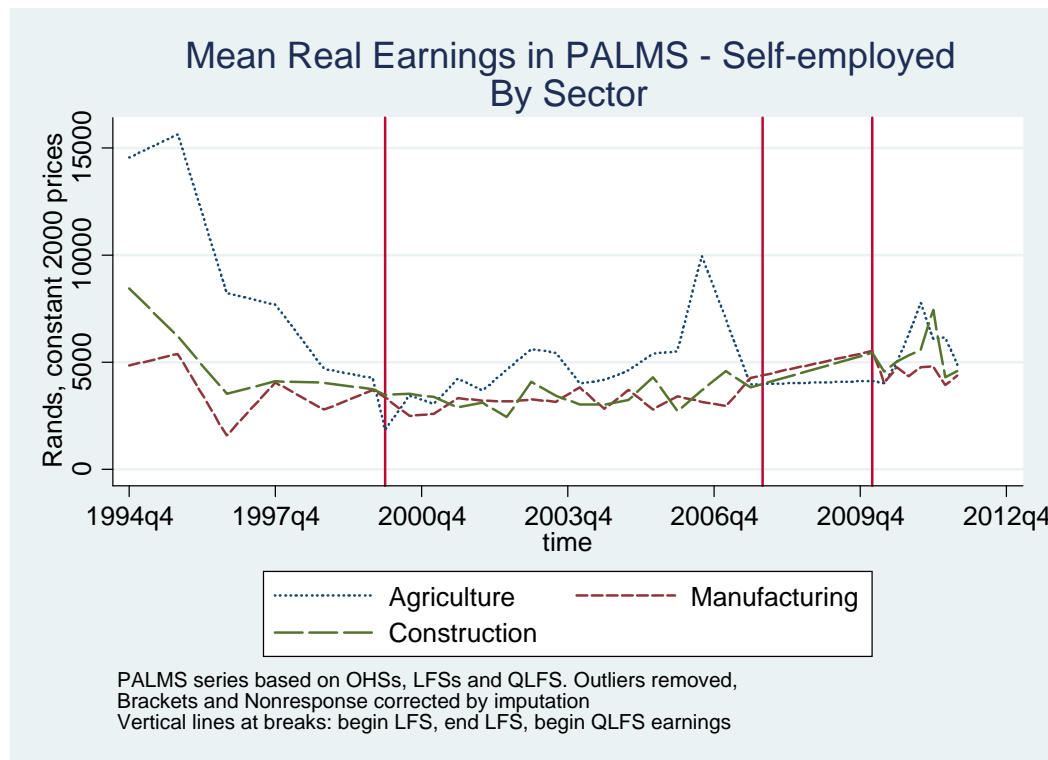


Figure 13. Real Earnings of the Self-employed by Sector - Part 1

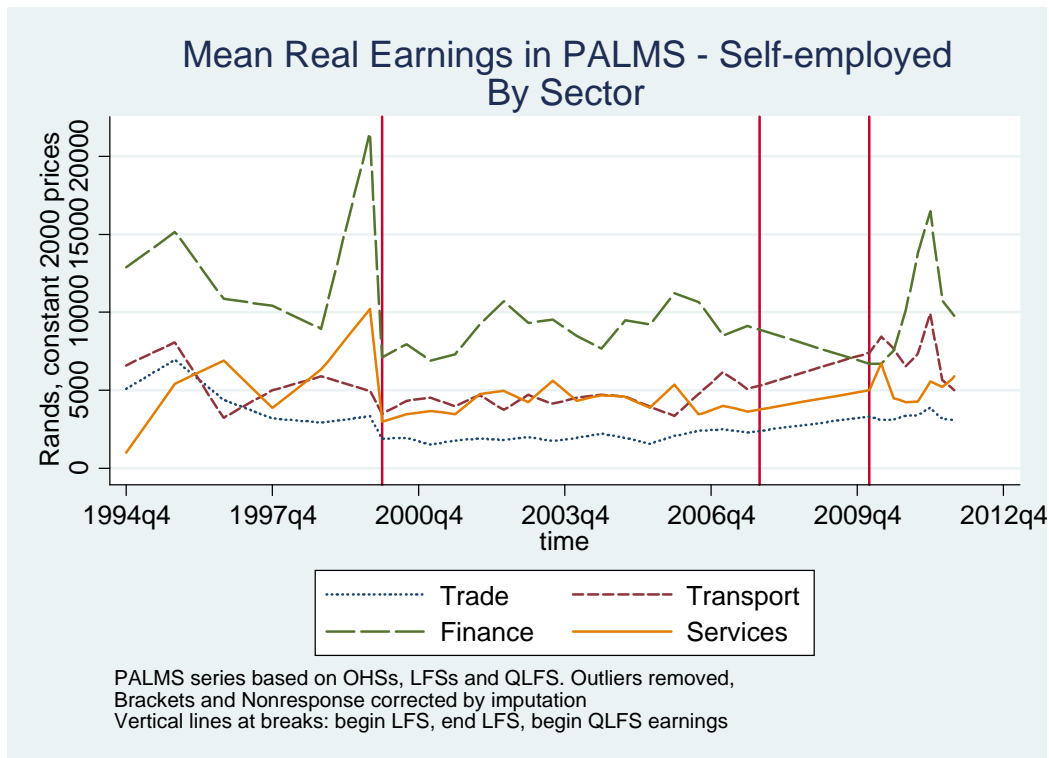


Figure 14. Real Earnings of the Self-employed by Sector - Part 2

3.3 The earnings distribution

In this discussion we focus on the changes of real earnings at the 10th, 25th, 50th, 75th and 90th percentile of the earnings distribution. We then calculate the ratio of the 10th, 25th, 75th and 90th percentiles to the median, to get a sense of changes in inequality in the distribution. We also calculate the ratio of the median to the mean.

3.3.1 Real earnings by selected percentiles of wage workers

The evolution of earnings at different points in the distribution of wage workers is shown in Figure 15. It is immediately apparent that there has been strong real earnings growth at the 90th percentile and, indeed, at the 75th. Median real earnings, by contrast seem to have been fairly stagnant. The picture at lower points of the distribution is not very clear from Figure 15.

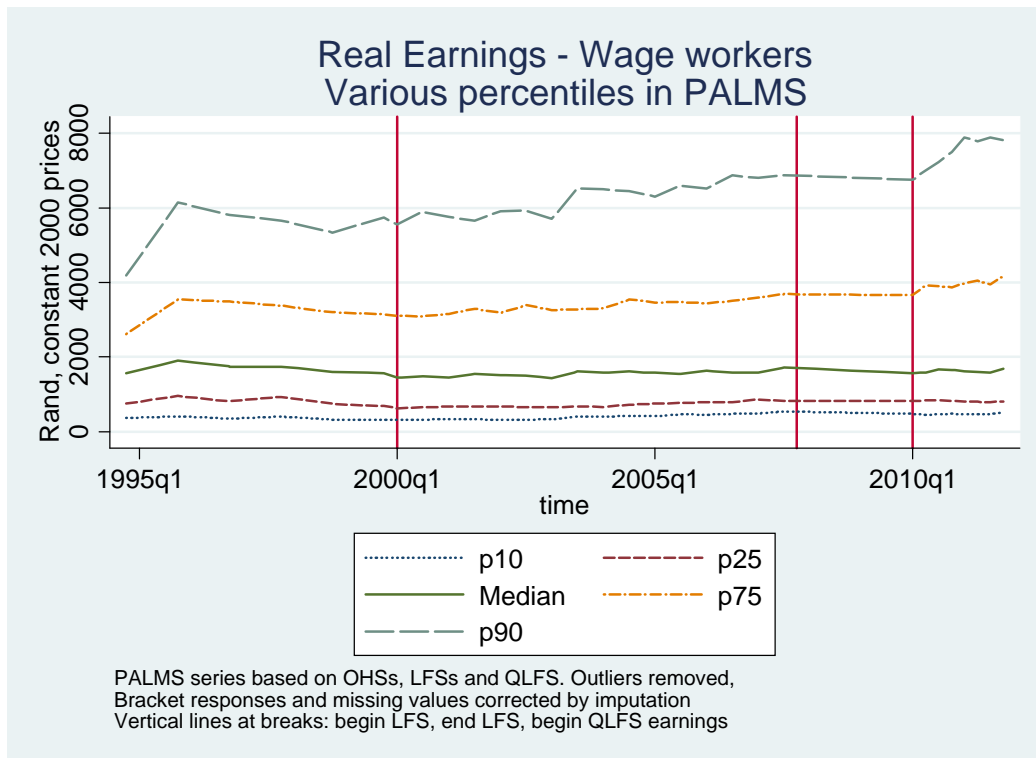


Figure 15. Real Earnings of Wage workers over time at different percentiles of the distribution

The relative position of the different percentiles in relation to the median is shown in Figure 16. This picture confirms the fact that the 90th percentile has moved away from the median. In fact the rate of the increase is greater than at the 75th percentile. This suggests that the top tail of the earnings distribution is “fanning out”. The end result of these changes is that the gap between the mean and the median of the earnings distribution has increased (see Figure 17).

The picture at the bottom of the distribution (right panel of Figure 16) is more complicated. The 25th percentile seems to have remained at around half of the median. The 10th percentile, however, has shown real earnings growth.

The overall pattern therefore seems to be a compression of the distribution at the bottom and a stretching of the distribution at the top. What mechanisms underpin these patterns is beyond the scope of this paper. It does, however, suggest that the practice of releasing mainly median figures for the QLFS may result in at best a partial picture of the evolution of earnings. Of course a focus on the mean only would be equally limiting.

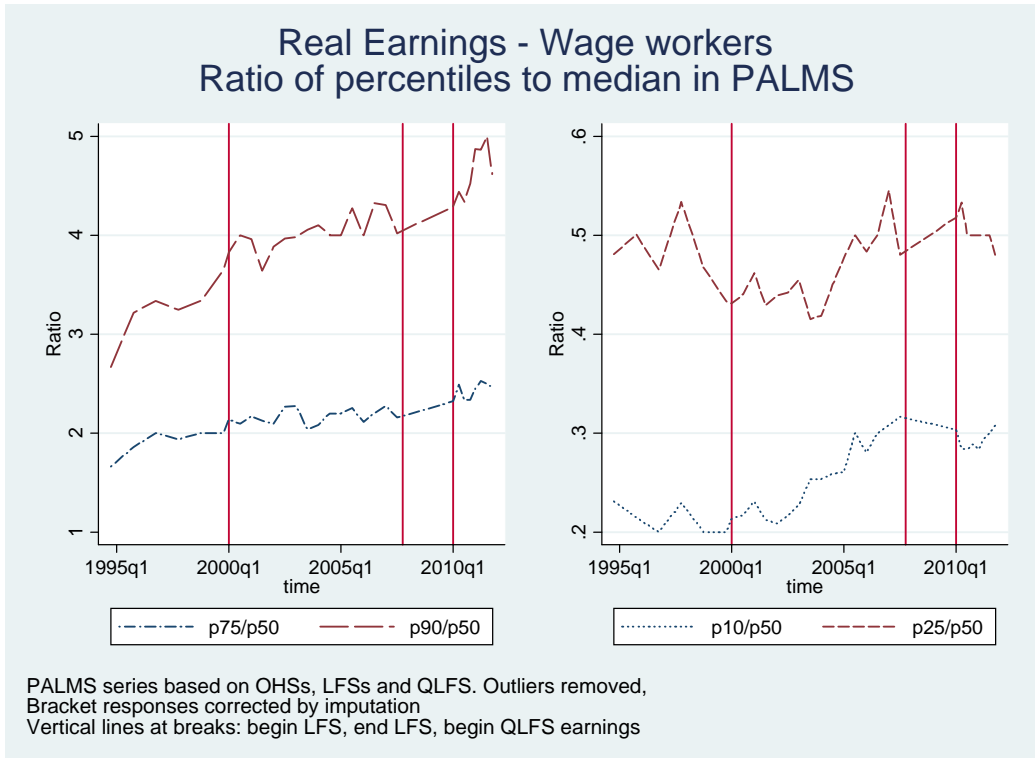


Figure 16. Change in position of different percentiles relative to the median

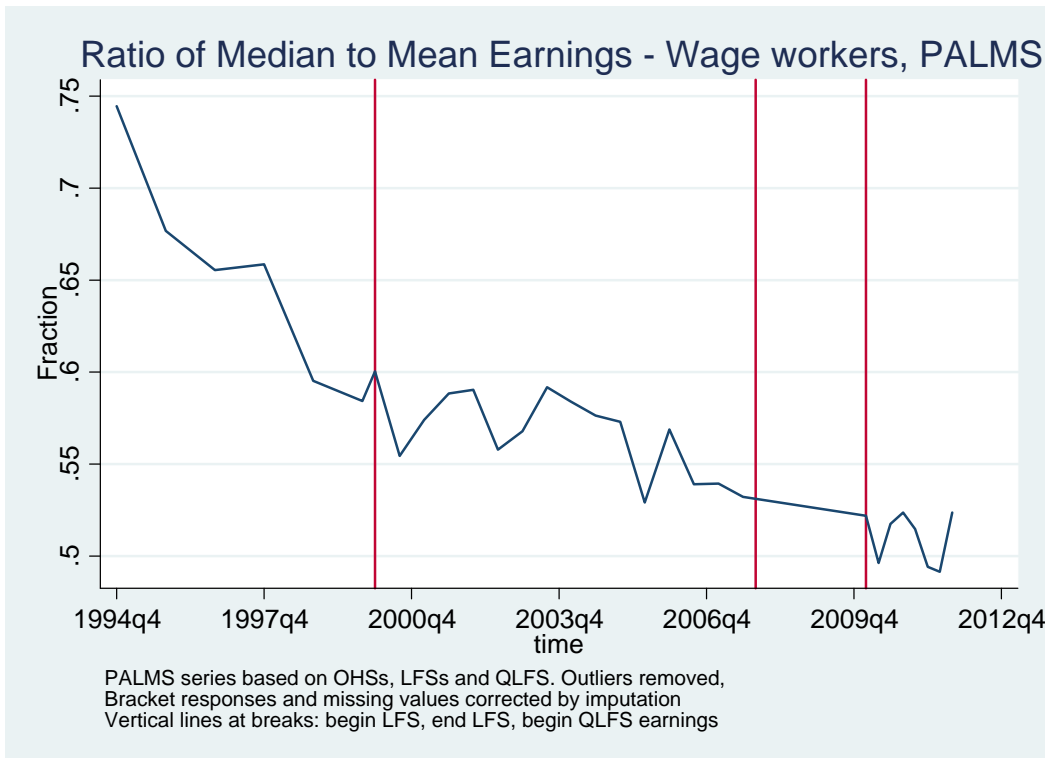


Figure 17. The gap between median and mean earnings among wage workers seems to have widened over time.

3.3.2 Real earnings by selected percentiles of the Self-employed

The picture for the self-employed is considerably different. Figure 18 shows the real earnings at various positions of the distribution. Earnings at the 90th percentile show steady growth from the LFS surveys onwards. Earnings at the 90th percentile show steady growth from the LFS surveys onwards. Over the period of the OHSs, however, the median, 75th percentile and 90th percentile all seem to grow strongly only to collapse down in the LFSs. The simplest explanation for this pattern is that the absolute increase in the number of self-employed (particularly in the bottom tail) measured in the surveys over this break led to a mechanical adjustment downwards in these thresholds.

Looking at the relative positions of these percentiles (shown in Figure 19), there is no analogous process of stretching of the distribution at the top. Ignoring the first two OHSs, the ratio of the 90th percentile to the median seems reasonably steady over the OHSs and LFSs and then shows some sign of dropping. The effect of this (as shown in Figure 20) is that the gap between the median and mean has shrunk somewhat. There is also no sign that the bottom of the distribution has become more compressed. Comparing the scales of Figure 19 and Figure 20 to those in Figure 16 and Figure 17 it is evident that the earnings distribution of the Self-employed is considerably more unequal than that of the wage distribution.

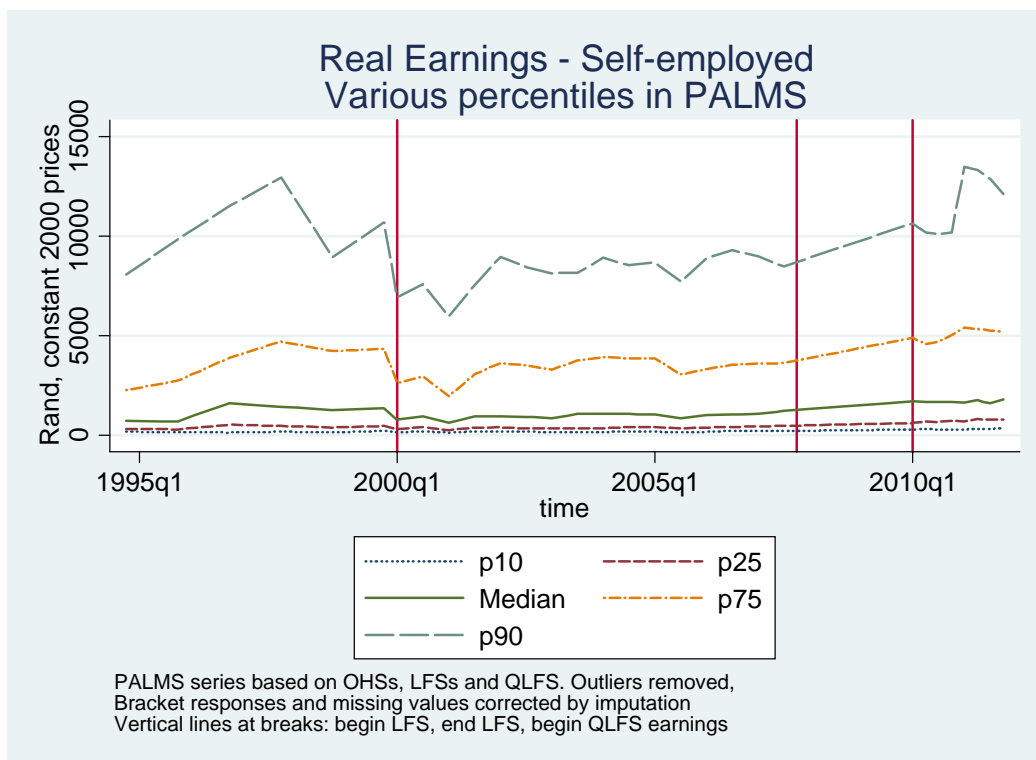


Figure 18. Real Earnings of the Self-employed at different positions of the earnings distribution over time

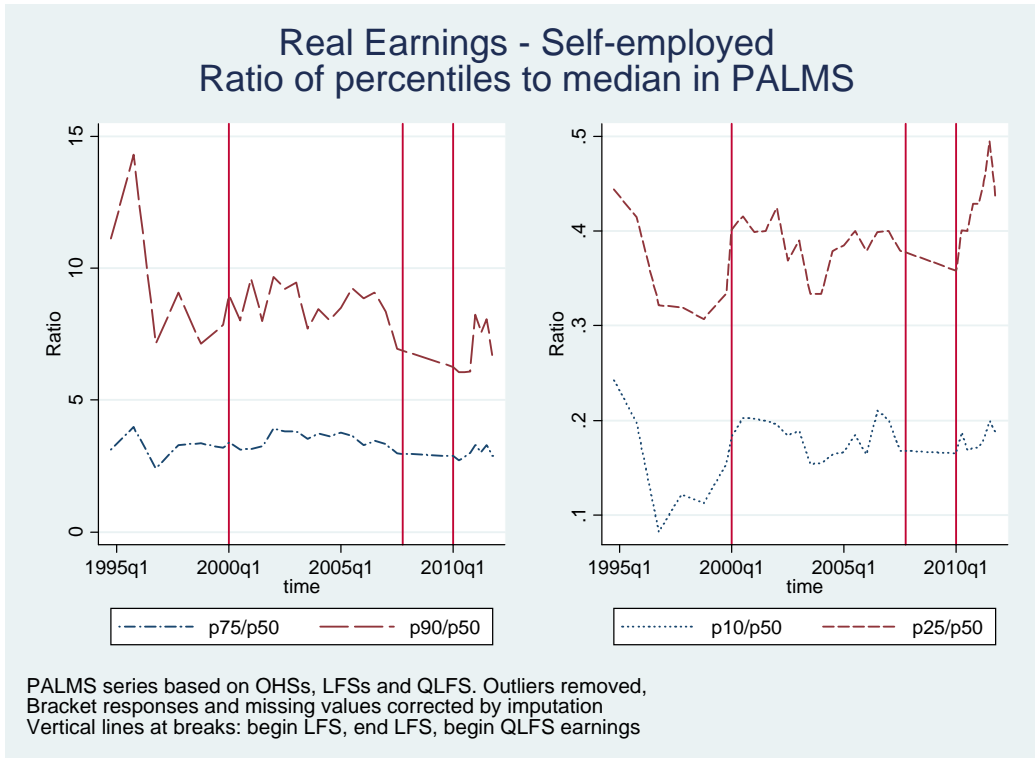


Figure 19. Change in position of different percentiles of the Self-employed earnings distribution relative to the median

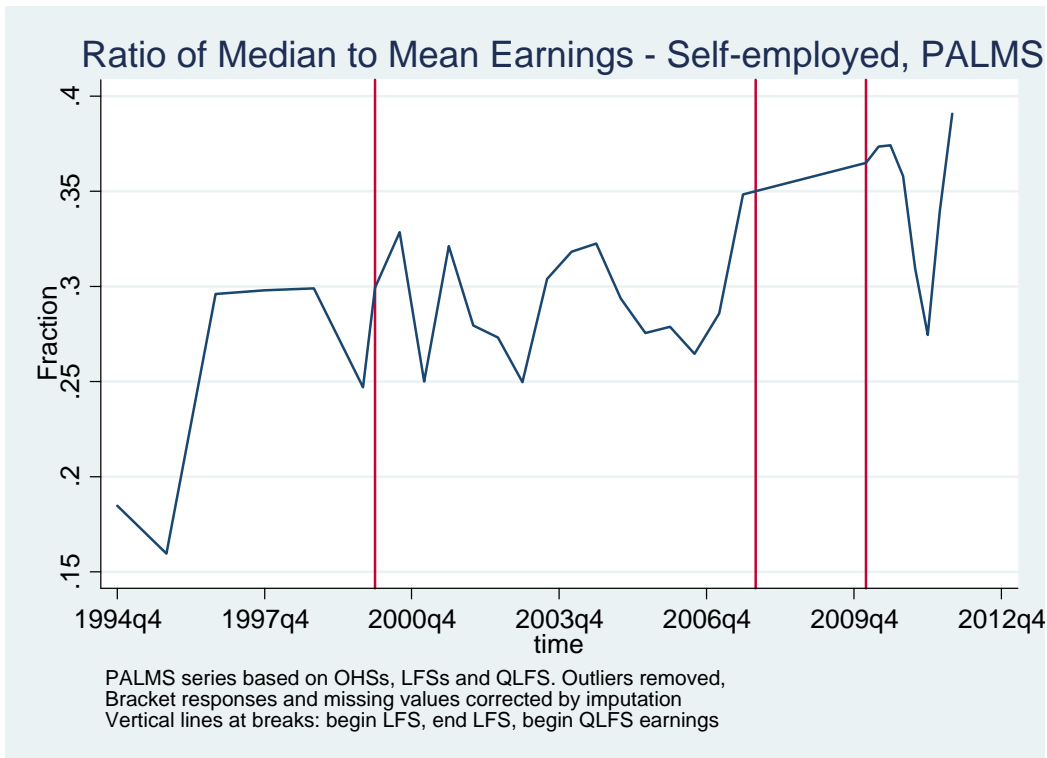


Figure 20. The ratio of median to mean earnings among the Self-employed

4. Employment

At various stages in the discussion so far we have referred to a big shift in the number of workers enumerated in the surveys between the last OHS (October 1999) and the first Labour Force Survey (February 2000). In this section we will analyse that issue further.

