

**AUTHORS:**Tom A. Moultrie<sup>1</sup>   
Rob E. Dorrington<sup>1</sup> **AFFILIATION:**<sup>1</sup>Centre for Actuarial Research,  
Faculty of Commerce, University of  
Cape Town, Cape Town, South Africa**CORRESPONDENCE TO:**

Tom Moultrie

**EMAIL:**

tom.moultrie@uct.ac.za

**HOW TO CITE:**Moultrie TA, Dorrington RE. Problems  
and concerns with the 2022  
South African census. *S Afr J Sci.*  
2024;120(7/8), Art. #18585. <https://doi.org/10.17159/sajs.2024/18585>**ARTICLE INCLUDES:**

- 
- Peer review
- 
- 
- Supplementary material

**KEYWORDS:**census, demography, South Africa,  
migration, age-sex distribution**PUBLISHED:**

31 July 2024



# Problems and concerns with the 2022 South African census

**Significance:**

A census provides a snapshot of a population at a point in time. In addition to describing the basic demographic and socio-economic characteristics of a population, census data provide denominators for estimating rates; inform resource allocation and policy, development and infrastructural investment; and are used to define sampling frames for other national inquiries. Although the 2022 South African census attempted to correct for a national undercount of over 30%, it was not entirely successful. We identify a number of concerns with the estimates released, which suggest that the data may not be fit for purpose.

## Introduction

The 2022 South African census, released on 10 October 2023, estimated the population as 62.03 million in February 2022.<sup>1</sup> The Post-Enumeration Survey (PES) report<sup>2</sup> revealed the undercount to be 31% (and 72% of the Indian and 62% of the white population groups) – the highest reported by any country and more than double that of the previous census.

An undercount of this magnitude is consequential – not only in terms of the greater uncertainty surrounding estimates of the population size, but also because of the implications for the reliability of estimates for smaller populations, particularly those defined at a granularity finer than the dimensions used to stratify the PES.

The estimated size of the population surprised demographers, both locally and internationally, particularly taking into account the assumed impact of the COVID-19 pandemic on both mortality and migration. Prior to the release of the census, the estimates of the size of the South African population at the date of the 2022 census, based on population projections from a variety of credible sources, ranged from 57.2 million<sup>3</sup> to 60.3 million (Stats SA's mid-year population estimates (MYPE)<sup>4</sup>).

Even under near-ideal conditions, running a national census is a costly, complex, and logistically demanding undertaking. The US National Academies describe the US Census as the “federal government’s largest and most complex peacetime operation”<sup>5</sup>. The 2022 South African census is estimated to have cost ZAR3.2 billion.<sup>6</sup>

Many developed countries are in the process of replacing their national censuses through linking of a variety of essentially complete administrative databases and official registers. However, given the generally poor quality of administrative data in developing countries, such an exercise would be foolish, and censuses take on even greater importance.

While the census offers a snapshot of the demographic and socio-economic characteristics of a country’s inhabitants at a point in time, the census also serves a multiplicity of other purposes. Apart from providing denominators for estimation of rates, it allows us to track changes in the population (and its characteristics) over time; it provides essential data for policy and planning (not least in the spheres of education, health, and infrastructure); it provides the sampling frame for state- and privately funded surveys; and – of particular relevance to South Africa – it is the single-most important constituent of the Equitable Share Formula used to allocate revenue from the National Treasury to the provinces.

This Commentary draws attention to and summarises the most pertinent results of a detailed Technical Report published by the South African Medical Research Council (SAMRC)<sup>7</sup> and lays out several concerns with the data that have been released from the 2022 South African census.

## Issues with the 2022 South African census

### *Balance equation (national)*

One of the most basic tools for evaluating a census or population projection is the ‘balance equation’. For censuses, the balance equation asserts that if all numbers are correct, the numbers in the current census (nationally, by age, by sex, by province, by population group, etc.) should be equal to those of the previous census, plus the births, less the deaths, plus in-migration less out-migration (i.e. net in-migration) over the intercensal period.

The work of the SAMRC-UCT (University of Cape Town) collaboration<sup>8,9</sup> which monitored deaths in South Africa during the COVID-19 pandemic, provides a solid understanding of the number of deaths in South Africa through to 2023. Likewise, the trend in the number of births is relatively well understood based on the results of previous censuses and surveys, analysis of the data on registered births released by Stats SA, as well as births recorded by the District Health Information System (DHIS), each corrected for late or under-reporting. Although the total numbers of intercensal births and deaths remain approximate, they are probably accurate, nationally, to within one hundred thousand.

