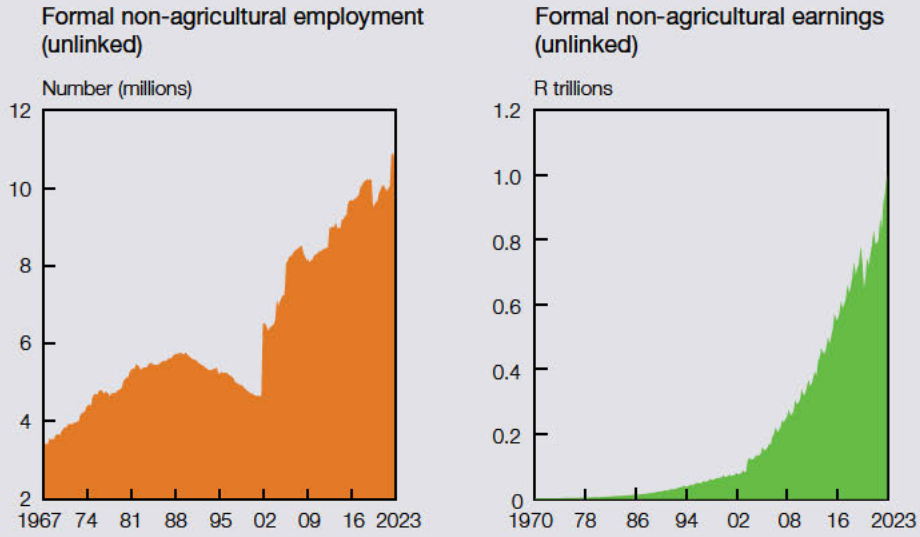
Box 1 Comparing the linking of enterprise-surveyed labour market statistics by Statistics South Africa and the South African Reserve Bank

Statistics South Africa (Stats SA), in line with international best practice, periodically updates its business register to reflect recent economic developments, incorporating improved coverage, such as the addition of new enterprises, the removal of dormant enterprises as well as reclassifications.

Stats SA's business register is the primary sampling frame from which all their economic sample surveys, including the enterprise-based *Quarterly Employment Statistics (QES)* survey, are drawn. The business register is a database containing information on all the businesses registered for income tax and value-added tax (VAT) with the South African Revenue Service (SARS).

These sample refreshes have caused several level shifts[1] in the *QES* survey statistics, resulting in structural breaks in the time series measuring formal non-agricultural employment and gross earnings.

1 Level shifts occurred in the third quarter of 2002, the fourth quarter of 2004, the second quarter of 2006 and the second quarter of 2013. From 2015 onwards, Stats SA began to draw annual samples, which resulted in structural breaks occurring in the second quarter of each successive year, except in 2020 and 2022, with the most recent one in the second quarter of 2023.



Not seasonally adjusted
Source: Stats SA

The analysis of economic indicators such as labour market statistics requires long time series, and structural breaks should be prevented where possible. The standard approach to preserving time-series continuity and consistency following a major change to a survey is to create linkages where the series breaks.[2] It is typically recommended that statistical agencies conduct parallel surveys for a few overlapping periods to enable data users to assess the impact of a survey change and to facilitate appropriate adjustments to the historical time series. However, there are no coordinated international guidelines or best practice on how to address structural breaks and the linking of time series. This increases the risk of varying linking methodologies, thus leading to different versions of the same time series being produced by different users and secondary data providers.[3]

Both Stats SA and the South African Reserve Bank (SARB)[4] statistically link the *QES* survey statistics to remove the structural breaks caused by the sample adjustments to ensure long and continuous time series of formal non-agricultural employment and gross earnings for the purposes of macroeconomic analysis and policy formulation. Both institutions' linking methodologies are based on a linking factor, calculated at the overlapping quarter where the series break occurs, which is used to adjust the data based on the old sample frame to the level based on the new sample frame.

The main difference between the linking methodologies of Stats SA and the SARB relates to the period over which the linking factor is applied. The SARB applies the linking factor over the entire history of the time series, at each structural break, while Stats SA applies the linking factor only over the period since the previous structural break – which is similar to the method applied by the Bureau of Labor Statistics (BLS) in the United States (US).[5] However, the BLS method seems to be more appropriate when the level shifts have historically been fairly small, which is not the case for the *QES* survey, as the continuous process of improving the coverage of the survey due to an evolving business register has resulted in fairly large structural breaks. According to the BLS, their average annual benchmark revisions (level shifts) have historically been very small for total non-farming employment, ranging from -0.7% to 0.3%. By contrast, the structural breaks in the *QES* survey's formal non-agricultural employment have ranged from 5.9% in 2013 to 0.5% in 2019, with the most recent one amounting to 7.0% in 2023.