4.1 Total Employment

The picture is seen very clearly in Figure 21 where we graph total employment, total numbers of wage workers (left axis) and the total number of self-employed (right axis). The shift up is clearly visible in every series. It is also evident that 1995 showed higher levels of employment when compared to the surveys around it.

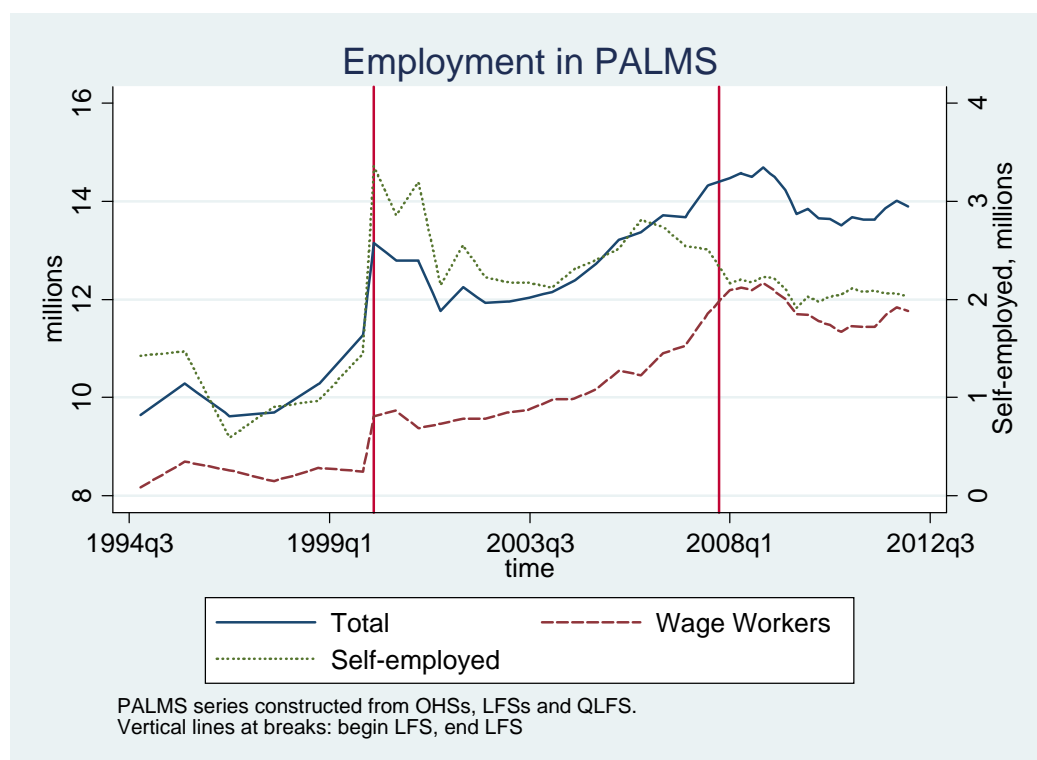


Figure 21. Total employment

The picture emerges even more starkly in Figure 22 where we graph the change of employment. In order to standardise for the different gaps between the surveys, we have annualised the changes. This accentuates the shift between the OHSs and LFSs even more, because it was concentrated in a half year period (October to February).

Ignoring the break in the series, there is clear evidence both in Figure 21 and Figure 22 of steady increases in employment over the entire period until the financial crisis of 2008/2009.

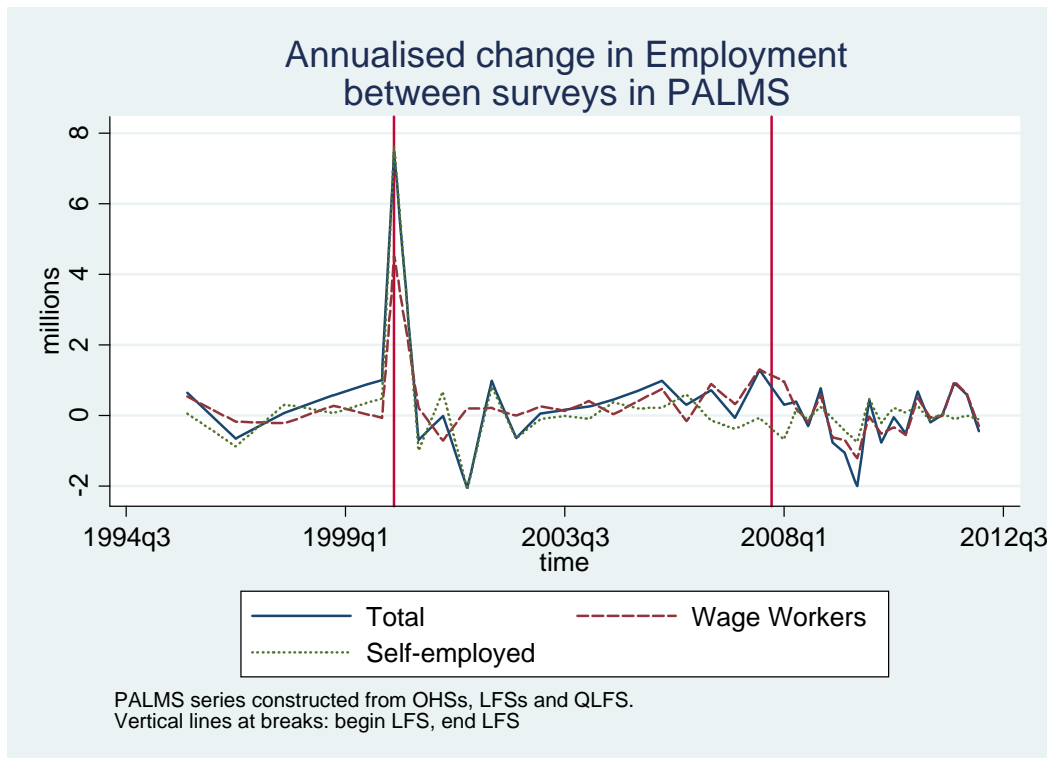


Figure 22. Annualised change in employment - the big shift in February 2000 is clearly evident

4.2 Employment by Sector – Wage Workers

The sectoral figures provide more information. In Figure 23 and Figure 24 we provide the employment trend by major industry code for wage workers. It is evident that OHS94 and OHS95 found much more agricultural employment than any other survey since. Wittenberg (2004) argued that the drop is unlikely to have been a real one. It is much more likely to have been an artefact of skewed sampling frames. Of course it must also be recognised that industry codes are likely to be recorded with some error, given that they are reconstructed from the answers supplied by the respondents.

Domestic service (shown in Figure 24) was obviously recorded very badly in the early surveys. For reasons outlined above it is likely that domestics were underenumerated in these surveys. Those domestic workers that **were** captured in the surveys were probably classified into the “service” sector. It is interesting to note in Figure 24 that several series (domestic work, trade and services) all show breaks at the OHS/LFS border. This is brought out more clearly in Figure 25, which records the corresponding annualised changes in employment. So although the major changes at the OHS/LFS break are concentrated among the self-employed (as we show below), it is worth noting that the LFSs were also more successful at capturing wage workers. Indeed that is also evident in Figure 21.

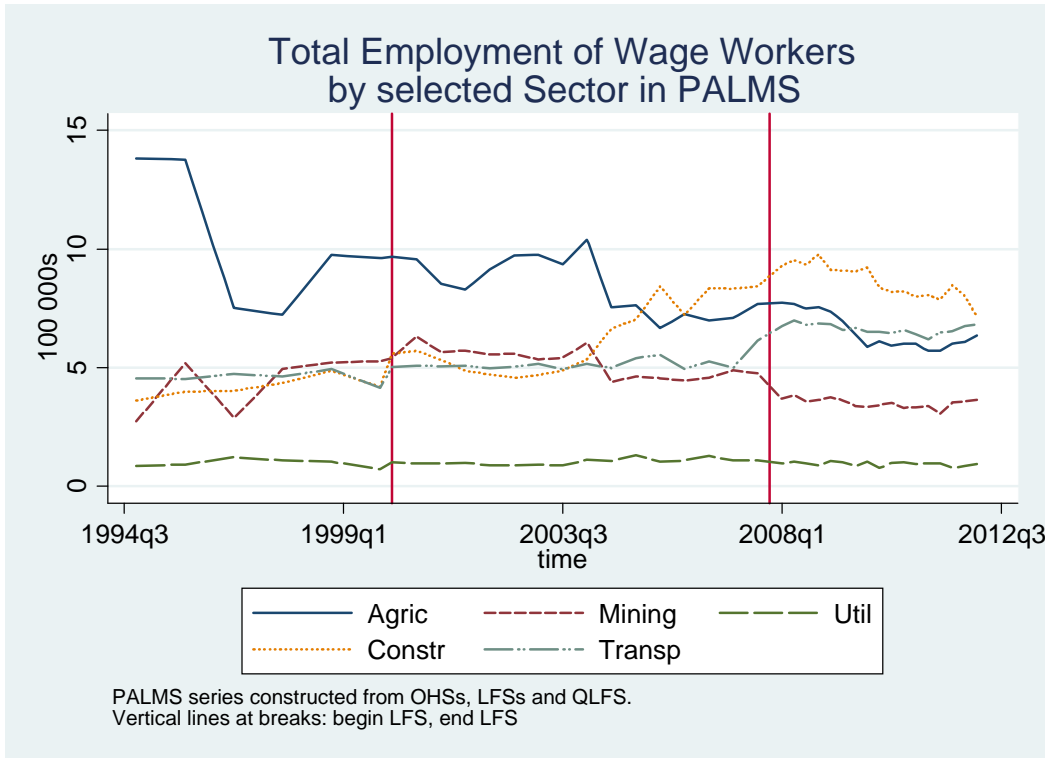


Figure 23. Employment among Wage Workers, by Sector - Part 1

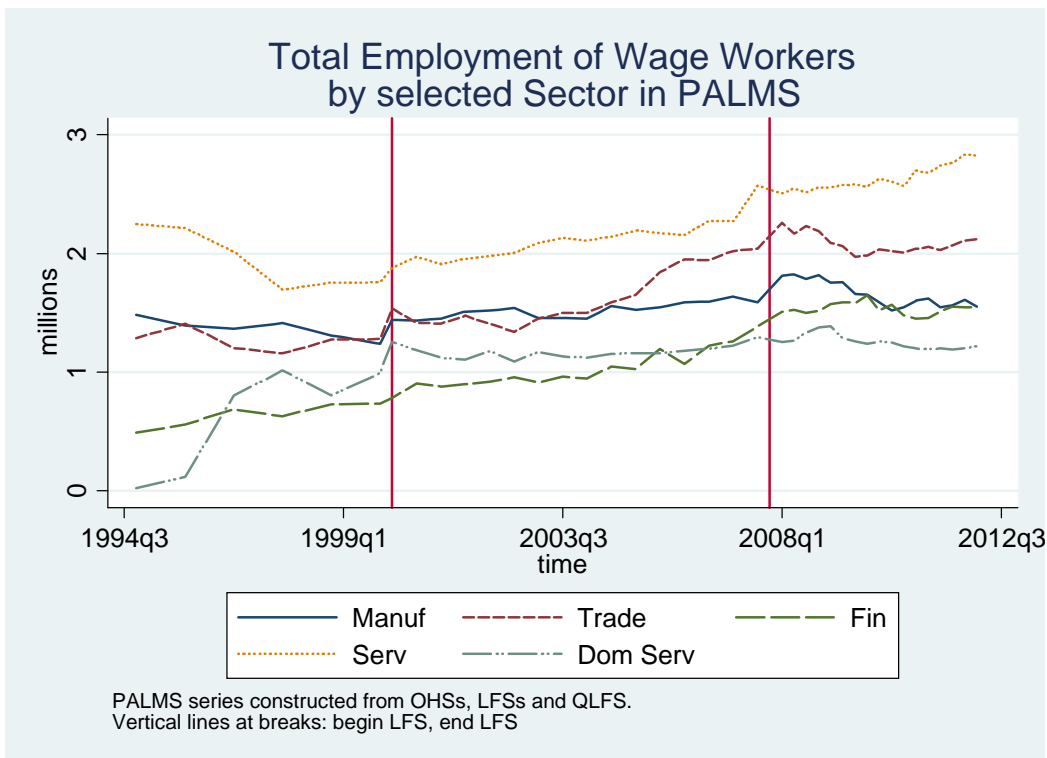


Figure 24. Employment among Wage Workers, by Sector - Part 2

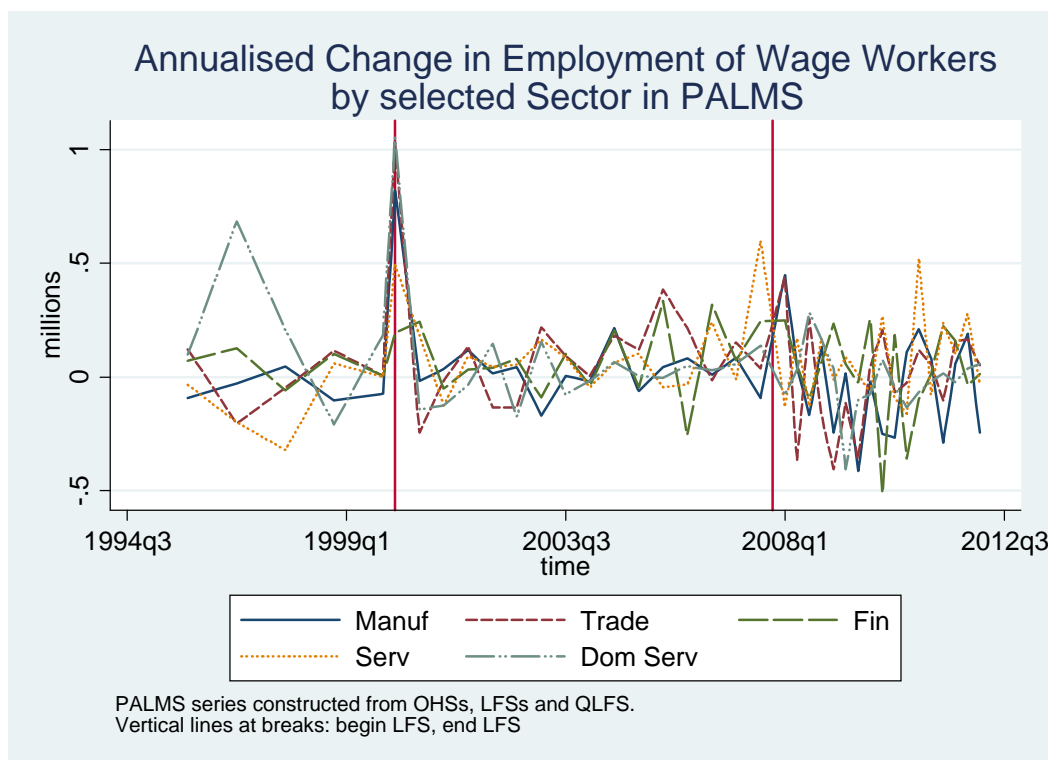


Figure 25. Changes of employment among Wage Workers for selected sectors

4.3 Employment by Sector – Self-employed

The picture for the self-employed is given in Figure 26 and Figure 27⁵. The most dramatic series is that for self-employment in agriculture (largely subsistence). Figure 26 shows that more than a million extra people were recorded in this sector between October 1999 and February 2000. Looking at the other sectors, it is evident that trade, manufacturing and transport also all show breaks at this point. In most of the other sectors, however, the increase in recorded self-employment levels starts in 1999, which agrees with the Kerr and Wittenberg (2013a) argument that a shift in sampling methodology led to better enumeration of workers living in back rooms. The break at the OHS/LFS boundary is, however, a different break – produced by a changing definition of what constitutes “work” and much more probing of informal sector activity.

⁵ We don't provide figures for the mining, utilities and domestic services sector, because the numbers of self-employed in these sectors are very small.

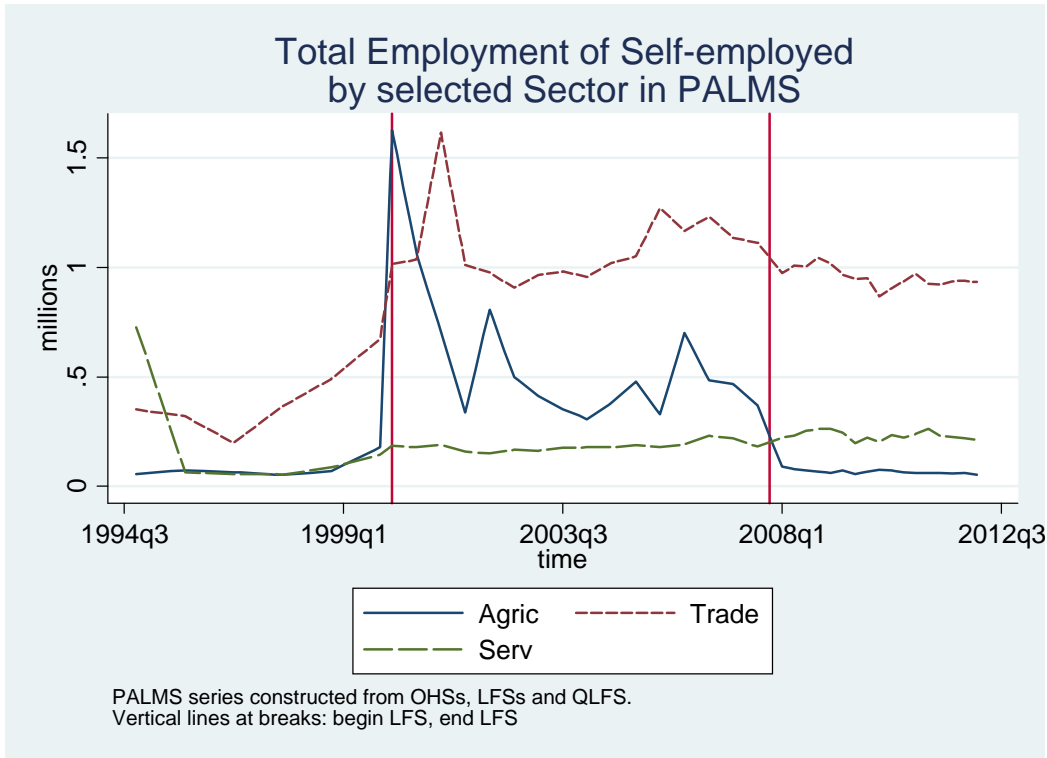


Figure 26. Numbers of self-employed by Sector - Part 1

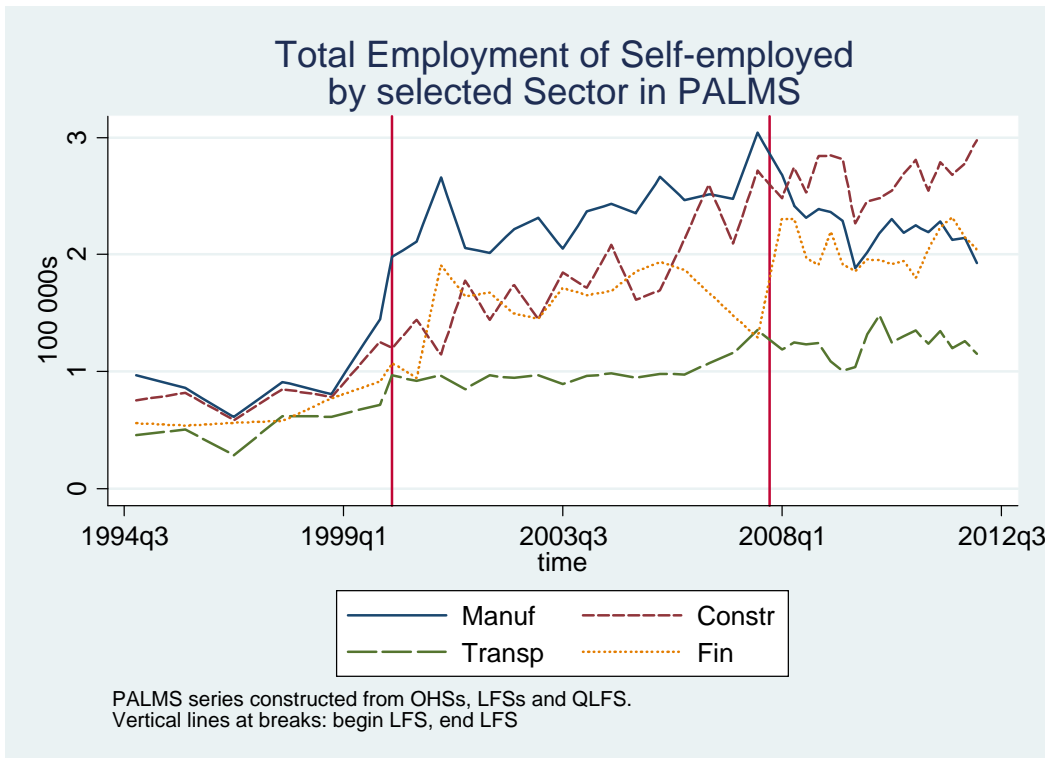


Figure 27. Numbers of self-employed by sector - Part 2

5. Working hours

The variable recording average hours worked in the last week is subject to some measurement error: the maximum recorded is 190, which is physically impossible. For the analysis below we set anything above 98 hours (14 hours per day, 7 days a week) to missing.

5.1 Total Hours worked

A different perspective on employment in the economy is provided by looking at the total hours worked in a week. One might imagine that changes in economic activity might not filter through to employment levels if individuals either work longer or shorter hours.

5.1.1 Total Hours

Figure 28 provides the aggregate picture, comparable to the total employment graph (given in Figure 21). The discontinuity at the start of the LFSs is still very evident, although it is not quite as pronounced as in the earlier figure. The subsistence agriculture workers that were more assiduously enumerated in the LFSs clearly work fewer hours than individuals in other sectors, a point we'll see below (Figure 34).

It is also worth noting the reduction in total hours (and indeed total employment) after the onset of the financial crisis in 2008. This is almost definitely a real change and not one induced by measurement changes.