Using the data published in the 2022 census release<sup>1</sup>, together with the results from the 2011 census and our estimates of the numbers of intercensal births and deaths in the country, we can reasonably accurately reconstruct the implied dynamics of the South African population over the intercensal period. Table 1 demonstrates this reconstructive process at a national level.



**Table 1:** Reconstruction of the South African population using the balance equation, at a national level

Reported census 2011 population (1)	51 770 561
+ Estimated intercensal births (2)	12 418 342
- Estimated intercensal deaths (3)	5 924 464
Implied immigration (4)	3 763 063
Reported census 2022 population (5)	62 027 502
Reported immigration (place of residence) (6)	618 910
Reported immigration (place of birth) (7)	461 542
Reported/implied migration = (6)/(4)	16%
Reported/implied migration = (7)/(4)	12%

Notes, sources, and formulae:

(1) Table 2.1<sup>1</sup>

(2) Own estimate (probably accurate to nearest 100 000)

(3) Own estimate (probably accurate to the nearest 100 000)

(4) (5) - ((1)+(2)-(3))

(5) Table 2.1<sup>1</sup>

(6) Table 3.1<sup>1</sup>; the reported value of in-migration in Limpopo is miscalculated based on the data shown

(7) Own estimate (subtracting estimated survivors of foreign-born identified in the 2011 census from those reported as foreign-born in the 2022 census)

To reconcile the estimates from the two censuses, net immigration (i.e. immigration less emigration) over the intercensal period would have to have been approximately 3.7 million, assuming both census populations were accurately estimated.

Not only does this number seem high given the restriction on travel during the COVID-19 pandemic, but it implies that only 16% of these migrants were identified by the census when asked where they were at the time of the previous census and even fewer (12%) when asked about place of birth. This level of under- and mis-reporting is significantly higher than in the previous intercensal period.

The estimate of the cumulative foreign-born (2.4 m) also stands in stark contrast to the foreign-born populations reported in the 2011 census (2.2 m) or the 1.769 m net immigrants over the period 2011–2021 assumed by Stats SA in their most-recently released population projections and repeated without comment in a report on migration in South Africa which also made extensive use of the 2022 census data.<sup>4,10</sup>

### Comparison with previous census results

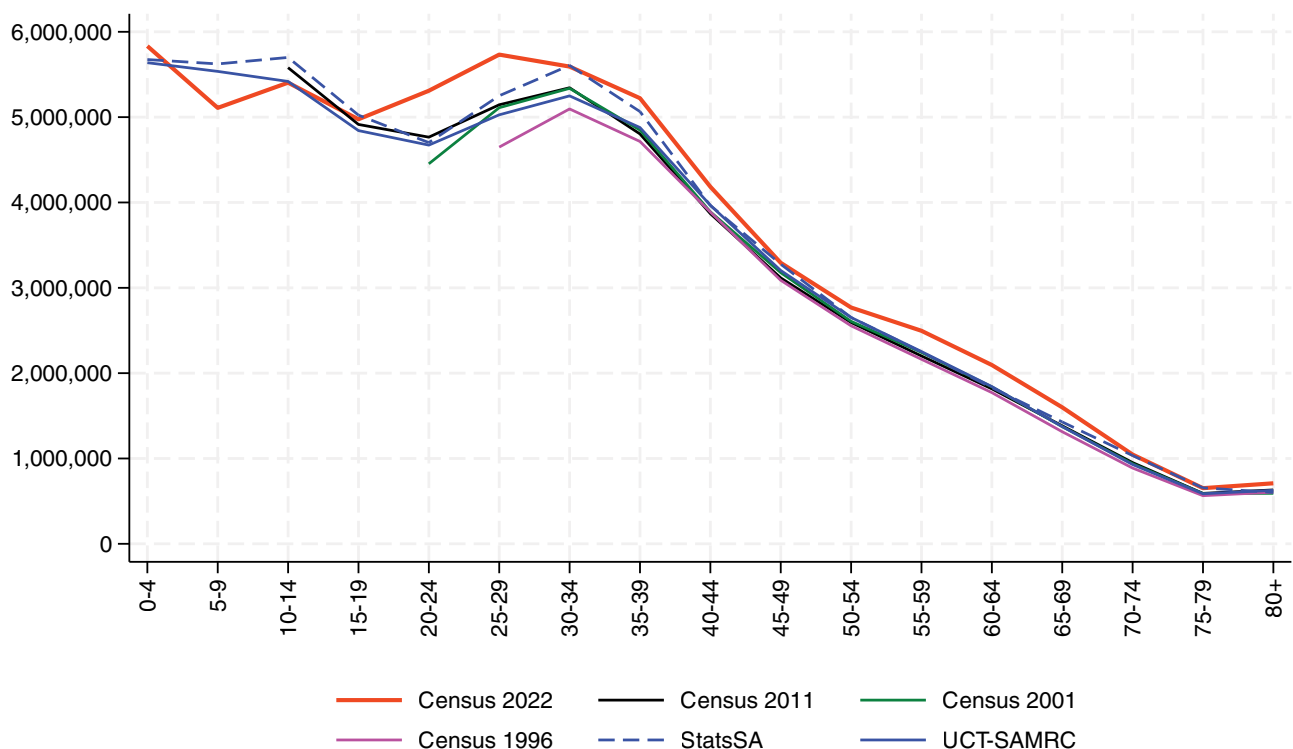
A second investigation was to compare the results from the most recent census to those projected from previous censuses. The sources of any observed discrepancy between the two numbers must arise from errors in either or both the census counts, or errors in the estimates of mortality or migration. Consistency in the numbers of the population by age observed in successive censuses, allowing for mortality, points to a general coherence of the data.