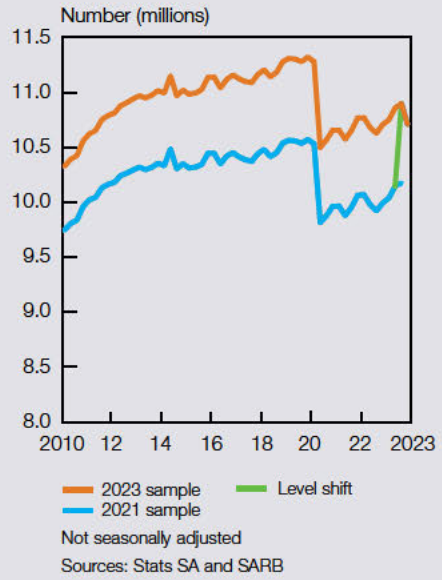
2 United States Census Bureau, 'The Impact of Classification Revisions on Time Series', *Economic Classification Policy Committee: Issues Paper No. 5*, Washington, D.C.: United States Census Bureau, 1993. https://www.census.gov/eos/www/naics/history/docs/issue_paper_5.pdf

3 R McKenzie, 'Country comments on OECD paper: Review of linking practices for the index of industrial production in OECD countries', paper presented at the Organisation for Economic Co-operation and Development (OECD) Short-term Economic Statistics Working Party (STESWP), Paris, 26–28 June 2006. https://www.oecd.org/sdd/fin-stats/36987634.pdf

4 See the box titled 'Statistical linking of formal non-agricultural employment and earnings time series' published in the March 2017 edition of the *Quarterly Bulletin* (*QB*), available at https://www.resbank.co.za/en/home/publications/publication-detail-pages/quarterly-bulletins/boxes/2017/7743.

5 C D Manning and J R Stewart, 'Benchmarking the Current Employment Statistics national estimates', *Monthly Labor Review,* October 2017, pp 1–8. https://www.jstor.org/stable/90015087

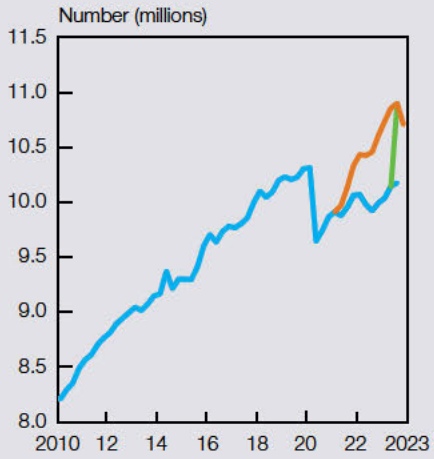
SARB linked formal non-agricultural employment



Not seasonally adjusted

Sources: Stats SA and SARB

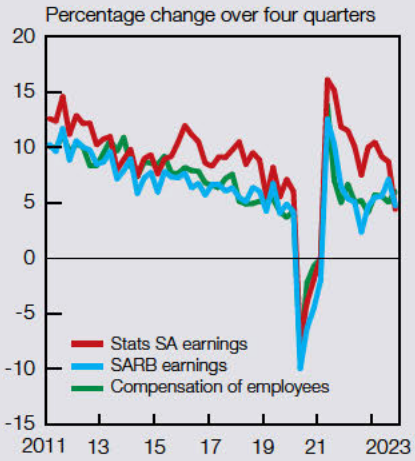
Stats SA linked formal non-agricultural employment



The accompanying graphs show the linked total formal non-agricultural employment time series by both Stats SA and the SARB following the most recent sample change in the second quarter of 2023. This sample change resulted in an upward adjustment of 714 900 jobs in the level of employment and R84 million in gross nominal earnings.

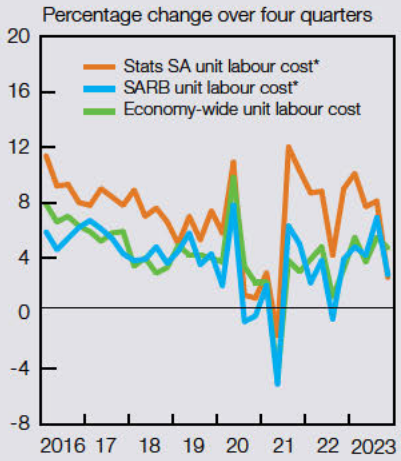
Ideally, the most suitable linking method would ensure that the cyclical movement, seasonal pattern and long-term trend of the original time series are retained in the linked series, while also ensuring minimal revisions to historical growth rates. This is, however, not possible in practice, and the choice of a linking method usually involves a trade-off between these objectives.

Formal non-agricultural earnings and compensation of employees



Sources: Stats SA and SARB

Nominal unit labour cost



* Formal non-agricultural sector

By applying the linking factor over the full history of the time series, the SARB implicitly assumes that only a small portion of the level shift resulted from growth in employment and earnings, with the bulk resulting from the improved quality of the business register and not new economic activity. While the SARB's linking methodology achieves most of the ideal linking objectives, it still results in a marginally steeper long-term trend in the linked time series. Despite this, a comparison between the year-on-year growth rates in formal non-agricultural earnings derived from the SARB's linked series and those calculated from the broader national accounts measure of 'compensation of employees' shows that these two indicators have tracked each other closely over time. The growth rates in both measures are, however, well below those in Stats SA's linked earnings time series.