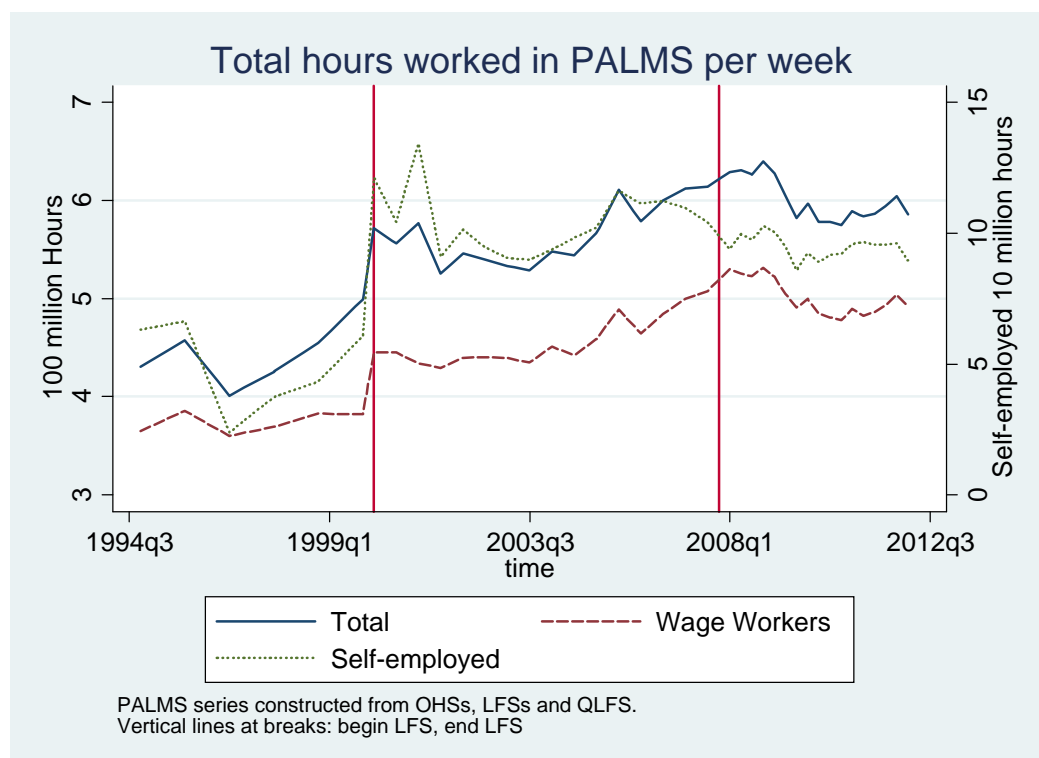


Figure 28. Another take on employment in South Africa - total hours worked in the economy, per week

5.1.2 Total Hours by Sector

The aggregate sectoral pictures are shown in Figure 29 and Figure 30 respectively. The long-term decline of activity in agriculture (except for the anomalous blip in 2000) is clearly in evidence, as is the long-run increase in construction, trade, services and financial services.

In summary, the “total hours worked” series provides information parallel to that given by the employment numbers, although distorted a little less by the higher levels of informal and subsistence activities recorded in the LFSs. We might therefore expect productivity numbers calculated per hour to be more robust than productivity per worker.

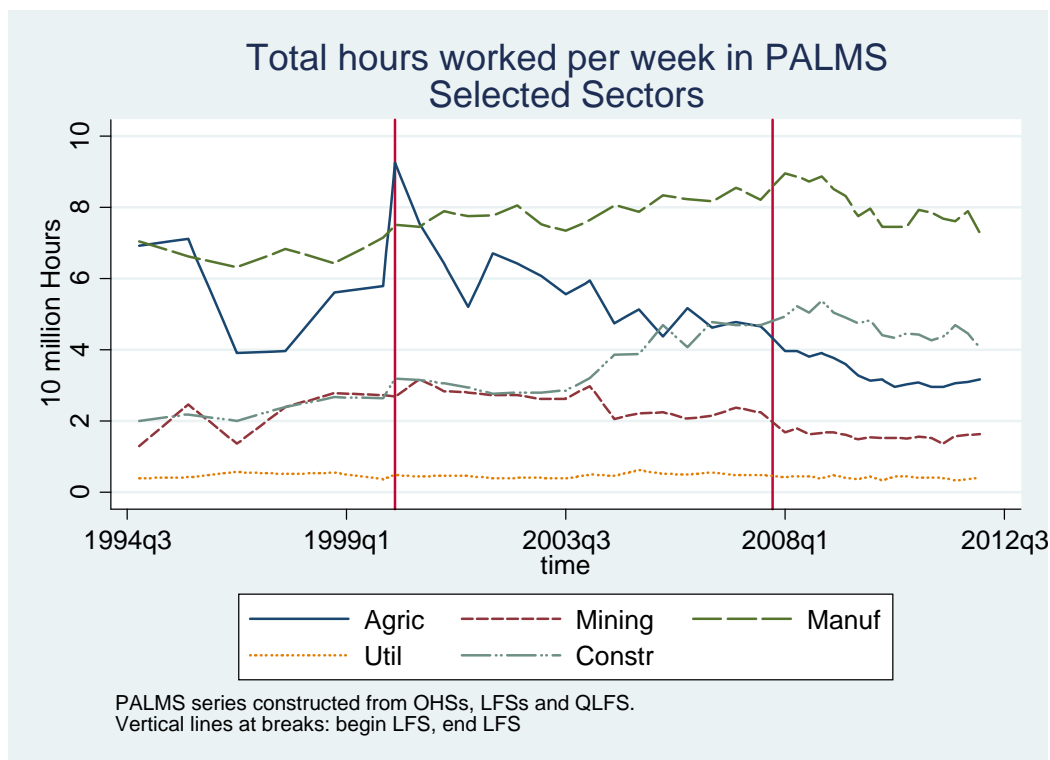


Figure 29. Total hours worked per week by sector - Part 1

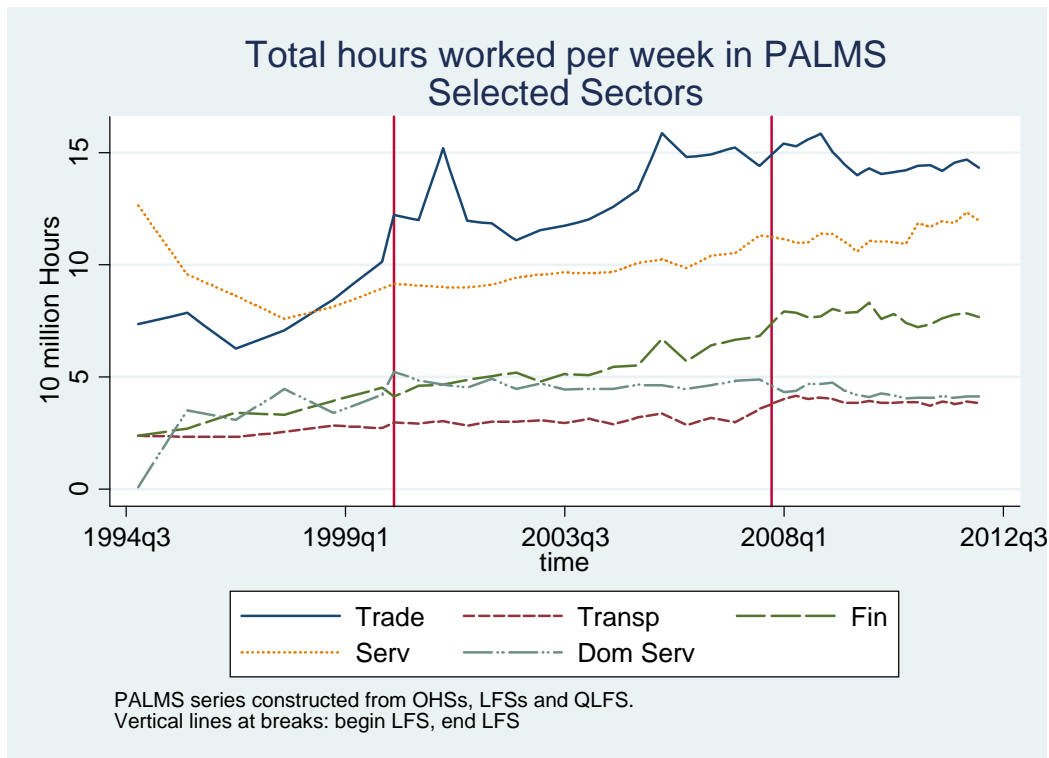


Figure 30. Total hours worked per week by sector - Part 2

5.2 Average hours: overall and by employment status

The aggregate pattern is shown in Figure 31. There are several noteworthy features. Firstly, the recorded hours worked increased over the OHSs. The coverage problems associated with these surveys therefore also seem to have resulted in the underenumeration of people working relatively long hours. Given who was likely to be missed (single persons living in backrooms) this is not altogether surprising. Secondly, there seems to be a long-term decrease in the hours worked by wage workers over the period of the LFSs and QLFs. The hours worked by the self-employed are markedly lower in the LFS period than before or after. This is a reflection of the greater enumeration of more marginal forms of economic activity.

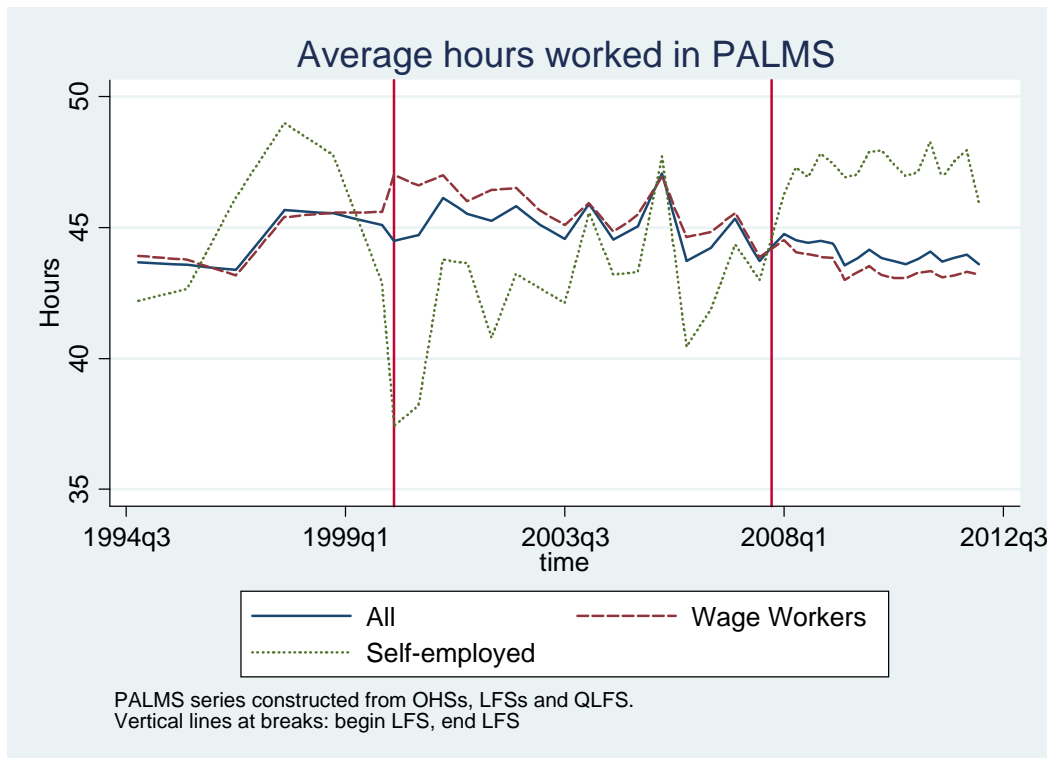


Figure 31. Hours worked - by employment status

5.2.1 Average Hours by Sector – Wage workers

The sectoral trends in hours worked are shown in Figure 32 and Figure 33. The trend towards a reduction in hours is noticeable across broad sectors of the economy. Interestingly retail (“trade”), transport and finance are exceptions to this pattern. Hours among domestic workers have come down much more sharply than elsewhere. This suggests a move towards more part-time forms of employment.

Comparing the sectors, agricultural, mining and transport workers work longer hours per week than workers in manufacturing and construction.

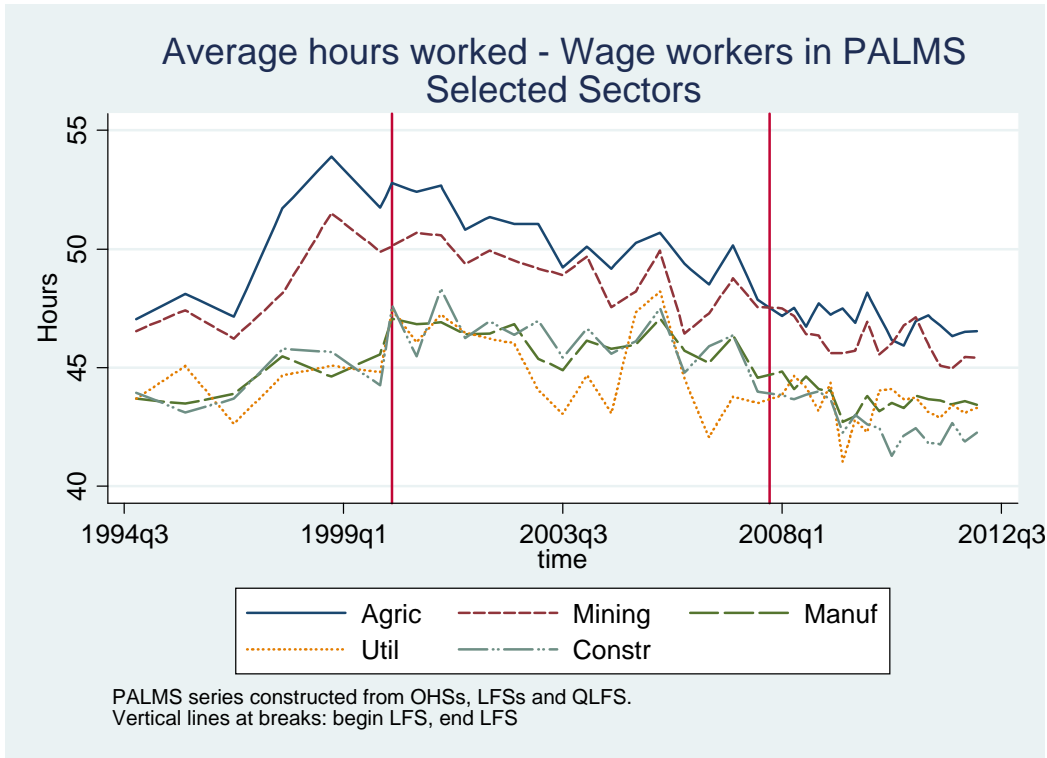


Figure 32. Average hours worked by wage workers, selected sectors - Part 1

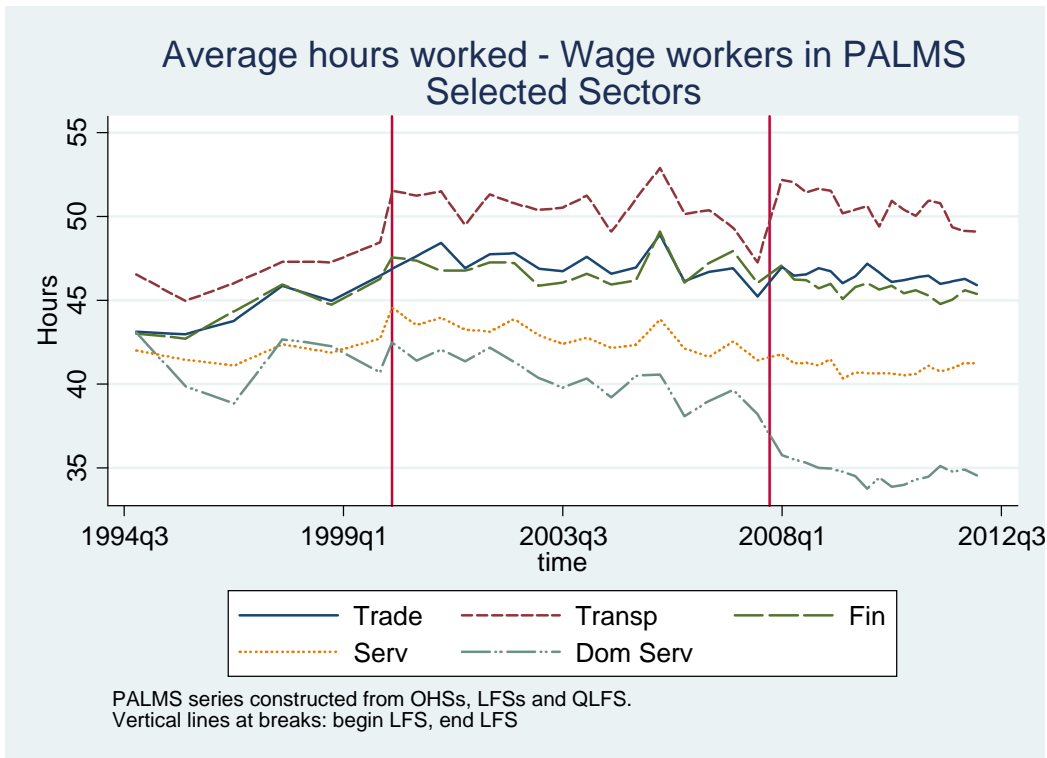


Figure 33. Average hours worked by wage workers, selected sectors - Part 2

5.2.2 Average Hours by Sector – the Self-employed

The pattern among the self-employed is shown in Figure 34 and Figure 35. The pattern in agriculture is remarkable. The LFS series for this sector is clearly very different than the series in the OHSs and QLFSs. None of the other sectors show any analogous breaks. In contrast to the wage workers, there is no long-run trend in reduction of hours. Indeed compared to wage workers the hours worked in many sectors seem to be on the long side.

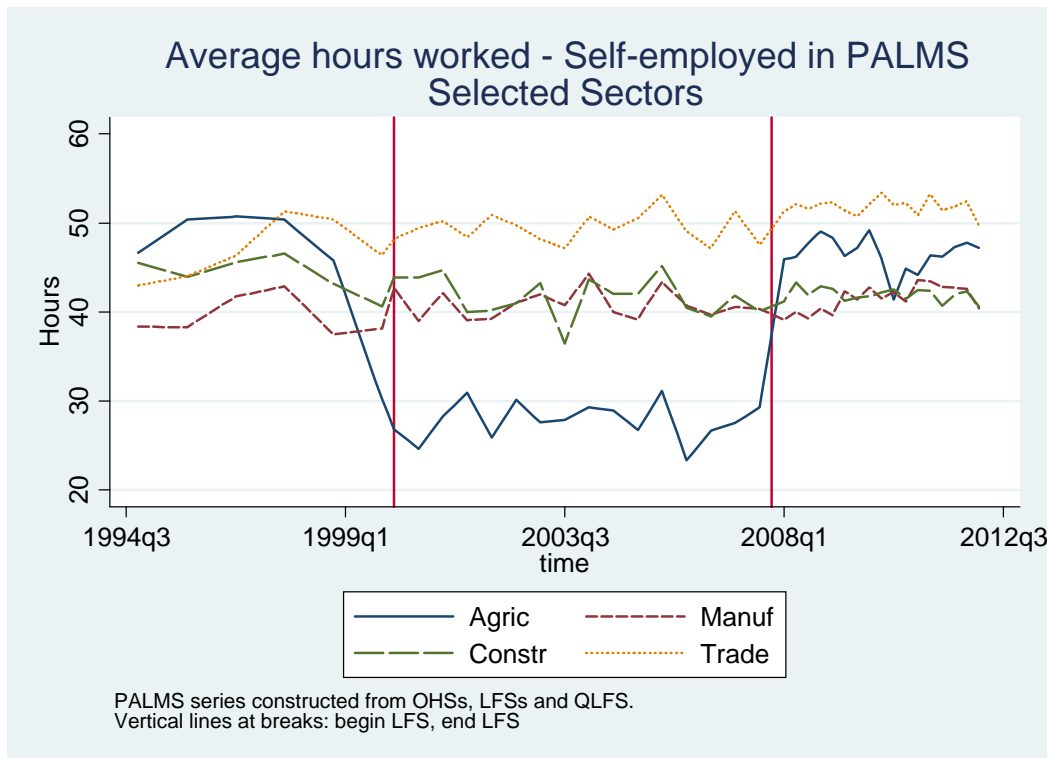


Figure 34. Average hours worked among the Self-employed, by Sectors - Part 1

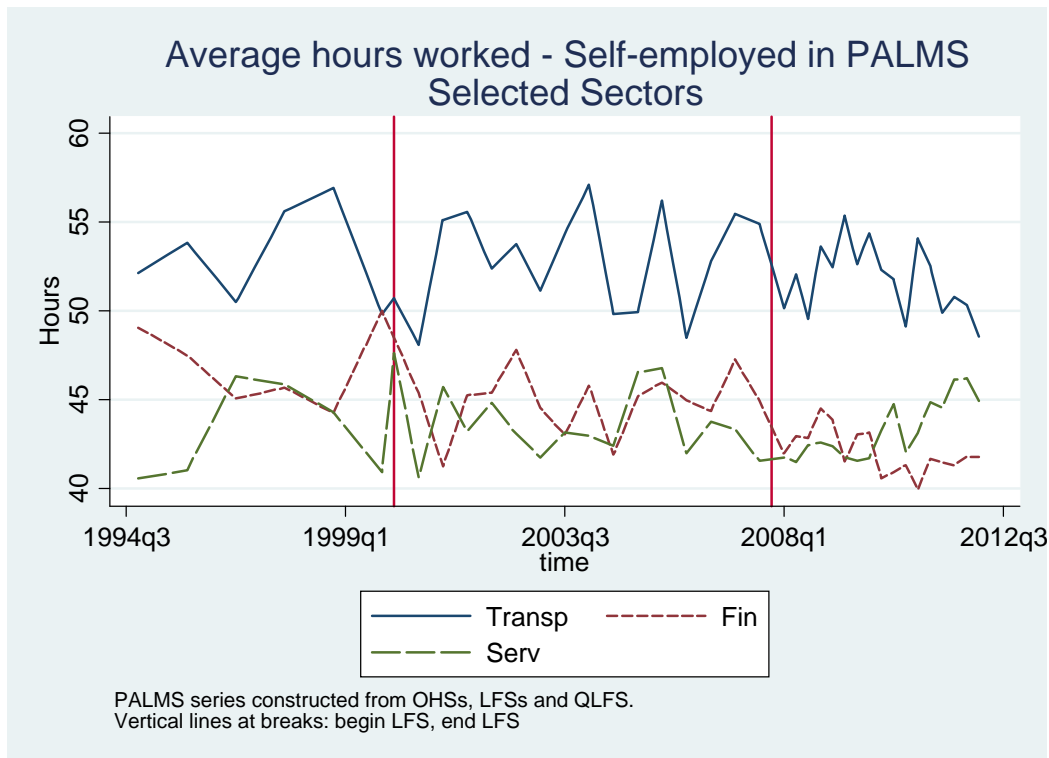


Figure 35. Average hours worked among the Self-employed, by Sector - Part 2

5.3 Earnings per hour

5.3.1 Overall trends

The aggregate trends for the period as a whole are shown in Figure 36. It is evident that there has been steady growth overall. Again the break between the OHSs and the LFSs is evident. It is also clear that growth has been stronger among the self-employed than among employees.

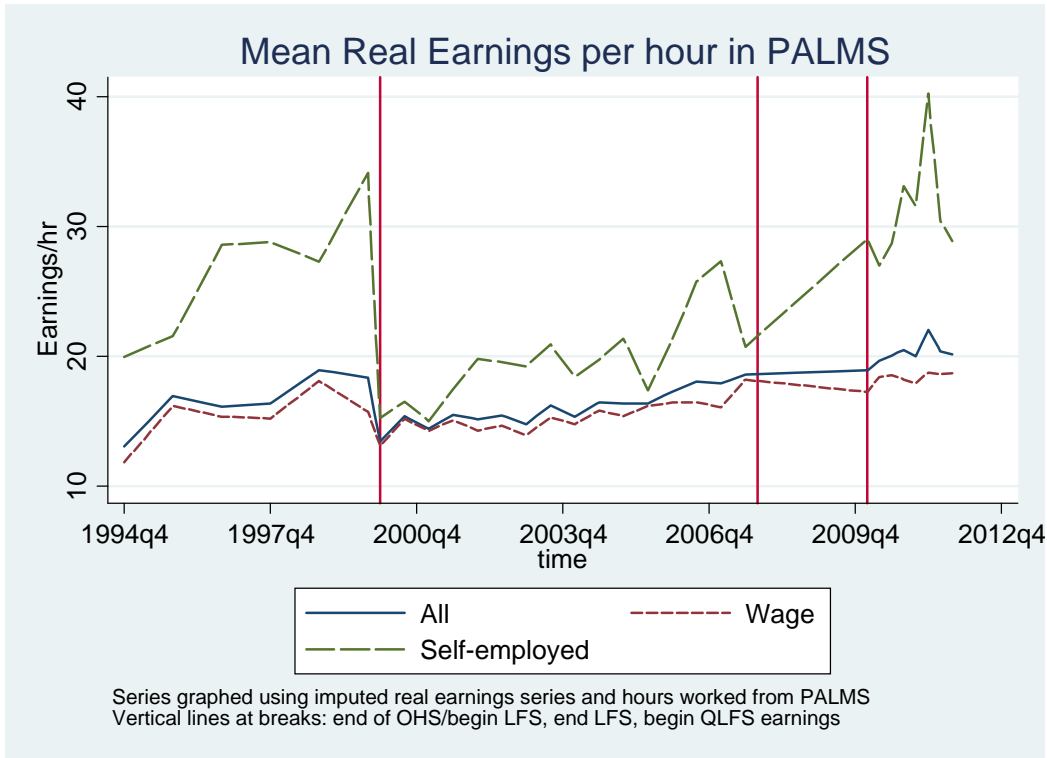


Figure 36. Real earnings per hour, 1994-2011

5.3.2 Earnings per hour among employees by Sector

Figure 37 and Figure 38 parallel the evidence given in Figure 11 and Figure 12. Compared to the earlier graph, agricultural workers now earn as little as domestics once their hours are taken into consideration. Service workers (including government employees) seem to be the highest earners since around 2003.

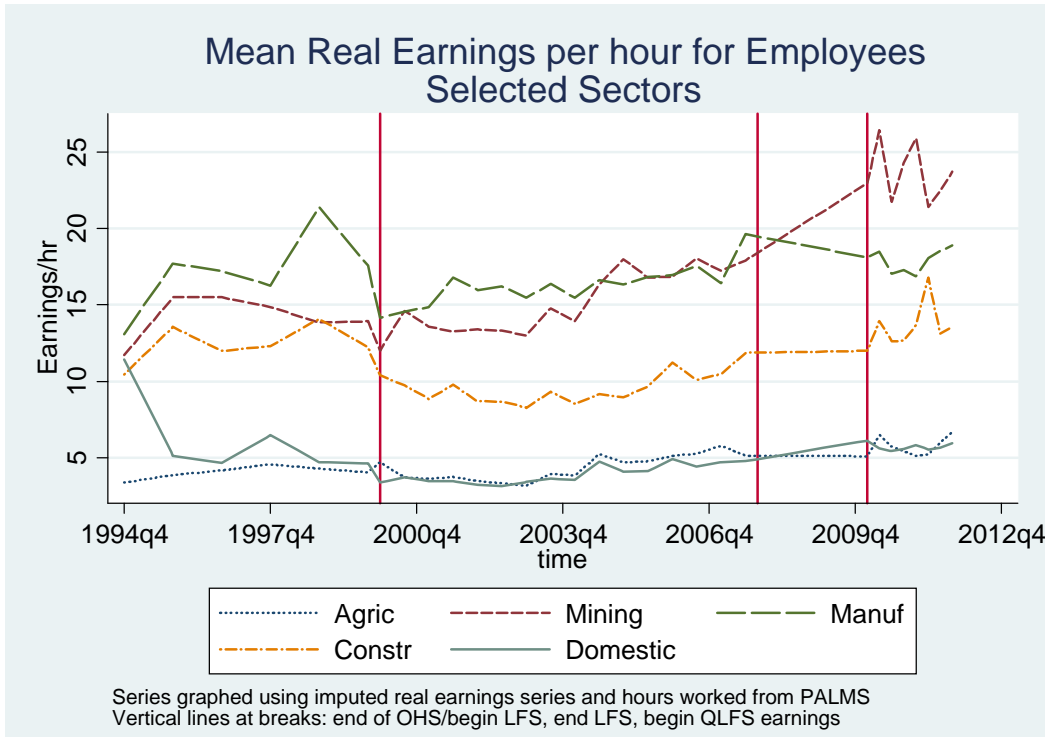


Figure 37. Real earnings per hour by sector among employees, 1994-2011 - Part 1

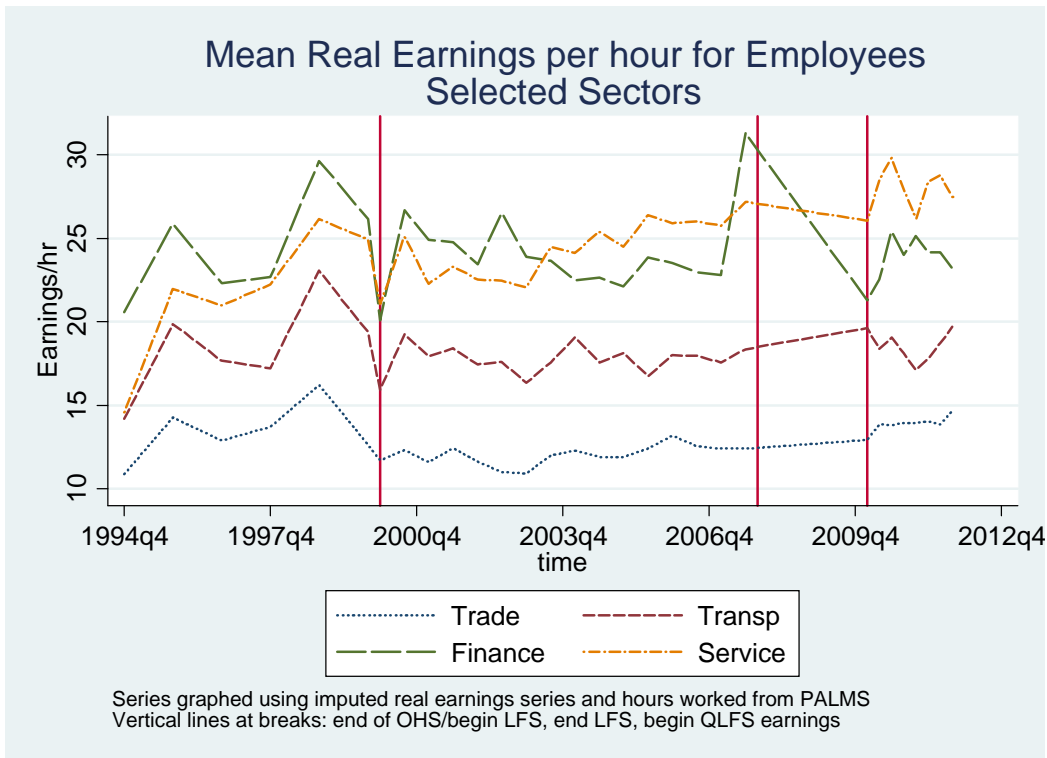


Figure 38. Real earnings per hour by sector among employees 1994-2011 - Part 2

5.3.3 Earnings per hour among the self-employed by Sector

Figure 39 and Figure 40 provide the sectoral trends among the self-employed. It is again fairly clear (compare to Figure 14) that earnings among the self-employed in the Financial sector are higher than in other sectors of the economy. There is less strong evidence of growing real wages, suggesting that the aggregate trends shown in Figure 36 are driven at least in part by the changing sectoral composition of the South African economy.

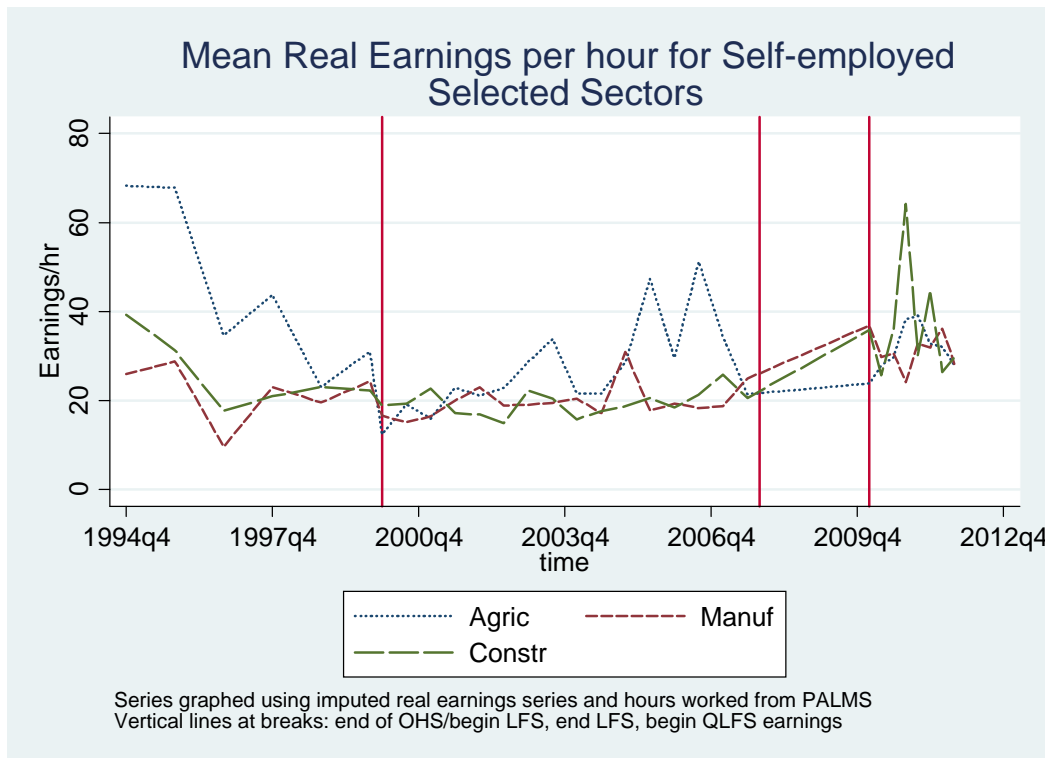


Figure 39. Real earnings per hour by sector among the self-employed 1994-2011 - Part 1

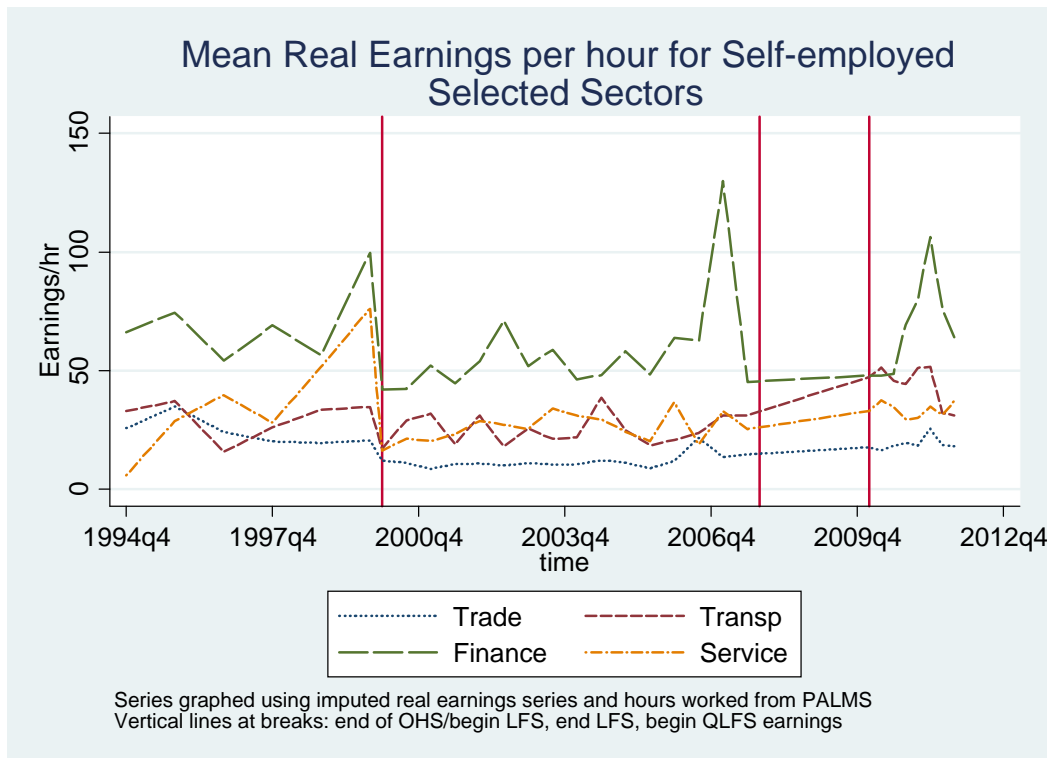


Figure 40. Real earnings per hour by sector among the self-employed 1994-2011 - Part 2

6. Reconciling the firm level wage series with the micro household data series

One of the key puzzles in South African labour economics is the big difference between the average earnings figures reported in the QES and the QLFS. As noted earlier, median earnings reported in Statistics South Africa’s “Labour Market Dynamics in South Africa in 2011” were around R3000 per month (Table 4-38, p4-51). By contrast the average wage reported in the QES for the same period was in excess of R12 000.

These differences can have huge consequences. One idea is that of a minimum wage tied to the average national wage. Depending on whether the QLFS median or the QES average was used this would make very dramatic differences to the level of the minimum wage. Getting a sense of what drives these estimates so far apart may help one understand something about the structure of the South African economy as well as something about the quality of the information in the QLFS and QES respectively.

Our line of attack is to systematically tick off the measurable differences, in particular changing the definition of the sample, so that the coverage of the QLFS becomes comparable to that of the QES. We note that mean wages can be written as

$$mean = \frac{\sum earnings}{wage\ earners} \quad (1)$$

Differences can therefore arise both due to differences in the numerator and in the denominator. To anticipate the findings below, we will show that the QLFS does not find enough earnings (the

numerator is too small) while it finds more employees (the denominator is larger). We then construct some hypotheses as to what sort of errors, both of inclusion and exclusion could explain the gap.

Before exploring these issues further, it is useful to contrast some of the characteristics of the “macro” labour series with those obtained from “micro” data.

6.1 Macro and micro data

The most widely used macro labour series is that published by the South African Reserve Bank, based on the firm surveys of Statistics South Africa. The employment series is published in index form, so the difference in levels between the series is not evident. In Figure 41 we have graphed the SARB total employment series and the equivalent series from the household surveys. Agricultural employment and employment in domestic work have been stripped out of the household survey series, to make coverage more comparable. The micro series has been converted to index form so that the starting point of the two series would be identical⁶.

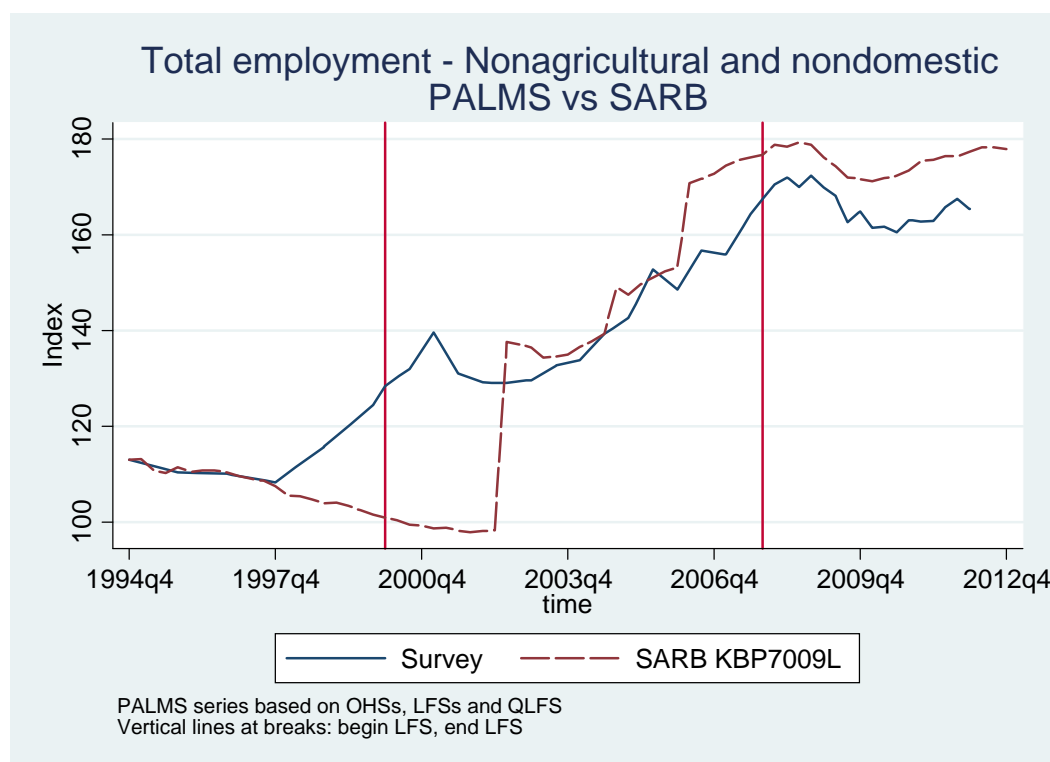


Figure 41. Comparing the firm surveys to the household surveys

The first point to emerge from this comparison is that the “macro” series is also subject to major breaks. The upward revision in employment estimates due to the revision of the business register in the third quarter of 2002 is very noticeable. Secondly the employment trends diverge in the late 1990s. Whereas the household survey data records higher levels of employment (at least in part due to better coverage), the firm surveys suggest job losses through to mid-2002. It should be noted that we have repeatedly suggested that the early parts of the OHSs were probably overestimating employment (particularly OHS95), so although the two series seem to overlap impressively up to 1997, it is not clear that the job losses in the OHSs over this period are entirely believable.

⁶ Note that the SARB series is benchmarked to June 2000. The micro series is set so that October 1994 is 113. This is the “benchmarking” for this series. It is clearly not benchmarked to June 2000.

Comparable earnings series are given in Figure 42. In this case the series from the household surveys has also been indexed to June 2000⁷. It is clear that the firm level series estimates much higher levels of earnings growth over the entire period than the micro series does. In the light of our discussion of wage inequality above, this is a sign that the firm data has a higher representation of high income individuals (earning above the median) than the microdata series does.

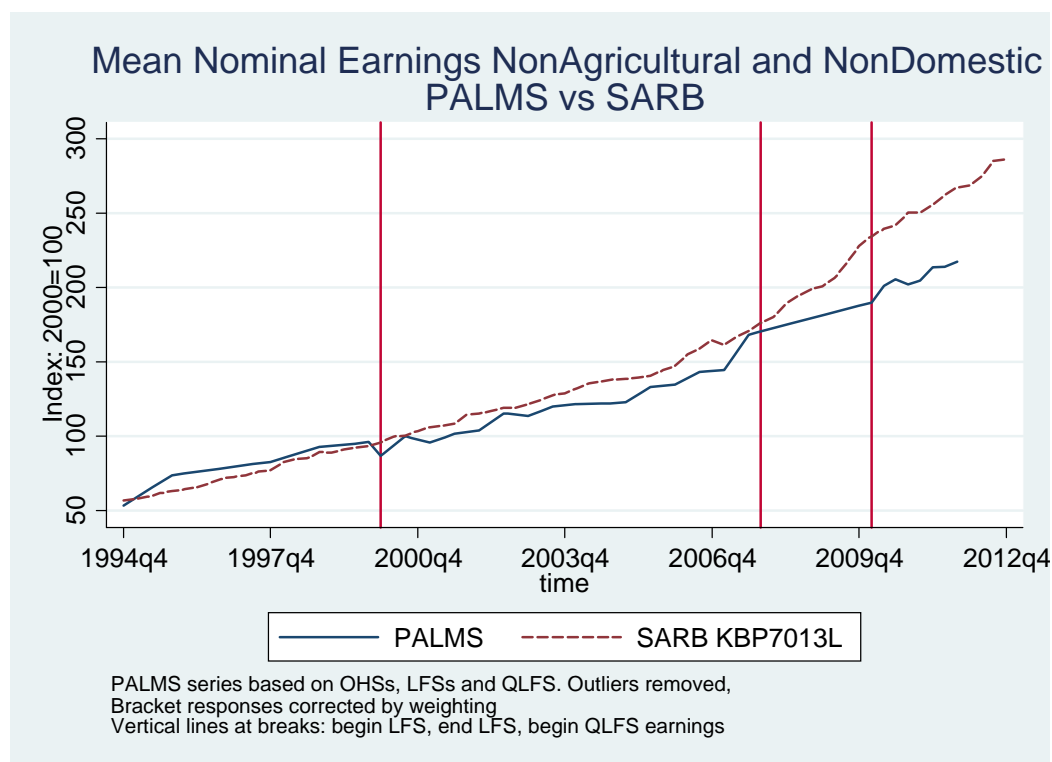


Figure 42. Nominal earnings over time in the firm surveys and household surveys.

Given this divergence in behaviour over the period as a whole it is not surprising that it isn't easy to get agreement even for a particular time period. In order to explore some of the reasons for the divergence we will focus on data from 2011, so that breaks in the series will not complicate the discussion.

⁷ June 2000 is so clearly anomalous in the case of the SARB employment series that this was not a good benchmark for comparing the two series.

6.2 The measurement used: Medians and means

The point of departure for this analysis is the “raw” comparison alluded to previously - median earnings in the QLFS are less than a quarter of the mean earnings reported in the QES⁸. The salient comparisons are shown in Table 1. We see that, indeed, the median in the QLFS hovers around R3000 per month, while the average monthly earnings in the QES are over R12 000.

The earlier discussion should have alerted us to the fact, however, that the median and the mean are quite different in these surveys. The more appropriate comparison is therefore between the mean of around R6 500 in the QLFS and the mean of over R12 000 in the QES. Nevertheless even this difference is very large.

Table 1. Raw comparisons between the QES and QLFS

Quarter	QLFS		QES	QLFS			QES		
	Median	Mean	Mean	Gross (millions)	Employ (1000s)	Implied Mean ⁹	Gross (millions)	Employ (1000s)	Implied Mean
2011 Q1	3 000	6 426	12 262	86 469	13 626	6 346	106 116	8 289	12 802
2011 Q2	3 000	6 989	12 123	94 141	13 630	6 907	107 882	8 300	12 998
2011 Q3	3 000	6 526	12 923	89 196	13 857	6 437	113 438	8 358	13 572
2011 Q4	3 250	6 598	13 284	91 283	14 006	6 517	121 473	8 381	14 494

Considering the expression for the mean (equation 1), it is evident that the differences between the two means (columns 3 and 4) could be due to either differences in the numerator or the denominator. In Table 1 we break up this comparison by looking at these elements. It is evident that there are very large differences in both. The estimated total sum of earnings in the QLFS is around R90 billion, while the monthly figure¹⁰ for the QES is around 20% higher. The work force in the QLFS, however, is over 60% greater than the equivalent QES figure. As a check we show that the ratio of the two figures reproduces broadly the difference shown in the third and fourth column. They are not precisely the same in the case of the QLFS, since individuals with missing earnings are in the denominator, but contribute nothing to the numerator. It is not clear where the divergence in the case of the QES arises.