The SARB's linked statistics are also used to derive two important labour market indicators, namely labour productivity and unit labour cost (ULC) in the formal non-agricultural sector. Growth in the SARB's linked formal non-agricultural nominal ULC and growth in the economy-wide ULC, calculated with the compensation of employees, have historically correlated fairly well, with growth in both measures well below growth in formal non-agricultural nominal ULC calculated with Stats SA's linked earnings time series.

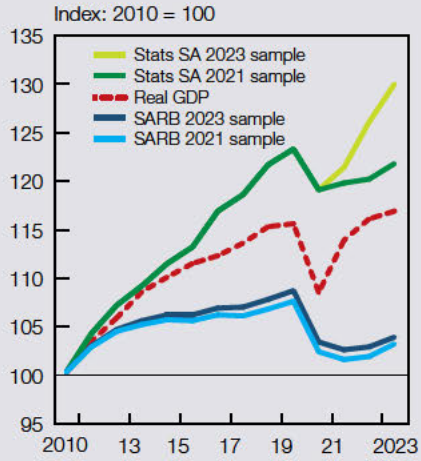
Formal non-agricultural earnings and compensation of employees

	Percentage change over four quarters		
	Stats SA's linked earnings	SARB's linked earnings	Compensation of employees
March 2020	6.0	4.1	4.2
June 2020	-7.5	-10.0	-8.7
September 2020	-4.0	-6.3	-2.2
December 2020	-1.9	-4.5	-0.7
March 2021	0.1	-2.0	-0.1
June 2021	16.1	12.5	13.8
September 2021	15.1	10.3	7.0
December 2021	11.8	6.5	5.0
March 2022	11.5	5.3	6.7
June 2022	10.1	5.1	5.1
September 2022	7.5	2.4	5.2
December 2022	10.0	4.7	4.1
March 2023	10.5	5.5	5.7
June 2023	9.1	5.6	5.6
September 2023	8.7	7.1	5.0
December 2023	4.4	4.7	6.0

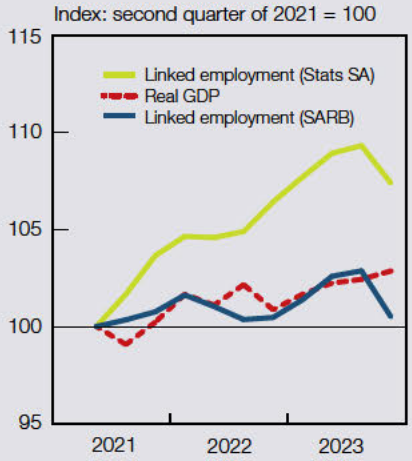
Source: Stats SA

By applying the linking factor starting only from the previous sample change, Stats SA's linking methodology results in a much steeper long-term trend since the previous sample change, as it attributes the whole level shift to growth in employment and earnings without considering any qualitative improvements in the business register. This essentially builds an upward bias into the linked series, especially given the significant level shifts, which tends to overstate the growth rates in the time series since the previous sample change and distorts the cyclical movements.

Annual formal non-agricultural employment and output



Quarterly formal non-agricultural employment and output



Sources: Stats SA and SARB

The accompanying graphs depict the annual average level of real gross domestic product (GDP) together with the linked formal non-agricultural employment time series by both Stats SA and the SARB, based on the 2021 and 2023 sample frames. According to Stats SA's linked time series, formal non-agricultural employment generally increased at a much faster pace than real GDP. By contrast, the SARB's linked employment time series increased at a similar pace as real GDP over the past two years, from when the previous sample change occurred. The SARB's linked employment time series shows that 171 000 formal non-agricultural jobs were created between the second quarter of 2021 and the first quarter of 2023, compared to a significant increase of 786 600 jobs over the same period according to Stats SA's linked time series. This seems highly unlikely given the weak economic growth over this period, as the average quarterly growth rate in real GDP amounted to only 0.3% over this period.

In addition, following the pandemic-related distortions in 2020 and 2021, year-on-year growth in formal non-agricultural nominal earnings recorded double-digit rates, ranging between 11.5% in the first quarter of 2022 and 10.5% in the first quarter of 2023, according to Stats SA's linked earnings time series. By contrast, growth in nominal earnings was between 5.3% and 5.5% over the same period according to the SARB's linked series, which is in line with the 4.7–5.9% range in the compensation of employees measure.

Changes in the business register are typically reflected in the differences between the frames as there is much variation in the sampled units, which contributes to the size of the level shifts caused by the sample adjustments. Objectively, random samples are drawn for smaller and medium enterprises, while large enterprises are completely enumerated. The level shifts are therefore most likely a reflection of qualitative differences resulting from the updated business register rather than actual growth in employment and gross earnings, as Stats SA's linked time series does not seem to align well with other macroeconomic indicators.