6.3 Coverage differences

6.3.1 Individuals covered – Employees, Agriculture and Domestic Work

There are several important differences in the population covered by the QLFS and the QES respectively. Firstly, the QES employment numbers do not cover the self-employed or working proprietors, i.e. they only cover employees. Secondly, the QES excludes the agricultural sector. It is also obvious that domestic workers will not be covered by any firm survey. In Table 2 we have adjusted the QLFS sample to correspond to this restricted population. The average earnings for this

⁸ All QES statistics are taken from Statistics South Africa, Statistical Release P0277, December 2011. The QLFS statistics are my own calculation from Statistics South Africa microdata, with the cross-entropy weights contained in PALMS.

⁹ The implied mean is smaller than the estimated mean (in column 3) because individuals who did not report any earnings swell the numerator but not the denominator. When the mean gets estimated in the usual way any cases with missing earnings are excluded. In this exercise we focus separately on the numerator and the denominator. The issue of missing information (in general) is discussed later.

¹⁰ The figures reported in Statistical Release P0277 are quarterly, so were divided by three to get comparable numbers.

population is marginally higher (about 4% on average across the four quarters). Total earnings have dropped considerably, mainly because the high earning self-employed (professionals etc.) have been removed. This widens the gap in the denominator. The shortfall is now of the order of R45 billion! The earnings reported in the QLFS would have to be higher by around 40% to close this gap. By contrast the difference in the denominator has shrunk. The QLFS numbers are still higher by about 18%.

Table 2. Adjusting the QLFS to correspond to the population covered by the QES

Quarter	QLFS			QES		
	Gross (millions)	Employ (1000s)	Implied Mean	Gross (millions)	Employ (1000s)	Implied Mean
2011 Q1	63 662	9 680	6 577	106 116	8 289	12 802
2011 Q2	66 522	9 672	6 878	107 882	8 300	12 998
2011 Q3	68 059	9 886	6 884	113 438	8 358	13 572
2011 Q4	70 131	10 028	6 994	121 473	8 381	14 494

6.3.2 Sectoral coverage – a detailed look at the other sectors

A more detailed look at the sectoral comparisons is given in Table 3. The figures presented in this table are for the first quarter of 2011. It is evident that the mismatch is much bigger in some sectors than in others. Looking at the gap between the sectoral contributions to total earnings (i.e. the numerator), it is clear that there are particularly large gaps in the Finance and Services sectors. The total quantum of earnings “lost” between them is R27 billion.

Table 3. Comparisons by Sector

2011 Quarter 1	QLFS			QES			Difference		
	Gross (millions)	Employ (1000s)	Implied Mean	Gross (millions)	Employ (1000s)	Implied Mean	Gross (%)	Employ (%)	Mean (%)
Mining	3 031	338	8 959	6 523	511	12 765	115.2%	51.0%	42.5%
Manufacturing	9 052	1 622	5 581	13 330	1 160	11 492	47.3%	-28.5%	105.9%
Utilities	787	97	8 107	1 438	59	24 373	82.7%	-39.2%	200.6%
Construction	3 563	806	4 422	3 867	413	9 362	8.5%	-48.7%	111.7%
Trade	9 764	2 055	4 752	13 893	1 645	8 445	42.3%	-20.0%	77.7%
Transport	3 735	620	6 022	5 904	360	16 399	58.1%	-42.0%	172.3%
Finance	11 984	1 457	8 224	26 677	1 798	14 837	122.6%	23.4%	80.4%
Services	21 648	2 680	8 078	34 484	2 343	14 718	59.3%	-12.6%	82.2%
Total	63 564	9 675	6 570	106 116	8 289	12 802	66.9%	-14.3%	94.9%

Interestingly the QES also records higher levels of employment in the Finance sector, suggesting that the gap may be a combination of underreporting of earnings as well as certain high earning individuals being simply missed in the QLFS. We will explore this issue further below.

6.3.3 Type of firms

There are some additional coverage differences between the QLFS and QES. Only formal firms of a particular size will be captured by the QES. Correcting the QLFS to align it with the QES is not all that easy. One straightforward filter is by the number of employees. Firms with only one employee are unlikely to be enumerated in the QES. Taking them out of the QLFS sample raises the average wage a little more. The figures are shown in the left panel of Table 4. One point to note is that the gap in the numerator has now increased a little more also.

Table 4. Adjusting the characteristics of the firm

Quarter	QLFS, 2 or more employees			QLFS, correct for part-time		
	Gross (millions)	Employ (1000s)	Implied Mean	Gross (millions)	Employ (1000s)	Implied Mean
2011 Q1	60 302	9 047	6 665	60 302	8 519	7 078
2011 Q2	63 002	9 022	6 983	63 002	8 578	7 345
2011 Q3	64 192	9 205	6 973	64 192	8 704	7 375
2011 Q4	66 336	9 330	7 110	66 336	8 857	7 489

6.3.4 Part-time work

In all the calculations thus far we have counted up the numbers of workers regardless of the hours they worked. It is unclear how firms would be inclined to report their work force, but it is likely to be in “full-time equivalents”, i.e. two part-time jobs would add one worker to the total. In Table 4 we also report a simple correction for part-time work: individuals who work less than 40 hours a week are reported as a fractional worker, with the fraction given by the ratio of their hours to 40, i.e. someone who worked 20 hours would be considered half a worker. This correction affects the denominator only; the numerator is as for the left panel of Table 4. This correction raises the average earnings more meaningfully; they are now around 12% above the “baseline” mean.

6.4 Possible sources of measurement error

The attempt to “close the gap” between the QLFS and the QES on the sample characteristics alone must be deemed to be a failure. We therefore need to consider what sorts of biases and measurement error there may be in the different instruments.

6.4.1 Missing high income earners in the QLFS

There is a lot of anecdotal evidence that household surveys find it difficult to penetrate affluent areas. Even with weighting adjustments for non-response, it is plausible that a handful of extremely rich individuals could significantly raise the QLFS earnings. Indeed a handful of “outliers” that were deleted in the data cleaning stage of this project managed to do that. Unfortunately we cannot determine how much the QES figures are swung by top management salaries, since there is no distributional information collected.

We can get a sense on how much we could plausibly have missed by comparing the top end of the income distribution to the 2011 tax tables. The tax tables record 63 000 “millionaires” (income per annum). The 2011 QLFS figures represent 31 000 legitimate millionaires, of whom around 7.5 thousand would have been “employees”. This suggests around a 50% undercount. The total monthly earnings of the 7.5 thousand “millionaires” is around R1 billion, suggesting that we may be missing earnings of around that magnitude. Going a step or two down the tax tables, there are around 220 000 tax payers with income between 500 000 and R 1 million (per annum). The QLFS has around 170 000 individuals in this income range of whom about 90 000 are employees. Assuming that there may be around another 25 000 individuals “missing” from the QLFS in this earnings bracket would increase total earnings by another R1.5 billion.

6.4.2 Other missing earners in the QLFS

The sectoral comparisons in Section 3.2 above suggest that the QLFS has missed around 170 000 mine workers. These are likely to be truly missing since it seems hard to believe that mine employment could be misclassified. Furthermore it is plausible that coverage of hostels and other single person living quarters may be less than perfect. The mining employment figure from the firm side is likely to be very accurate. 170 000 extra miners at the QLFS average wage for that sector of around R 9 000 would add another R 1.5 billion to the aggregate.

It is less plausible to make an equivalent adjustment for the mismatch between the QLFS and QES employment numbers in the Financial Sector. In this case it is at least equally plausible that the 340 000 “missing” Financial Sector workers are in the numbers somewhere, but just with the sector misclassified, e.g. in “Services”.

6.4.3 Misclassification of employees in the QLFS

It may be possible that some of the individuals classified as “self-employed” or as “employers” may in fact be incorrectly classified. This might be the case for (non-owner) managers in large firms, who might think of themselves as “employers”, even though it is the firm that pays the bills. This would certainly raise the total earnings, although it would also increase the denominator. It is impossible to undo this type of classification error on the data and it would be idle to speculate how much this might contribute to narrowing the gap. It is clear, however, that it cannot do so completely, since all the self-employed are included in the QLFS figures in Table 1 – and there is a R20 billion earnings gap even with the self-employed.

6.4.4 Missing earnings in the QLFS

It is quite possible that individuals underreport earnings in the QLFS. Several types of underreporting may happen. Firstly, the QLFS only asks for income from the “main job”. To the extent to which individuals work for more than one employer, they may in fact be earning considerably more. In essence they would appear potentially multiple times in the QES (with different employers) but all their earnings would be recorded. Of course this suggests that the QES employment numbers may suffer from some double counting.

Secondly, individuals in the QLFS are requested to report their pre-tax and pre-benefits income. It is not clear whether this is consistently done. Given that at least some of the responses would have come from proxy responders, it is possible that some of the information in the QLFS is actually about the cash portion of the earnings. Looking at the mining information in Table 3 again, the difference between the QES and the QLFS number could plausibly be attributed to the difference between “cost of employment” and take home pay. The QLFS mean is 30% lower than the QES figure. Reversing the perspective, the QLFS figures would have to be scaled up by more than 40% to get to the QES numbers. Is it plausible that the QLFS figures could be off by such a lot? One indication that creates some doubt is the fact (noted in the discussion of “real earnings” earlier) that when the questionnaires explicitly allowed for “take-home pay” response, the measured incomes were considerably lower, suggesting that the bulk of the respondents do not, in fact, seem to get the concept that badly wrong. Short of getting actual payroll data for a considerable section of the population (e.g. through a sample of tax records) it would be impossible to get a sense whether or not this is, in fact, the case.

6.4.5 Errors of inclusion in the QES

It is possible that the total earnings figure in the QES is too high. This would be the case, for example, if the people tasked with reporting to Stats SA included the earnings of any people other than “employees”. In some firms (e.g. large professional service firms) it is possible that some of the partners pay themselves a monthly salary as well as periodic shares of the profit. Would the person filling out the returns to Stats SA strip out “salaries” paid to partners or their medical aid contributions? This type of error would be the converse of the “misclassification” case considered earlier. Unfortunately short of auditing particular returns (which would require access to the confidential firm level information) it is also impossible to correct for this type of problem.

Other potential errors of inclusion would be if any of the returns included extraordinary payments that were not part of the standard wage or salary package. Again there seems to be little prospect of isolating how large these errors might be.

6.4.6 Errors of omission in the QES

It is equally possible that the denominator in the QES is too small. “Employees” should include casual and temporary labour, but it is not clear how accurately these would be captured, particularly if some of them get hired by the day, e.g. to deal with a Monday morning short-staffing problem. Their total wage bill is likely to be measured quite carefully. It is plausible that the person tasked with filling out the Stats SA returns may report only on the “usual” staff complement.

6.5 Portioning out the sources of error

Given the size of the gap in the numerator (i.e. around R40 billion in earnings) it is somewhat heroic to try to reconcile the figures. Furthermore we have not tried to apportion any of the gap to misclassification errors (Section 6.4.3) or to inclusion errors in the QES (Section 6.4.5) which may, in fact, be substantial.

6.5.1 Coverage and measurement differences

The differences in the sample and in the measurement of “average” income have been covered in Sections 6.2 and 6.3. We start, therefore with the final adjustment to the QLFS, given by the last column in Table 4. This leaves still an enormous gap between the average of around R7 000 and the QES figure of more than R12 000.

6.5.2 Possible errors in the QLFS

Table 5 presents a set of possible adjustments to the QLFS figures, beginning with the adjusted QLFS figures. We begin by adding in around 35 000 very high earners (based on the analysis in section 6.4.1) with an addition to total earnings of around R 2.6 billion. This figure is based on inflating up the actual total in the QLFS sample proportionately to the “missing” high income earners. We also add in 172 thousand mine workers, giving each one the average income from the QLFS. In the final panel we consider a 40% increase in the QLFS earnings, to account for a possible difference in the “post-deductions” versus “pre-deductions” reporting¹¹. The figure of 40% corresponds approximately to the difference between the QLFS average wage in the mining sector and the average

¹¹ Despite our doubt that the error could be of this magnitude.

wage calculated in the QES. The firm figures for the mining sector are likely to be as accurate as they can be, and this sector is likely to suffer less from sectoral misclassification than most of the others.

Table 5. Adjustments to bring QLFS closer to QES

QLFS adjust			High earners			Mine workers			Underreport earnings		
Gross (mill)	Employ (1000s)	Av	Gross (mill)	Employ (1000s)	Av	Gross (mill)	Employ (1000s)	Av	Gross (mill)	Employ (1000s)	Av
60 302	8 519	7 078	62 898	8 554	7 353	64 445	8 727	7 385	90 223	8 727	10 339
63 002	8 578	7 345	65 597	8 613	7 616	67 144	8 785	7 643	94 002	8 785	10 700
64 192	8 704	7 375	66 787	8 739	7 642	68 334	8 912	7 668	95 668	8 912	10 735
66 336	8 857	7 489	68 932	8 892	7 752	70 479	9 065	7 775	98 670	9 065	10 885

The final figure is closer to the QES figure, but still some way short.

6.5.3 Possible errors in the QES

We have much less detailed information on the QES, so this is more hypothetical even than the discussion of the last section. Given the large gaps between the (adjusted) employment figures in the QLFS and the QES, it is plausible that employment is somewhat undercounted in the QES. Allowing for a 5% undercount brings the two figures into closer alignment. The effect of this would be to cut the QES implied average income by just under 5%. (See Table 6)

Table 6. Allowing for underreporting of employment in the QES

QES			5% underreport workers		
Gross (millions)	Employ (1000s)	Implied Mean	Gross (millions)	Employ (1000s)	Mean
106 116	8 289	12 802	106 116	8 703	12 192
107 882	8 300	12 998	107 882	8 715	12 379
113 438	8 358	13 572	113 438	8 776	12 926
121 473	8 381	14 494	121 473	8 800	13 804

6.5.4 Decomposing the difference

Even after all of these adjustments there is still a sizable gap. Undoubtedly there will be other measurement issues (like the misclassification errors that we flagged) that we can't address. It is also possible that the imputation algorithms used for missing information in both the QLFS and the QES may contaminate some of these results. Lastly both the QES and the QLFS need weights to gross up to national figures. Smallish problems on a survey could "balloon" if the weights on those problematic observations turn out to be high.

In Table 7 we present the results of this accounting exercise. The initial difference between the mean and the median turns out (unsurprisingly after the earlier discussion) to matter a lot. The coverage issues turn out (surprisingly) to explain relatively little of the overall gap. Making the correction for part-time work seems important. High income earners are almost definitely underrepresented in the microsamples, but their presence is not enough to shift the distribution (at least not if the "missing" high earners have incomes similar to the ones that we do see in the data). Of the remaining possible explanations the one that would get us closest to closing the gap is the difference between a "cash"

wage and the full cost of employment. The 40% raising factor that we used (based on the mining sector) would be equivalent to a 30% underreporting of full incomes (on average). That is not implausible given the tax rates and the costs of other benefits, although it is somewhat difficult to believe that the bulk of the respondents would get the question that badly wrong.

Table 6. Apportioning the gap between the QLFS and QES figures

Quarter	Gap	Mean vs Median	Coverage	Firm size	Part time	High earners	Mine workers	Under-report earnings	Under-report workers	Residual
2011 Q1	9 262	36.1%	2.5%	1.0%	4.5%	3.0%	0.3%	31.9%	6.6%	14.2%
2011 Q2	9 123	42.8%	-0.3%	1.2%	4.0%	3.0%	0.3%	33.5%	6.8%	8.8%
2011 Q3	9 923	34.6%	4.5%	0.9%	4.0%	2.7%	0.3%	30.9%	6.5%	15.5%
2011 Q4	10 034	32.6%	4.7%	1.2%	3.8%	2.6%	0.2%	31.0%	6.9%	17.0%

Even after all of those adjustments, there is still a residual gap of around 15%. It is hard to know what to make of it. What is really required is the sort of hard data (e.g. PAYE information) that would give us both accurate aggregate earnings information as well as good measures of its distribution. That would help us understand the quality of both the QLFS and QES earnings information.

It is important to note that even if incomes are underreported, it is highly unlikely that this would change the distributional picture discussed earlier. The people facing the highest tax rates and with the biggest set of benefits are also the people at the top end of the distribution. If anything, the “missing” information is likely to exacerbate that picture.

Indeed the importance of the “mean vs median” in Table 7 as well as the role played by high earners is a reflection of the extent of inequality in this society. This was an issue already highlighted in relation to the increase in the gap between the 90th percentile and the median in the wage distribution, as well as the systematic widening of the gap between the median and the mean. Indeed the more rapid growth of earnings in the firm surveys suggests that the firm surveys have a better handle on high income earners. A key difference between the QLFS and the QES is that the QLFS has much better coverage of low earning forms of work (including domestic work and agriculture) and much worse coverage of the small layer of very high earners. Unfortunately the high earners are quite important for the mean.

7. Productivity

In this section we will discuss the estimation of annual labour productivity measures both for the economy as a whole and sectorally.

7.1 Sources used in the measurement of annual GVA¹²

This discussion will focus on the estimation of **annual** GVA and GDP. A short discussion on the quarterly figures will follow at the end. There are three processes involved:

- Estimating the output of the economy (and of the sectors) from the production side

¹² The information in this section is based on South African Reserve Bank (2010), Statistics South Africa (2012), Statistics South Africa (2009) and telephonic discussions with Mr Gerhard Bouwer and Mr Patrick Kelly of Statistics South Africa. The latter two are not responsible for any errors of interpretation on my part.

- Estimating the elements of expenditure
- Getting the two sides to balance through the construction of “Supply and Use Tables”

The information used to measure production comes in two types:

- **Annual surveys**

Chief among these is the series “Annual Financial Statistics” (Statistical release P0021). This is a survey of around 12 000 firms from the Business Register designed to be representative of all sectors, except for Agriculture, the Financial Sector and General Government Services. This survey asks detailed questions about turnover, costs (including taxes paid and subsidies received), inventories etc. These surveys are stratified by sector and size of firm and are designed to give results accurate to two digit SIC level.

- **Benchmark surveys**

Statistics South Africa periodically runs “Economic Large Sample Surveys” (ELSS) for most sectors of the economy. These provide detailed information on the structure of particular sectors at a finer level of disaggregation, i.e. they provide results down to four digit SIC level. These surveys are also extracted from the Business Register, i.e. they are designed to be representative of all firms that are registered for tax.

Every five years the South African Reserve Bank and Statistics South Africa do a benchmarking exercise for the National Accounts. The last benchmarking was done in 2009, the next one is due in 2014. During the benchmarking process the latest ELSS available will be used. The purpose of the benchmarking is to fix, as accurately as possible, the nominal value of economic activity in that year. The annual updates then take that structure as reference (which is why the annual figures are subject to revision following another benchmarking exercise). Furthermore in the benchmarking process the base period used for the “constant price” series is moved to a more recent year.

The following table presents the sources used, per sector, in the calculation of the GVA figures:

Sector	Sub-sector	Source	StatsSA Series
Agriculture, Forestry and Fishing	Agriculture	Benchmark: Census of Commercial Agriculture 2007 Annual Survey: Agricultural Survey (latest available 2011) Note: “Estimates from the census and survey results are verified against quarterly source data collected by the National Department of Agriculture. Estimates include the value of farm produce consumed by farmers for own account.” (SARB 2010, p.10)	P1102 P1101
	Forestry & Fishing	Annual: Annual Financial Statistics	P0021
Mining and quarrying		Benchmark: Mining LSS 2004 (most recent available Mining LSS was conducted in 2009) Annual: Annual Financial Statistics Note: “supplemented by working results of the Chamber of Mines in respect of the goldmining industry, information from the Department of Minerals and Energy and sample surveys	P2001 (also Report 20-01-02) P0021

		conducted by the SARB” (SARB 2010, p.10)	
Manufacturing		Benchmark: Manufacturing LSS 2005 (most recent Manufacturing LSS is 2011) Annual: Annual Financial Statistics	P3002 (also Report 30-02-04) P0021
Electricity, Gas and Water		Benchmark: Electricity, Gas and Water LSS 2006 (most recent available LSS was done in 2010) Annual: Annual Financial Statistics	P4001 (also Report 41-01-02) P0021
Construction		Benchmark: Construction LSS 2007 (most recent available LSS is 2011) Annual: Annual Financial Statistics	P5002 (also Report 50-02-01) P0021
Wholesale and retail trade, catering and accommodation	Wholesale trade	Benchmark: Wholesale Trade LSS 2005 (most recent available LSS is 2009) Annual: Annual Financial Statistics	P6101 (also Report 61-01-01) P0021
	Retail trade	Benchmark: Retail Trade LSS 2005 (most recent available LSS is 2009) Annual: Annual Financial Statistics	P6201 (also report 62- 01-02) P0021
	Motor trade	Benchmark: Motor Trade LSS 2006 (most recent LSS is 2009) Annual: Annual Financial Statistics	Report 63-01-01 P0021
	Accommodation	Benchmark: Accommodation LSS 2007 ¹³ (most recent available LSS is 2009) Annual: Annual Financial Statistics	P6411 (also Report 64-11-01) P0021
	Catering	Benchmark: Food and Beverages LSS ¹⁴ (most recent LSS is 2009) Annual: Annual Financial Statistics	P6421 (also Report 64-20-01) P0021
Transport Storage and Communication		Benchmark: Transport Posts and Telecommunication LSS 2006 (in 2010 this was split into a Transport and storage LSS; and a Posts and Telecommunication LSS) Annual: Annual Financial Statistics Note: Sources also include “Financial statements of major role players such as Transnet, Telkom and South African Post Office” (Statistics South Africa 2012, p.65)	P7000 (Report 71-02-01) (Report 75-01-01) P0021
Financial intermediation, insurance, real-estate and business services	Financial intermediation	Benchmark and annual: SARB surveys of financial firms	Not covered by Stats SA
	Real estate and	Benchmark: Real estate and business services LSS	Report

¹³ The SARB (2010) report states that the 2007 Accommodation LSS was used. Statistics South Africa (2012) lists the 2004 survey as being the source.

¹⁴ It is not clear whether the 2007 Food and Beverages LSS was available at the time of the 2009 benchmarking process. The SARB (2010) report doesn't mention it, nor does the Statistics South Africa (2012) GDP release. I can't find any references to a LSS between 2007 and the 1994 “Census of contract caterers” (Report 64-22-01).

	business services	2006 (Latest available LSS is for 2010) Annual: Annual Financial Statistics	80-04-01 P0021
Community, social and personal services	General government	Benchmark/ Annual: "Stat SA conducts a number of surveys on the three tiers of government that are used to derive the estimates. They are supplemented with information from SARB." (Statistics SA 2012, p.65). Surveys: Financial statistics of consolidated general government Financial Census of Municipalities	 P9119.4 P9114
	Personal services	Benchmark: Personal Services Industry LSS 2004 (Latest available is LSS for 2008) Annual: Annual Financial Statistics	P9001 P0021

These sources provide information on the formal sector. The national accounts are adjusted (see e.g. Table G, p.8 of Statistics South Africa 2012) to make provision for:

- Tax evasion
- The informal sector
- Production for own use
- "Other non-observed"

An estimate of the value of unrecorded activity due to tax evasion comes from SARS seizure operations (Statistics South Africa 2012, Table N, p.62). Information on other illegal activities (e.g. prostitution) comes from SAPS and from organisations working with the sector (SWEAT).

Information on the structure of the informal sector (particularly its distribution across the sectors of the economy) is obtained from the Survey of Employers and the Self-employed (SESE), Statistical series P0276. The 2004 survey was used during the "benchmarking" process. The Quarterly Labour Force Surveys are used in the annual updating process.

The quarterly GDP figures are derived independently, using volume indicators from Statistics South Africa monthly and quarterly series (e.g. P6420, the monthly survey of the food and beverage industry). Once a year they are reconciled with the annual figures.

7.2 The measurement of annual GVA

There are three measures of value added in the national accounts provided by the South African Reserve Bank: Gross Value Added at Factor Cost (series KBP6003J0), Gross Value at Basic Prices (KBP6645J/Y) and Gross Domestic Product at Market Prices (series KBP6006J). Statistics South Africa (2009, p.26) makes the distinction between Output at Basic Prices, Output at Producers' Prices and Output at Purchasers' Prices:

"The relationship between the above-mentioned concepts can be expressed as follows:

Output at basic prices
plus taxes on products (excluding VAT)
less subsidies on products

= Output at producers' prices
plus trade and transport margins
plus non-deductible VAT
= Output at purchasers' prices

Basic prices are the preferred method of valuing output of goods and services produced for the market, especially when a system of VAT is in operation.”

In the discussion below we therefore focus on GVA at basic prices. Conceptually GVA at basic prices is the sum of compensation of employees, net operating surplus, consumption of fixed capital, taxes on production (but not taxes on the product – those go into GDP at market prices), less subsidies on production (but not subsidies on the product – those also go into GDP at market prices).

In order to calculate these magnitudes it is necessary to value the total output and the various costs associated with production.

7.3 The measurement of the Cost of Employment

According to the glossary provided by Statistics South Africa, the Compensation of Employees “is defined as the total remuneration, in cash or in kind, payable by an employer to an employee in return for work done by the latter during the accounting period. It is recorded on a gross basis, before any deduction for income taxes, pensions, unemployment insurance and other social insurance schemes. It also includes other forms of compensation, namely commissions, tips, bonuses, directors' fees and allowances such as these for holidays and sick leave, as well as military pay and allowances. It excludes employers' social contributions.” (Statistics South Africa 2009, Table X, p.94).

The estimates provided in the Cost of Employment Series published in the National Accounts (e.g. Table 6 in Statistics South Africa, 2012) will be based in large measure on the Annual Financial Statistics. That publication notes that these cost figures may not tally exactly with the gross earnings figures supplied in the QES:

“Employment costs estimates produced in this publication are based on information as defined by the International Accounting Standards 19 (IAS19) from an accounting perspective. The Quarterly employment statistics (QES) survey produces estimates for employment and earnings from the payroll perspective.” (Statistics South Africa 2011, p.72)

The main difference here is that the former will tend to work on an accrual basis, whereas the latter will be more on a cash basis. There are also differences in sample coverage and differences in where the information is sourced (accounting officer versus payroll officer).

The published CoE series is derived after the reconciliation of the production and expenditure sides of the national accounts in the Supply and Use Tables. It is therefore a series which has been subjected to some additional consistency checks.

Nevertheless the QES gross earnings series and the CoE series published in the national accounts track each other remarkably well – not only in trend but also in levels. This is shown in Figure 43. One would not be far off to treat the CoE as another measure of the same “gross earnings” construct that is estimated in the QES. This is of course also somewhat troubling given that the QES “gross earnings” exclude agriculture and employees in the informal sector.

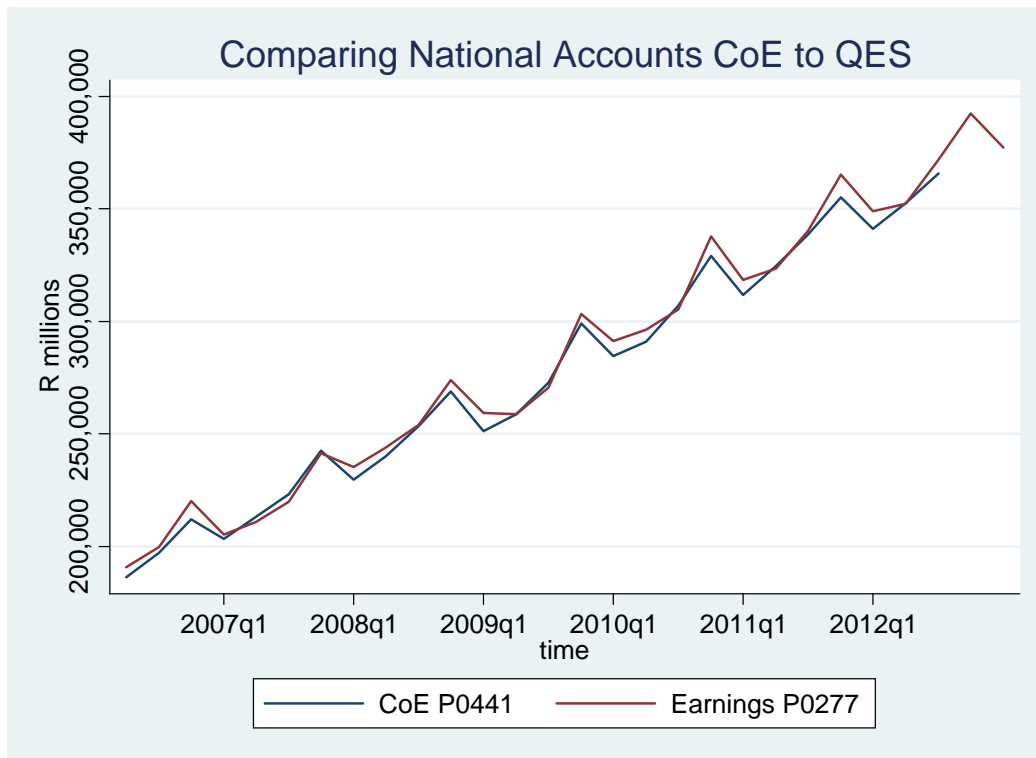


Figure 43. The National Accounts CoE series and the Gross Earnings Series from the QES show close agreement.

7.4 The evolution of labour productivity in South Africa 1994-2011

The definition of labour productivity is $GVA/labour$. Since GVA attempts to take the informal sector into account, the denominator should also incorporate the informal sector, i.e. it should use the full work force as measured in the OHS/LFS/QLFS surveys. We report on two measures of productivity, depending on whether the denominator is measured in terms of people (as per the employment series discussed in Section 4 or hours worked (Section 5.1). In either case we need to correct for the “spikes” in employment in 2000 arising from subsistence agriculture. In both cases we smooth over the spike.

7.4.1 Labour productivity using employment

In the case of employment, we reduce the total and the agricultural workforce in 2000 by 800 000. The effect of this adjustment is shown in Figure 44. The employment series graphed there is not strictly comparable to that given in Figure 21, because we have converted all employment to annual basis, by averaging the total estimates from surveys done in the same year.

There are no measures of “output” for the household sector. Correspondingly the employment in that sector should not be used in calculating labour productivity for the economy as a whole. The left hand panel in Figure 44 shows the employment series with the household sector removed.

Ideally there should be other adjustments to deal with the underenumeration of workers in the early OHSs and the relative over-enumeration among certain categories of self-employed in the LFSs. Unfortunately more extensive adjustments will require some guesses as to the counterfactual – e.g. to guess how many additional workers there would have been in backyard shacks in the early OHSs.

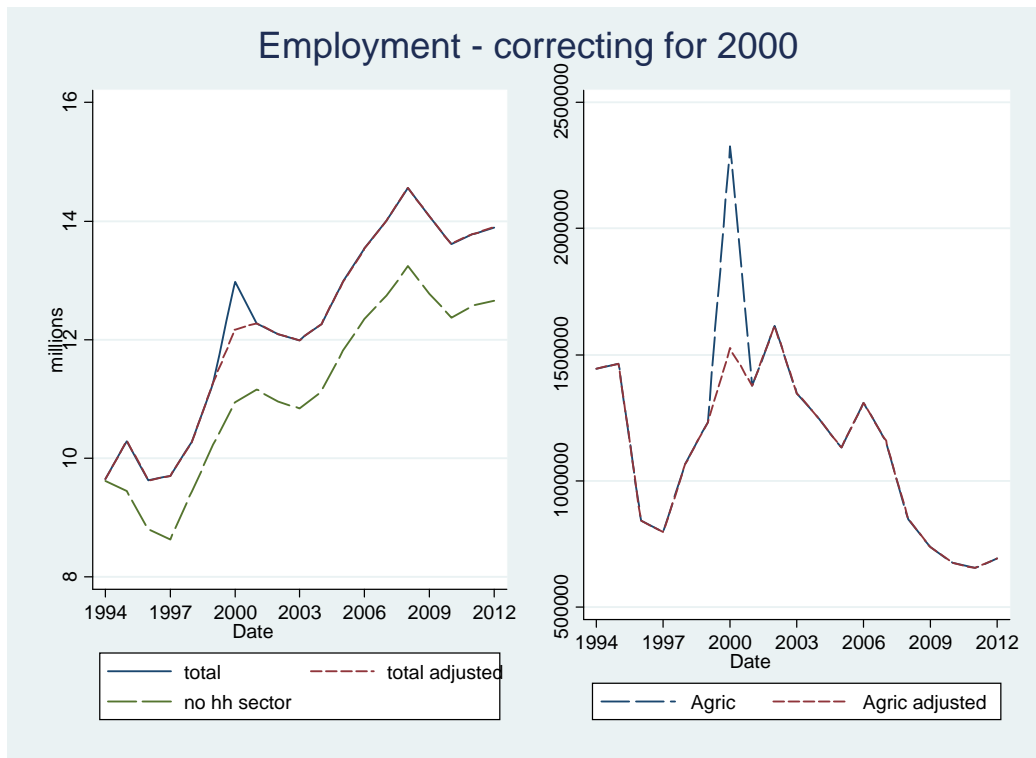


Figure 44. Annual Employment numbers

Using the Gross Value added series (at basic prices) from the South African Reserve Bank and total employment from PALMS we get the labour productivity series shown in Figure 45. Taking out the spike in 2000 reduces the downward blip between the October Household Survey period and the LFS/QLFS series. The increase and then decrease in labour productivity over the 90s is at least in part an artefact of an incorrect denominator. Taking the adjusted figure for 2000 (whether including or ignoring the household sector), the figures suggest a 30% increase in labour productivity since 2000.

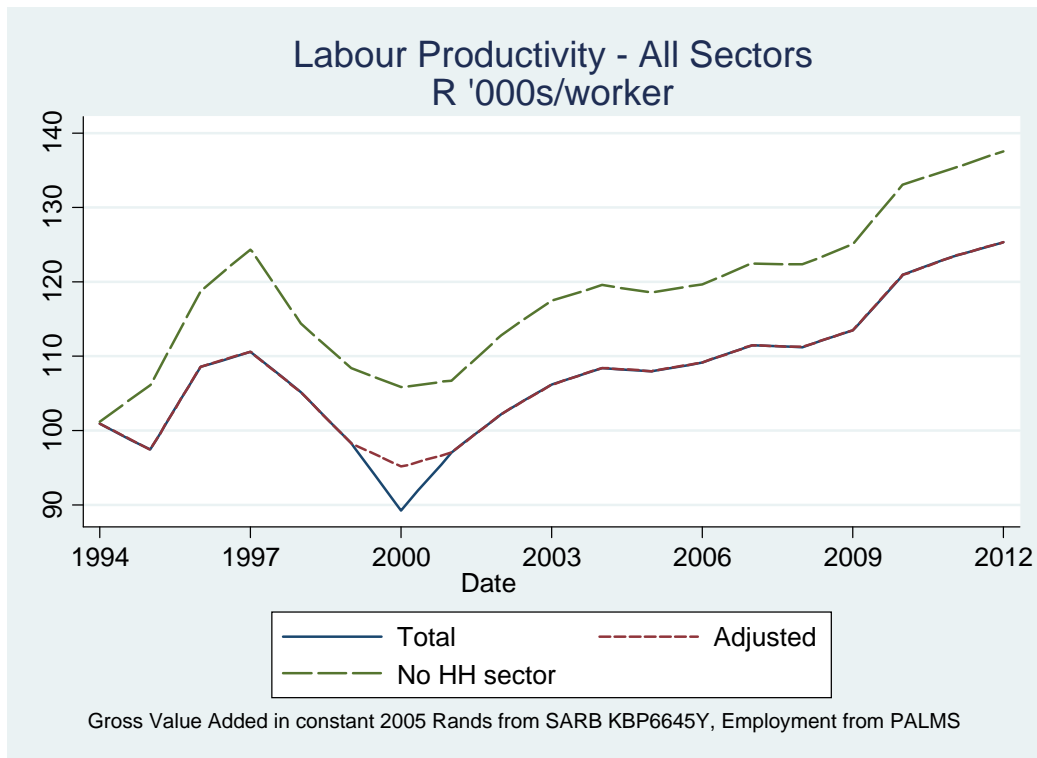


Figure 45. Labour productivity: 1994 to 2012

Labour productivity by sector is shown in Figure 46 and Figure 47. The blip downward in productivity in agriculture in 2000 is hardly noticeable in Figure 46. Overall there seem to have been strong growth in labour productivity in agriculture. Productivity in construction, trade and services seems to have been fairly flat across the entire period. The peak recorded in the services sector in 1997 is probably an artefact of problems in the denominator.

The sectors included in Figure 47 show growth in labour productivity since the late 1990s, with the single exception of the financial and business services sector. The first few points on the mining graph should be discounted because hostel dwellers were badly enumerated. The financial sector showed very strong growth in value added over this period, but even stronger employment growth. The result is the flat/decreasing appearance of the graph.

To characterise the picture crudely, the lower productivity sectors, i.e. those with productivity lower than R100 000/worker at the beginning of the period (shown in Figure 46) have typically also shown low productivity growth, whereas the sectors with higher initial productivity (shown in Figure 47) have also shown higher productivity growth – at least if one focuses on the period after 2000, where the employment series behaves better.

This divergence may be another sign of the increasing inequality remarked on in the first part of this report.

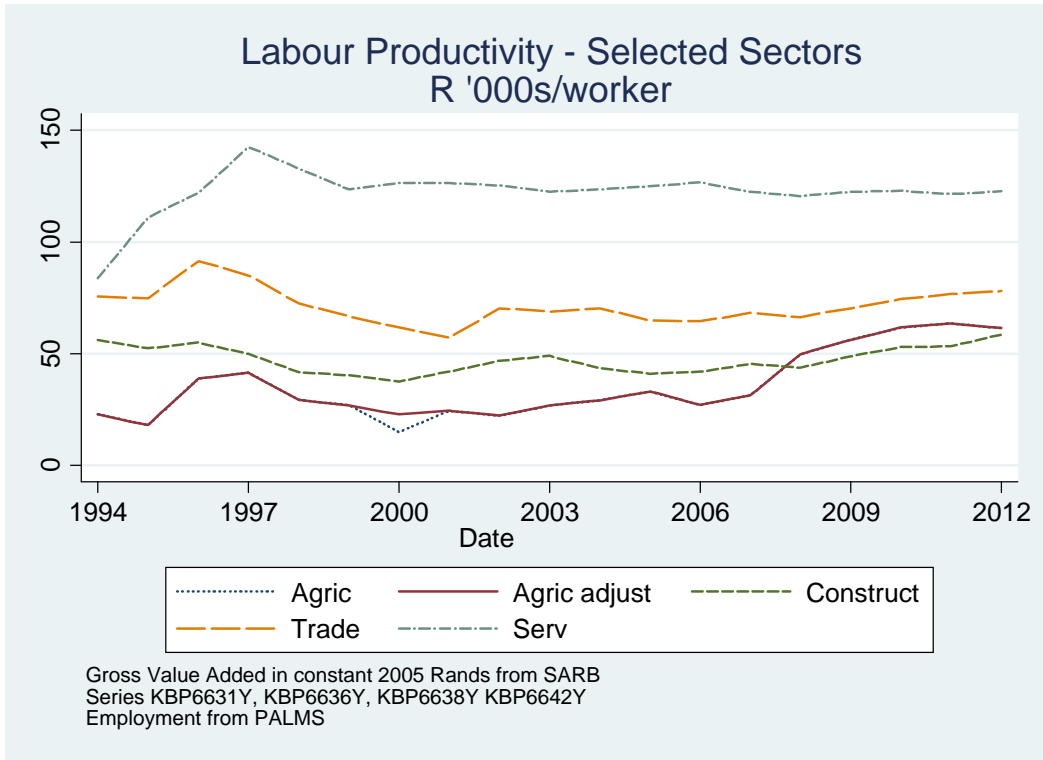


Figure 46. Labour Productivity by Sector 1994-2012 - Part 1

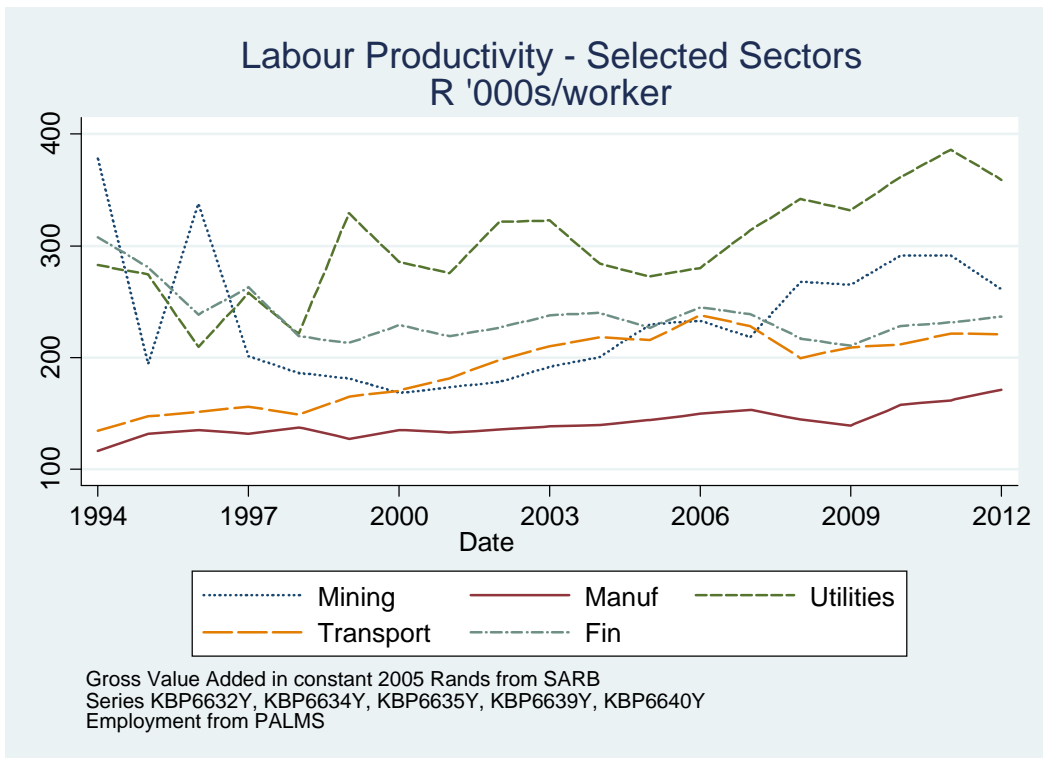


Figure 47. Labour Productivity by Sector 1994-2012 - Part 2

7.4.2 Labour productivity using total hours worked

In order to do the calculations in terms of hours, we multiply the weekly figures (from section 5.1) by 52 to get annual hours. The smoothing adjustment required to remove the 2000 spike is 1 billion hours. This suggests that the 800 000 “extra” agriculture workers worked around a 25 hour working week each. This is in line with the evidence shown in Figure 34. The effect is shown in Figure 48.

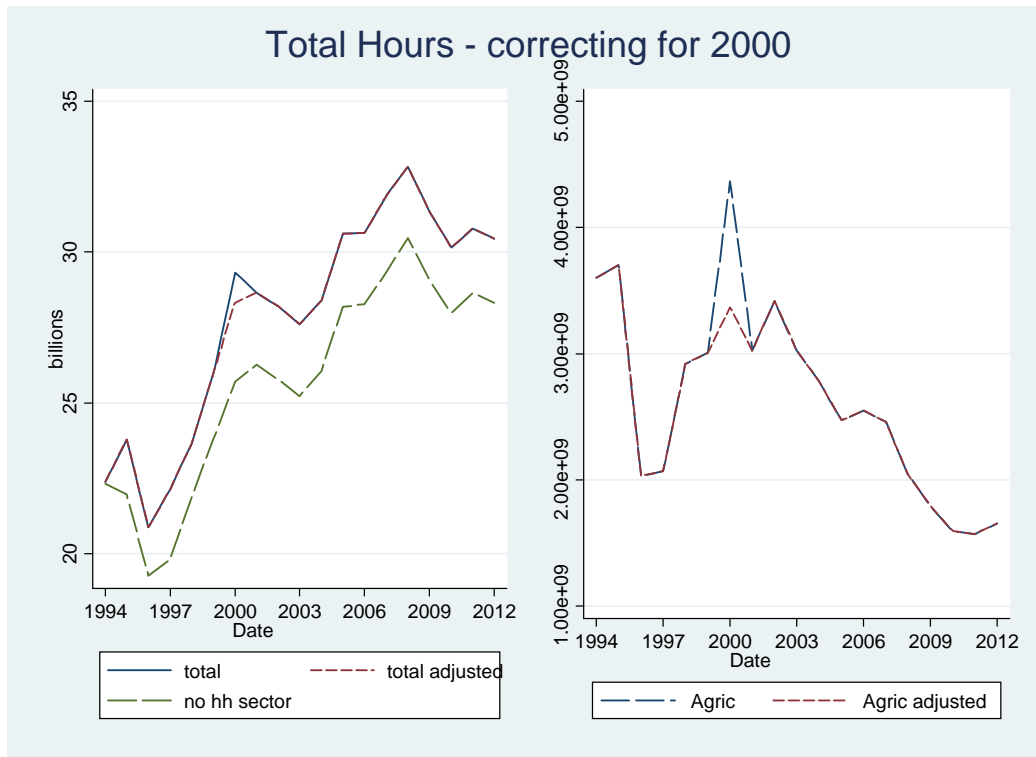


Figure 48. Smoothing over the spike in 2000

The overall labour productivity series is shown in Figure 49. Compared to Figure 45, the 90s look even odder. The increased enumeration of individuals working long hours has had a stronger effect on the denominator than the simple increase in employment would have suggested. Again it is clear that productivity has been increasing steadily since 2000. The increase is around a third (33%) using the series stripping out the household sector.

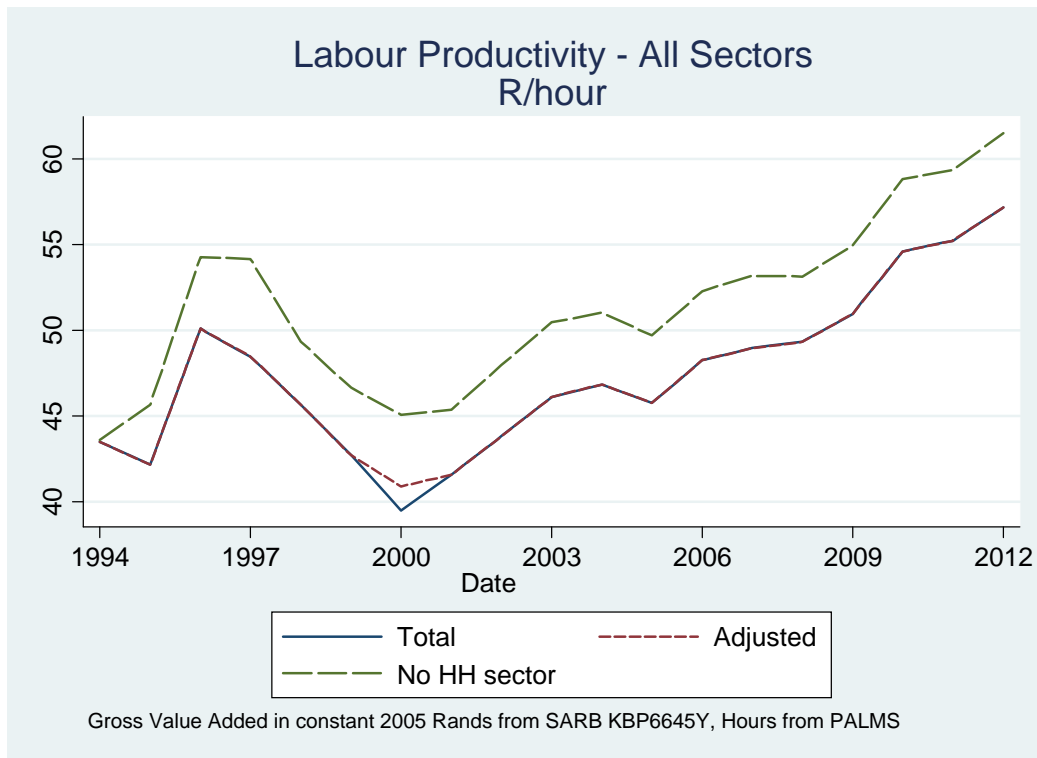


Figure 49. Labour productivity measured as output per hour worked

The sectoral patterns (analogous to Figure 46 and Figure 47) are given in Figure 50 and Figure 51. Note that the split between initially higher productivity sectors (in this case output above R50 per hour) and initially lower productivity sectors (output below that level) mirrors the split in terms of output per employee. The pictures also confirm the broad trends identified earlier. Sectors with higher value added (in terms of output per hour worked) at the beginning seem to have had stronger productivity growth, with the exception of the financial sector.

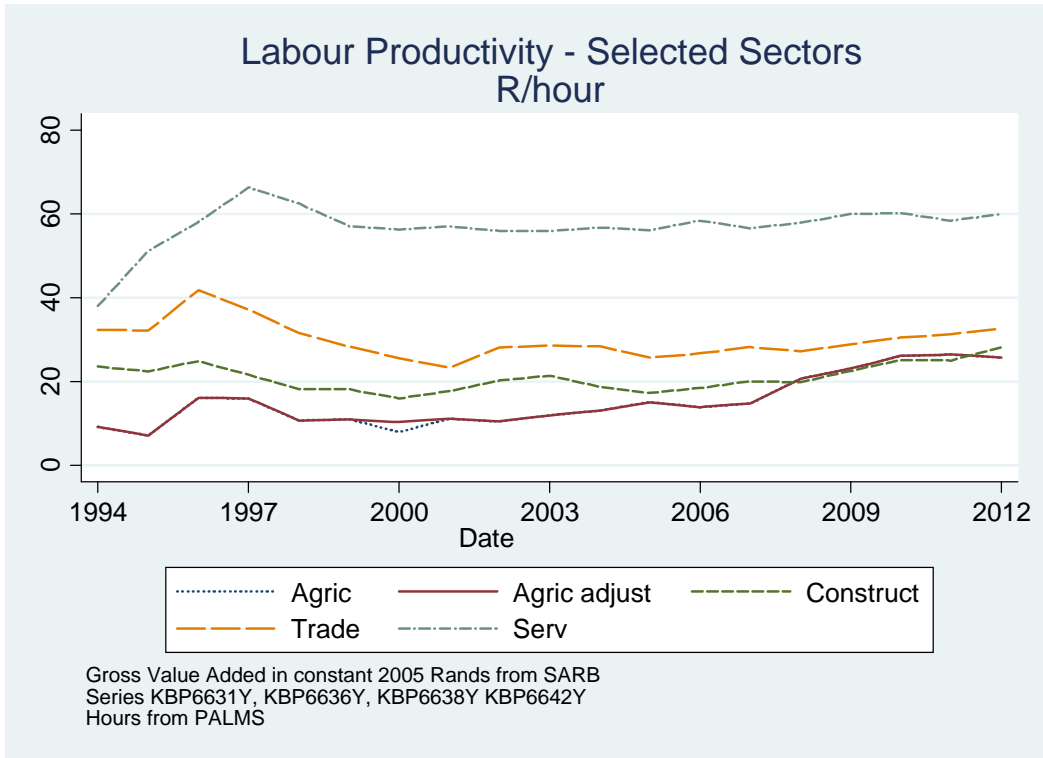


Figure 50. Labour productivity measured in output per hour worked by sector, Part 1

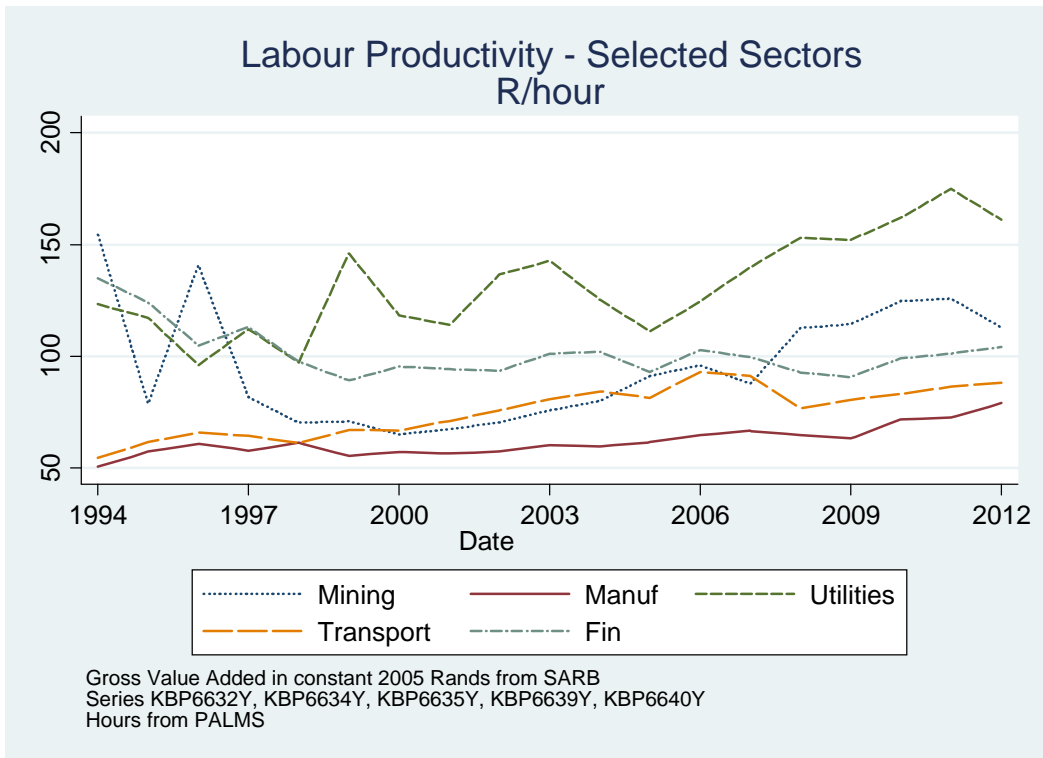


Figure 51. Labour productivity measured in output per hour worked by sector, Part 2

7.4.3 Comparing the productivity measures

It is clear that the two measures of labour productivity provide qualitatively similar information. We compare them directly by converting both to index form (with the year 2000 = 100). In both cases we have used the series stripping out the household sector and smoothing over the spike in 2000. The resulting aggregate trends are shown in Figure 52. The two series track each other very closely until 2005, when productivity measured in terms of hours worked improves more rapidly than when measured in terms of employment. This is evidently due to the long-term decline in average hours worked shown in Figure 31.

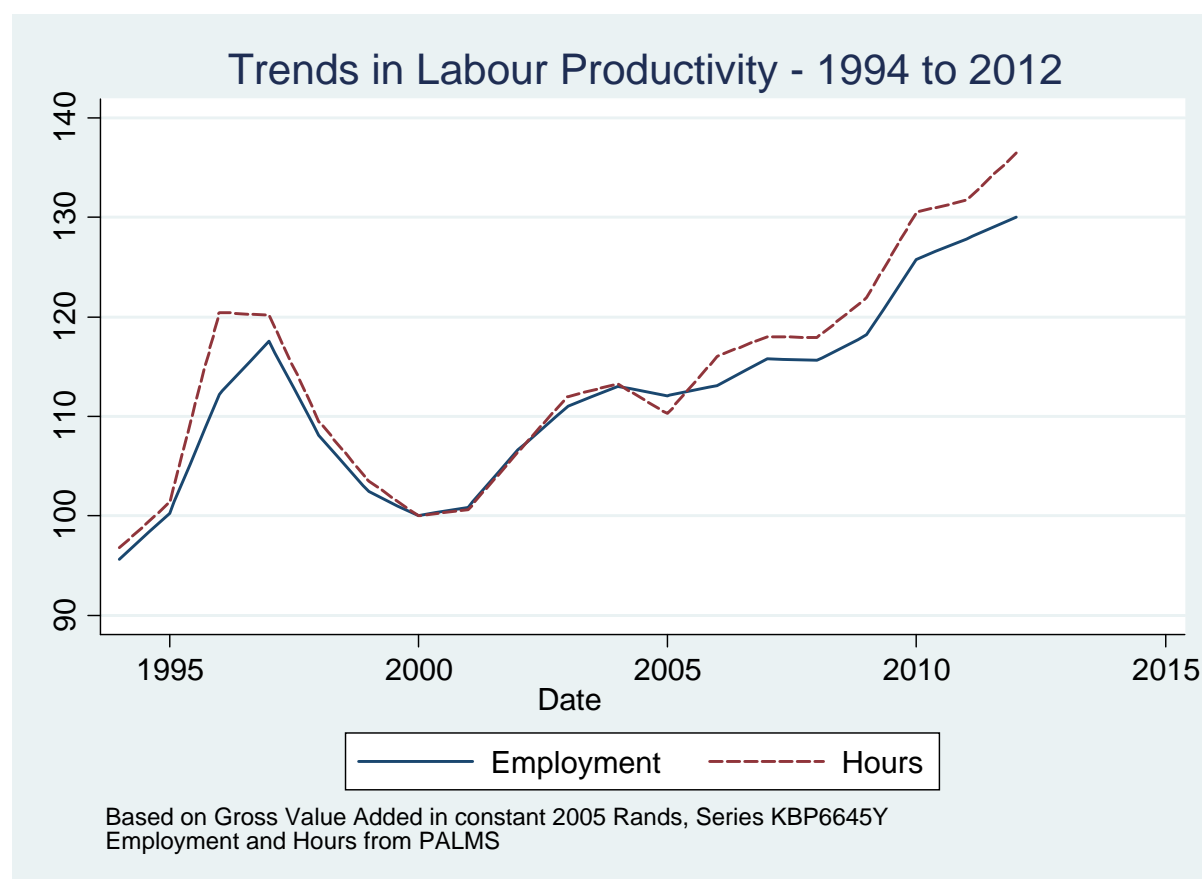


Figure 52. Comparing labour productivity trends when using hours worked or employment

7.5 The relationship between earnings and productivity

In one last comparison we construct indexes of earnings and earnings per hour (for the non-agricultural non-domestic sectors) and compare these to the labour productivity indexes constructed in the previous section. Figure 53 provides the evidence. There is little in these graphs to suggest that either earnings growth has outstripped productivity or *vice versa*.

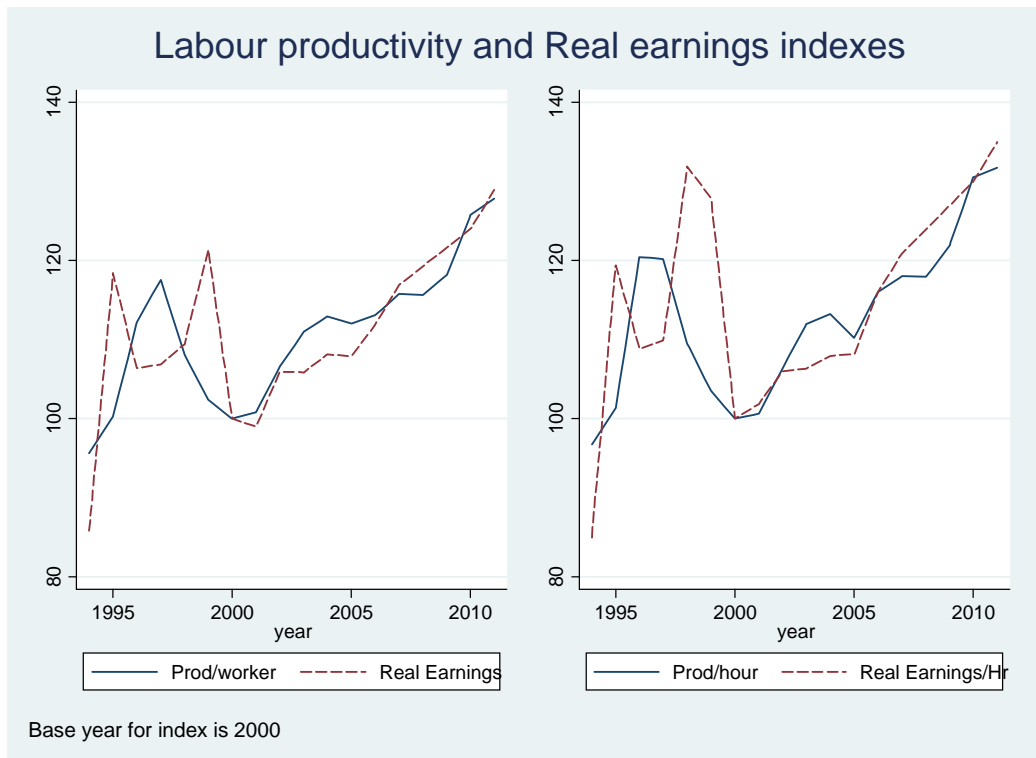


Figure 53. The evolution of productivity and real earnings, 1994 to 2011

8. Conclusion

Looking at long run trends in the South African labour market is bedevilled by breaks in the series – both in the microdata series emanating from household surveys as well as the firm level series that underpins the macro data distributed by the SA Reserve Bank and international agencies.

In the work underpinning this report, we have taken considerable care to highlight the discontinuities and to take them into consideration when discussing the long-run trends. Several clear patterns emerge through the murk:

- The “collapse” of earnings from the 1990s to the early 2000s is almost definitely an artefact of changing survey practice in the late 1990s and, in particular, a much greater drive to enumerate informal sector activities in the LFSs.
- While real earnings do not seem to have collapsed, median earnings have remained fairly static over the entire period. There seems to have been a widening of the distribution at the “top” and a compression at the bottom of the wage worker distribution.
- The gap between the median and the mean seems to have increased in the case of employees.
- The difference between the QES and QLFS average earnings data is hard to reconcile without positing large errors in coverage (missing high income earners), classification (employers/employees) or measurement (before/after tax or erroneous inclusion of non-employee costs in gross earnings). A different source of data would be useful to cross-check both series. Even allowing for sizable errors, it seems clear that the QES average earning figure of more than R12 000 per month is, in fact, a very poor reflection of conditions for the median worker even in the non-agricultural non-domestic sector.

-
- Labour productivity has increased in aggregate by around 30% since 2000. The aggregate trend, however, masks what appears to be a divergence between lower productivity sectors which have shown lower productivity growth and higher productivity sectors which also seem to have had stronger growth.
 - Looking over the entire period there is no strong evidence that average wages have outstripped productivity or *vice versa*

What are the implications of these findings for the policy debates? More attention needs to be paid to the factual basis on which arguments are advanced. For instance, it is not helpful to pick average wages from the QES but calculate poverty rates off the QLFS. Equally it is not helpful to dismiss claims about continuing poverty on the basis that average wages have risen. Indeed one of the key findings of this paper is that averages can conceal huge differences in experiences. This is true even of the finding that average productivity has increased. Indeed as we show the sectoral patterns are diverse. This would be even more pronounced at more disaggregated levels. At present we lack the data that would help us to do a finer grained analysis.

One of the key problems is that while we have a lot of household survey data, the firm microdata are not generally available for research. Without that we cannot complete the picture or hope to get a better insight into why the firm and the household data seem to diverge.

Besides better data it is also necessary to do some of the analytic work that goes beyond simply graphing trends against each other. Accurate descriptions of trends do not establish causal connections. More careful econometric work that tries to model some of the relationships is obviously necessary.

Appendix: Additional Statistics by Gender and Race

In order not to overburden the main text with too many results, some of the additional analyses done in terms of gender and race are reported here.

A.1 Employment by Gender and Race

In Figure 54 it is striking the extent to which the shift up in employment has a disproportionately strong influence on women's employment, with over a million extra female workers located between October 1999 and February 2000. The anomalous employment figures in 1995, by contrast, seem more strongly accentuated for men.

The racial breakdown shown in Figure 55 shows that the change is most dramatic for Africans (left hand axis). We can also see an increase among Coloureds, but in this picture not much evidence of change among Whites and Indians (all on the right hand axis). Again 1995 employment emerges as odd, when compared with other OHSs. There is little support in this figure for falling employment levels in the late 1990s which were of concern in some circles, unless 1995 is used as the anchor year for such comparisons.

A closer look at the numbers of self-employed by race (in Figure 56) shows that in fact the LFS found higher levels of self-employment across the board. Even among Whites there are around an extra 100 000 self-employed between October 1999 and February 2000, which is around a 30% increase concentrated in less than half a year.

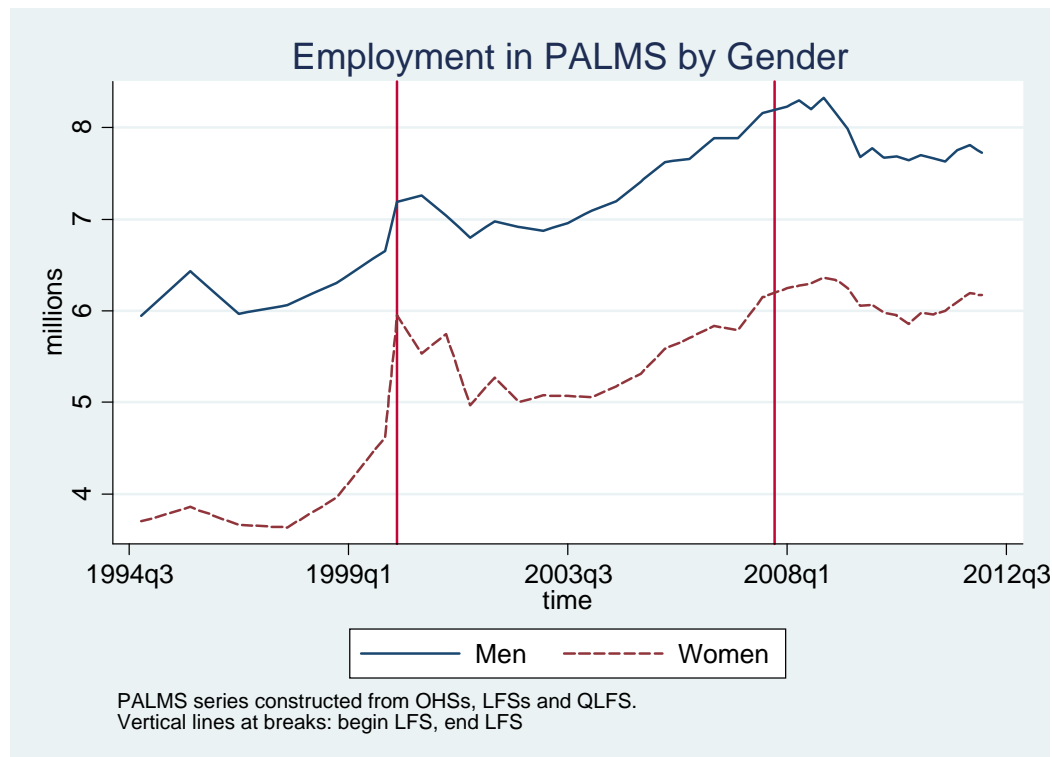


Figure 54. The shift up with the LFS is particularly noticeable among women

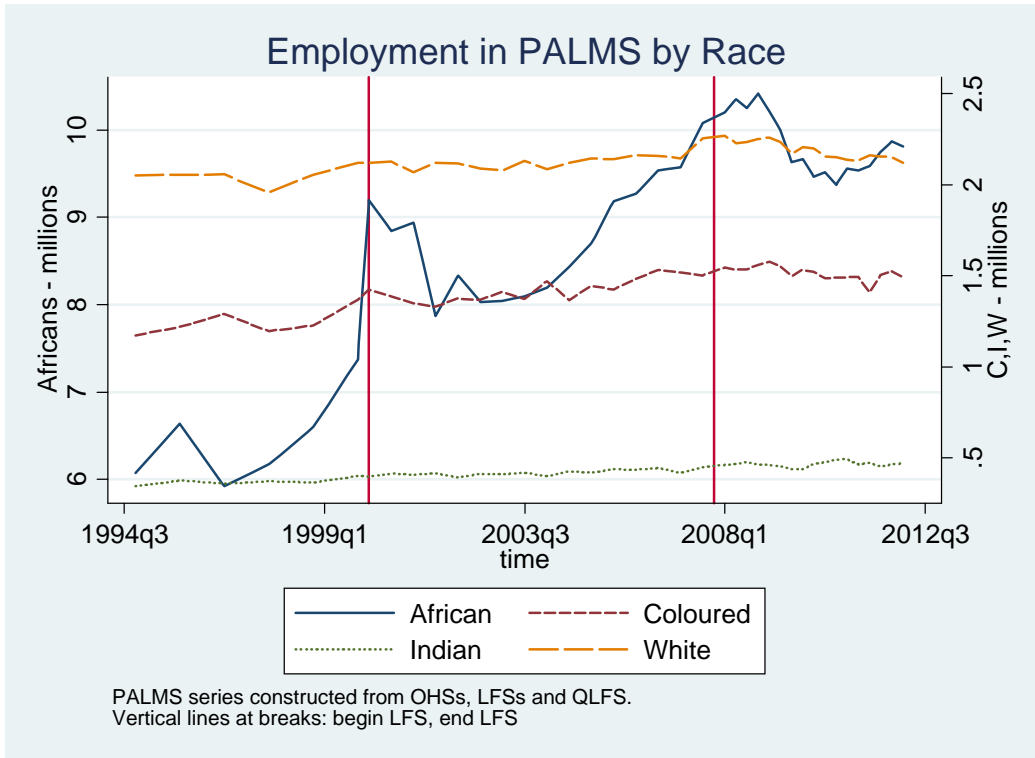


Figure 55. The shift up is most dramatic for Africans (left axis) but less noticeable among other groups (right axis)

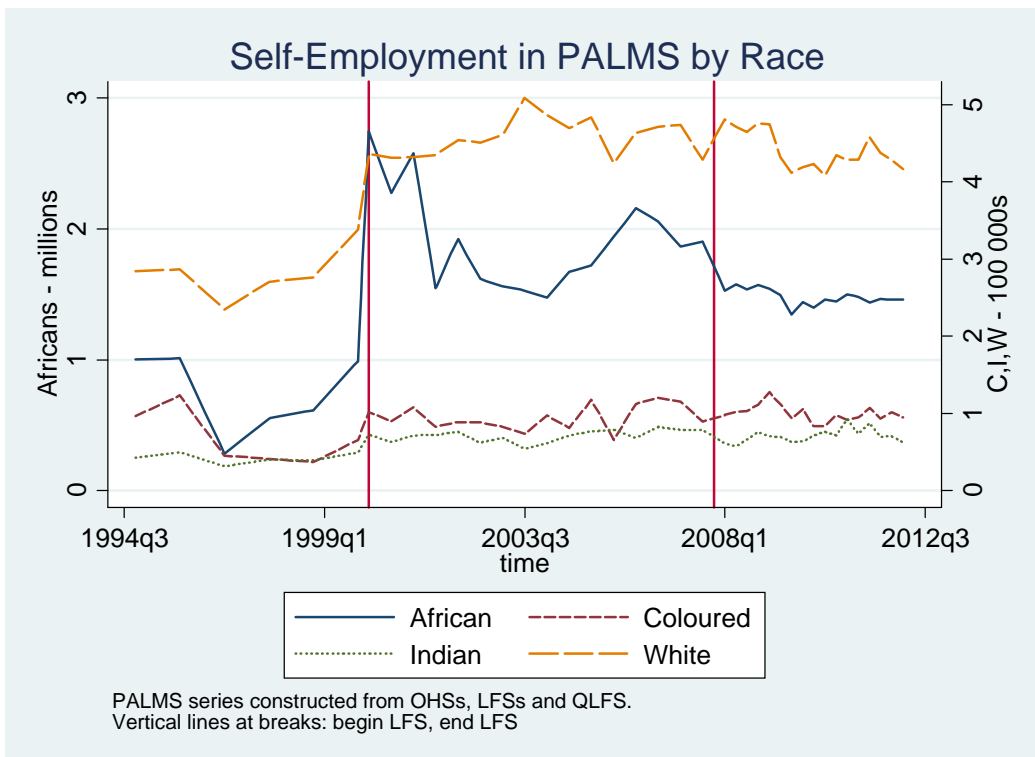


Figure 56. Self-employment levels by Race

A.2 Average Hours by Gender and Race

The average hours worked by gender and population group is shown in Figure 57 and Figure 58. It is evident that men work on average around 4 hours more per week than women. This is likely to be the result of more part-time work done by women. The picture by population group (Figure 58) is not so clear. In the QLFSs it appears that Africans work longer hours on average than Whites, or indeed Coloureds and Indians. The picture is more ambiguous for the period as a whole.

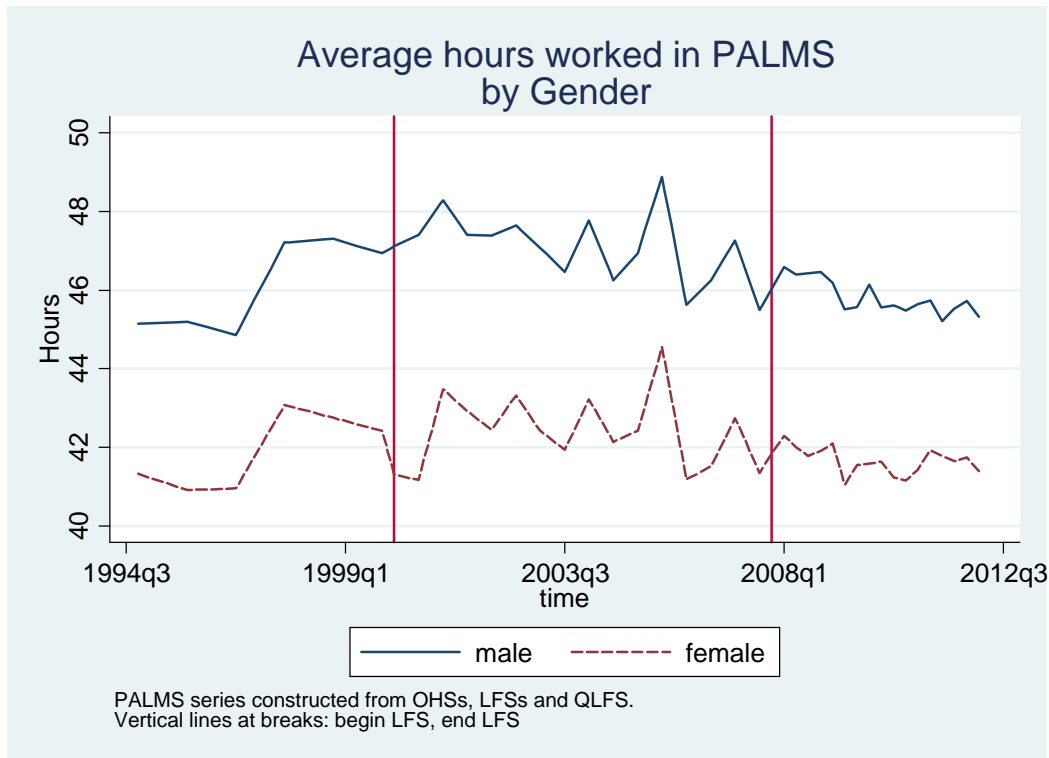


Figure 57. Average hours worked by gender

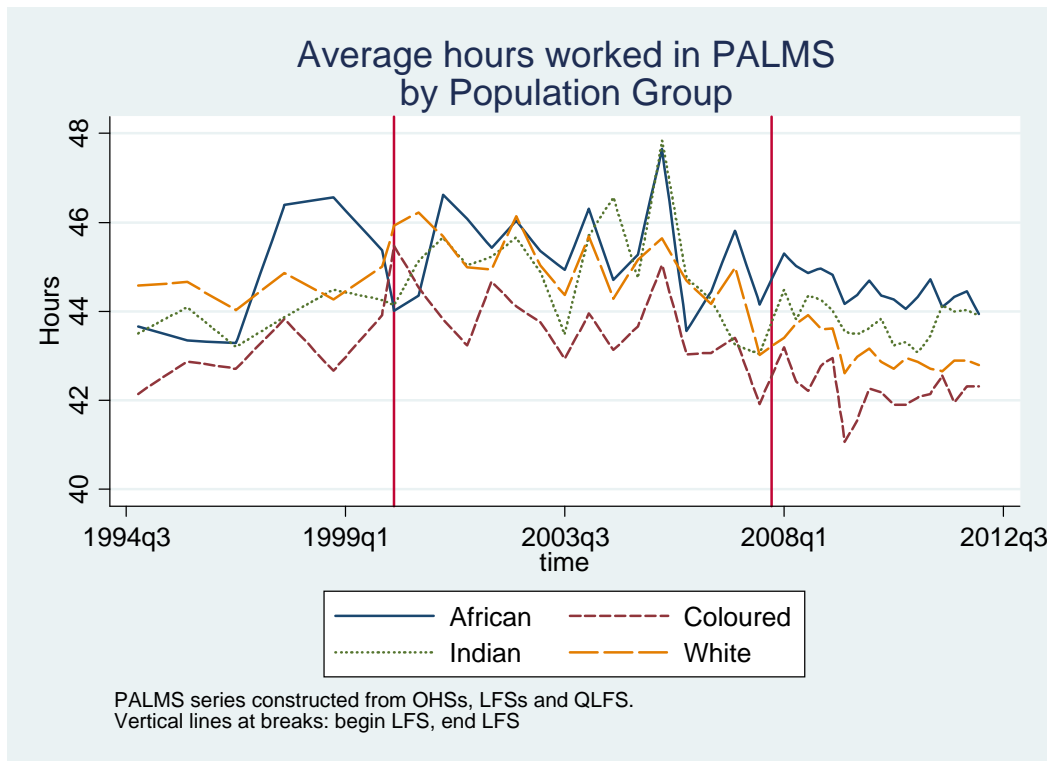


Figure 58. Average hours worked, by Population Group

A.3 Earnings per hour by Gender and Race

Figure 59 shows that the gender gap in earnings remains once we control for hours. The corresponding graph by race is shown in Figure 60. Again the qualitative picture does not change from that given before.

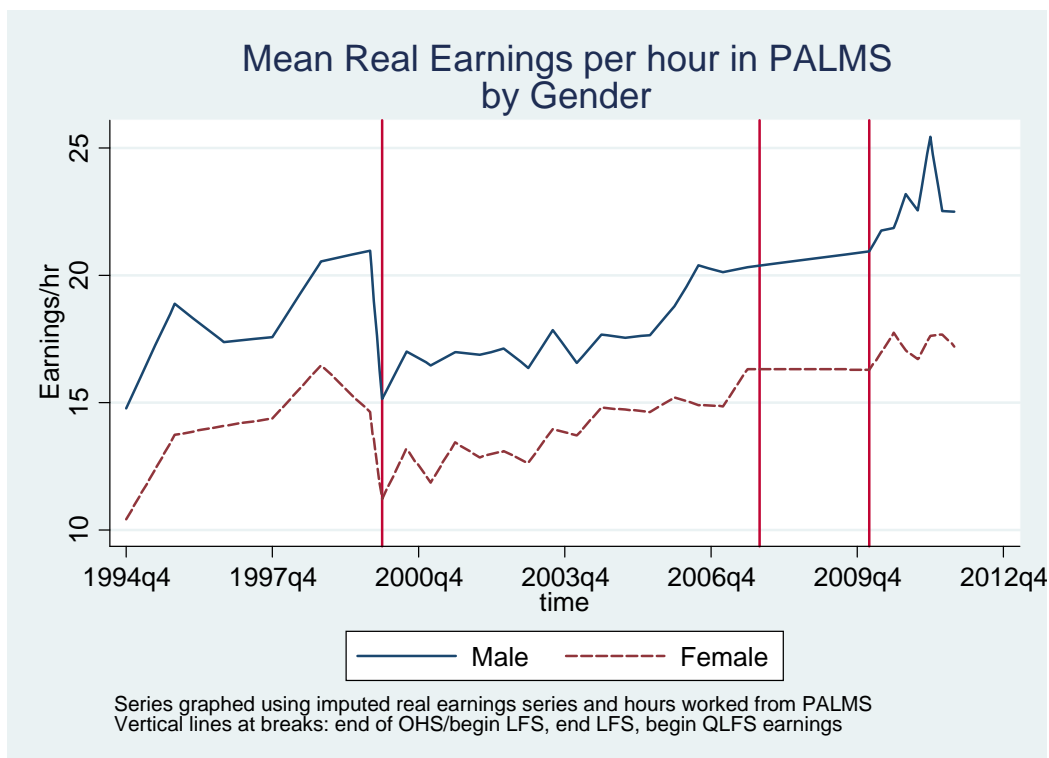


Figure 59. Real earnings per hour 1994-2011, by gender

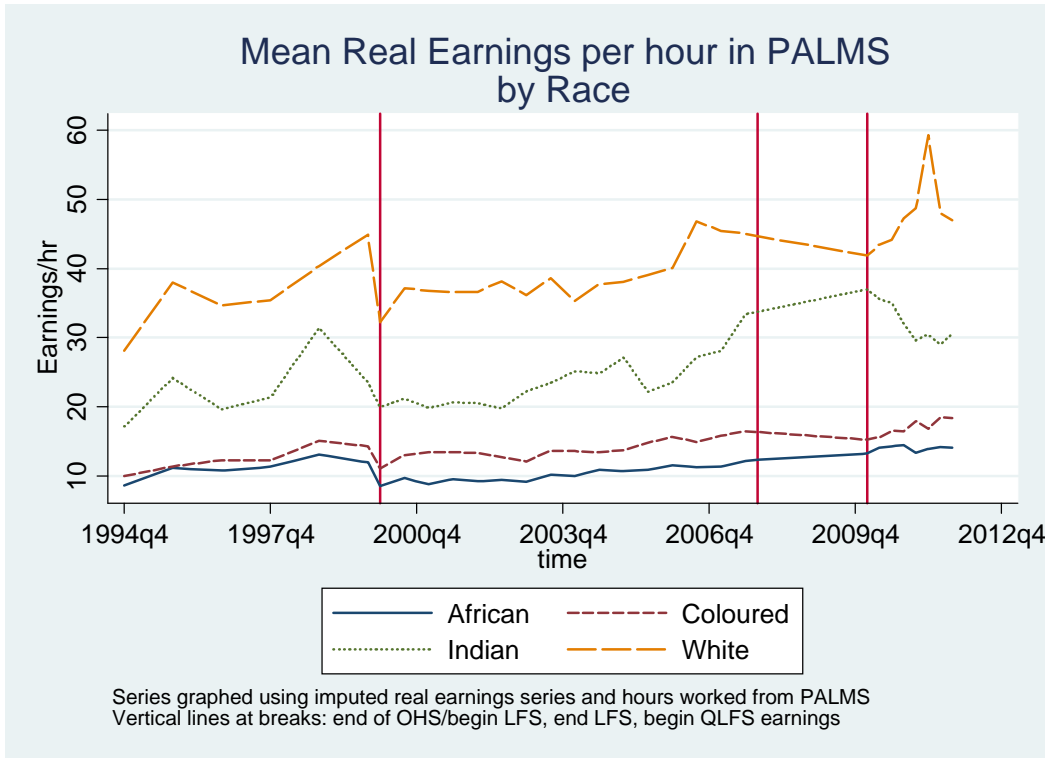


Figure 60. Real earnings per hour 1994-2011, by Race

References

- Branson, Nicola and Martin Wittenberg (forthcoming), “Re-weighting South African National Household Survey Data to create a consistent series over time: A cross entropy estimation approach”, *South African Journal of Economics*
- Burger, Rulof and Derek Yu (2007), “Wage trends in post-Apartheid South Africa: Constructing an earnings series from household survey data”, Working Paper 07/117, Development Policy Research Unit, University of Cape Town.
- Casale, Daniela (2004), “What has the feminisation of the labour market ‘bought’ women in South Africa? Trends in labour force participation, employment and earnings, 1995-2001”, *The Journal of Interdisciplinary Economics* 15, 251-275.
- Daniels, Reza (2013), “Questionnaire design and response propensities for employee income microdata”, Working Paper 89, SALDRU, University of Cape Town. Revised version. Retrieved from www.saldru.uct.ac.za on 18 April 2013.
- Heap, Anthea (2008), “Earnings Inequality in South Africa: Decomposing Changes Between 1995 and 2006”, unpublished Master’s dissertation, School of Economics, University of Cape Town.
- Kerr, Andrew and Martin Wittenberg (2013a), “Sampling Methodology and Field Work Changes in the October Household Surveys and Labour Force Surveys”, Working paper, DataFirst, University of Cape Town.
- (2013b), “Job Creation and Destruction in South Africa”, SALDRU/DataFirst Working Paper number 92.
- Klein, Nir (2012), “Real Wage, Labor Productivity, and Employment Trends in South Africa: A Closer Look”, *IMF Working Paper*, WP/12/92, International Monetary Fund.
- Mazumdar, Dipak and Dirk van Seventer (2002), “A Decomposition of Growth of the Real Wage Rate for South Africa: 1970-2000”, *South African Journal of Economics*, 70(6):1076-1102.
- Rubin, D. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.
- South African Reserve Bank (2010), *South Africa’s national accounts 1946 – 2009: An overview of sources and methods*, Supplement to the South African Reserve Bank Quarterly Bulletin March 2010, Pretoria: SARB.
- StataCorp (2013), *Stata 13 Multiple Imputation Reference Manual*, College Station, TX: Stata Press.
- Statistics South Africa (2009), “Gross Domestic Product: Annual estimates 1993 – 2008, Regional estimates 2000 – 2008, Third quarter: 2009”, Statistical Release P0441 (3rd quarter 2009), Pretoria: Statistics South Africa.
- (2011), “Annual Financial Statistics”, Statistical Release P0021, Pretoria: Statistics South Africa.
- (2012), “Gross domestic product: Annual estimates 2002 – 2011, Regional estimates 2002 – 2011, Third quarter 2012”, Statistical Release P0441 (3rd quarter 2012), Pretoria: Statistics South Africa.
- Vermaak, Clare (2012), “Tracking poverty with coarse data: evidence from South Africa”, *Journal of Economic Inequality*, 10:239–265.
- Wittenberg, Martin (2004), “The Mystery of South Africa’s Ghost Workers In 1996: Measurement And Mismeasurement in the Manufacturing Census, Population Census And October Household Surveys”, *South African Journal of Economics*, 72(5):1003-1022.
- (2008), “Income in the October Household Survey 1994”, Working paper, DataFirst, University of Cape Town.

-
- (forthcoming), “Data issues in South Africa”, in H. Bhorat, A. Hirsch, R. Kanbur and M. Ncube (eds), *Oxford Companion to the Economics of South Africa*, Oxford University Press.
- and Farah Pirouz (2013), “The measurement of earnings in the post-Apartheid period: An overview”, DataFirst Working Paper, University of Cape Town.
- Yu, Derek (2007), “The comparability of the Statistics South Africa October Household Surveys and Labour Force Surveys”, Stellenbosch Economic Working Papers 17/07, University of Stellenbosch, Economics Department. Available from www.ekon.sun.ac.za/wpapers/2007/wp172007/wp-17-2007.pdf, downloaded 8 June 2013
- (2009),” The comparability of Labour Force Survey (LFS) and Quarterly Labour Force Survey (QLFS)”, Stellenbosch Economic Working Papers 08/09, University of Stellenbosch, Economics Department. Available from <http://www.ekon.sun.ac.za/wpapers/2009/wp082009/wp-08-2009.pdf>, downloaded 8 June 2013.

Datasets

- Kerr, Andrew, David Lam and Martin Wittenberg (2013), Post-Apartheid Labour Market Series 1994-2012. [dataset]. Version 2.1. of the harmonised dataset based on Statistics SA's OHS, LFS and QLFS surveys, 1994-2012. Cape Town: DataFirst [data producer and distributor].

Conditions of Work and Employment Series

- No. 1 Quality of working life: A review on changes in work organization, conditions of employment and work-life arrangements (2003), by H. Gospel
- No. 2 Sexual harassment at work: A review of preventive measures (2005), by D. McCann
- No. 3 Statistics on working time arrangements based on time-use survey data (2003), by A. S. Harvey, J. Gershuny, K. Fisher & A. Akbari
- No. 4 The definition, classification and measurement of working time arrangements (2003), by D. Bell & P. Elias
- No. 5 Reconciling work and family: Issues and policies in Japan (2003), by M. Abe, C. Hamamoto & S. Tanaka
- No. 6 Reconciling work and family: Issues and policies in the Republic of Korea (2004), by T.H. Kim & K.K. Kim
- No. 7 Domestic work, conditions of work and employment: A legal perspective (2003), by J.M. Ramirez-Machado
- No. 8 Reconciling work and family: Issues and policies in Brazil (2004), by B. Sorj
- No. 9 Employment conditions in an ageing world: Meeting the working time challenge (2004), by A. Jolivet & S. Lee
- No. 10 Designing programmes to improve working and employment conditions in the informal economy: A literature review (2004), by Dr. R.D. Rinehart
- No. 11 Working time in transition: The dual task of standardization and flexibilization in China (2005), by X. Zeng, L. Lu & S.U. Idris
- No. 12 Compressed working weeks (2006), by P. Tucker
- No. 13 Étude sur les temps de travail et l'organisation du travail: Le cas du Sénégal. Analyse juridique et enquête auprès des entreprises (2006), by A. Ndiaye
- No. 14 Reconciling work and family: Issues and policies in Thailand (2006), by K. Kusakabe
- No. 15 Conditions of work and employment for older workers in industrialized countries: Understanding the issues (2006), by N.S. Ghosheh Jr., S. Lee & D. McCann
- No. 16 Wage fixing in the informal economy: Evidence from Brazil, India, Indonesia and South Africa (2006) by C. Saget
- No. 18 Reconciling work and family: Issues and policies in Trinidad and Tobago (2008), by R. Reddock & Y. Bobb-Smith
- No. 19 Minding the gaps: Non-regular employment and labour market segmentation in the Republic of Korea (2007) by B.H. Lee & S. Lee

-
- No. 20 Age discrimination and older workers: Theory and legislation in comparative context (2008), by N. Ghosheh
- No. 21 Labour market regulation: Motives, measures, effects (2009), by G. Bertola
- No. 22 Reconciling work and family: Issues and policies in China (2009), by Liu Bohong, Zhang Yongying & Li Yani
- No. 23 Domestic work and domestic workers in Ghana: An overview of the legal regime and practice (2009), by D. Tsikata
- No. 24 A comparison of public and private sector earnings in Jordan (2010), by C. Dougherty
- No. 25 The German work-sharing scheme: An instrument for the crisis (2010), by A. Crimmann, F. Weissner, L. Bellmann
- No. 26 Extending the coverage of minimum wages in India: Simulations from household data (2010), by P. Belser & U. Rani
- No. 27 The legal regulation of working time in domestic work (2010), by Deirdre Mc Cann & Jill Murray
- No. 28 What do we know about low-wage work and low-wage workers (2011), by Damian Grimshaw
- No. 29 Estimating a living wage: a methodological review (2011), by Richard Anker
- No. 30 Measuring the economic and social value of domestic work: conceptual and methodological framework (2011), by Debbie Budlender
- No. 31 Working Time, Health, and Safety: a Research Synthesis Paper (2012), prepared by Philip Tucker and Simon Folkard, on behalf of Simon Folkard Associates Ltd
- No. 32 The influence of working time arrangements on work-life integration or ‘balance’: A review of the international evidence (2012), by Colette Fagan, Clare Lyonette, Mark Smith and Abril Saldaña-Tejeda
- No. 33 The Effects of Working Time on Productivity and Firm Performance: a research synthesis paper (2012), by Lonnie Golden
- No. 34 Estudio sobre trabajo doméstico en Uruguay (2012), by Dra. Karina Batthyány
- No. 35 Why have wage shares fallen? A panel analysis of the determinants of functional income distribution (2012), by Engelbert Stockhammer
- No. 36 Wage-led or Profit-led Supply: Wages, Productivity and Investment (2012), by Servaas Storm & C.W.M. Naastepad
- No. 37 Financialisation and the requirements and potentials for wage-led recovery – a review focussing on the G20 (2012), by Eckhard Hein and Matthias Mundt
- No. 38 Wage Protection Legislation in Africa (2012), by Najati Ghosheh
- No. 39 Income inequality as a cause of the Great Recession? A survey of current debates (2012), by Simon Sturn & Till van Treeck

-
- No. 40 Is aggregate demand wage-led or profit-led? National and global effects (2012), by Özlem Onaran & Giorgos Galanis
- No. 41 Wage-led growth: Concept, theories and policies (2012), by Marc Lavoie & Engelbert Stockhammer
- No. 42 The visible face of Women's invisible labour: domestic workers in Turkey (2013), by Seyhan Erdoğan and Gülay Toksöz
- No. 43 In search of good quality part-time employment (2013), by Colette Fagan, Helen Norman, Mark Smith and María C. González Menéndez
- No. 44 The use of working time-related crisis response measures during the Great Recession (2013), by Angelika Kümmerling and Steffen Lehndorff