In addition, demographers make use of cohort-component projection models to provide a counterfactual population based on assumed patterns and trends in mortality, fertility, and migration. Demographers can estimate these components with reasonable accuracy (less so for migration) by applying an array of demographic techniques<sup>11</sup> to data from past censuses, national demographic surveys, and vital registration systems.

Again, consistency between the results of past censuses and population projections with the results from a more recent census increases confidence in the reliability of the more recent census data. Conversely, where the results from a more recent census depart markedly from a coherent and consistent prior set of results and population projections, the reliability of the more recent census is called into question.

### National population

Figure 1 shows the national population of South Africa by age as observed in the four post-apartheid censuses, as well as that used by



**Figure 1:** Census 2022 numbers by age compared to numbers projected from past censuses and models, at a national level.

the SAMRC-UCT collaboration to estimate excess COVID-19 deaths, and Stats SA (as released in their MYPE).

There is strong congruence between the estimates of the population (by age in 2022) from previous censuses (up to 2011), allowing for mortality and migration, and the two sets of population projections. The identified undercount of children aged under five in 1996<sup>11</sup> (who would be aged 26 to 30 in October 2022) is apparent. The MYPEs produced by Stats SA, especially between ages 30 and 39, are somewhat higher (and more in line with those from the 2022 census) than suggested by the SAMRC-UCT population projection. Although the Stats SA methodology for producing their population projections is somewhat opaque, this difference is likely mostly attributable to different assumptions regarding migration.

When one considers the age distribution of the difference between census numbers and those expected from the projection models (Figure 1), it can be observed that only a little more than 40% of the difference is to be found in the typical age range of migration (as evidenced from the previous intercensal period), namely people aged 20–39 at the time of the 2022 census. This again creates doubt as to the reliability of the data provided by the most recent census.

However, there are several other features that are of concern.

The first is an apparent underestimate in the 2022 census of the numbers of children aged 5–9. This is due to an unexplained undercount of children aged 5 last birthday (and if these children were counted at a different age, it was not in the 5–9 age group).

A possible explanation for much of the excess of the census estimates over the projections for ages 40+ is provided by a comparison of the census estimates by population group (proportionally reallocating the small numbers recorded as ‘Other’ to the four specified groups) to those projected from past censuses and by the SAMRC-UCT and Stats SA projections in Figure 2.

If one assumes that the excess of the census estimates of the African population over the projections in the age range 20–39 is due to migration unaccounted for by the projections, the estimates from the census for the African and coloured population groups are close to the numbers of the projections. Thus, much of the excess seen in the national population aged 40+ is to be found in the Indian and white population groups.

As noted, the estimated undercount of these two population groups was exceedingly high. It is quite likely that much of the excess of the census above the projections seen at the national level at these ages is due to an over-adjustment for undercount (with the true numbers lying somewhere between the SAMRC-UCT and Stats SA projections). The excess of the census over the projections amounts to 24% of the projected population for the Indian and 14% for the white population group.

### **Provincial populations**

Figure 3 shows the same comparisons for the provinces as that shown for the country in Figure 1. From this we see that the undercount of the 5–9-year-olds is apparent in all provinces and the excesses over age 50 is in all provinces except Gauteng (which is interesting, as this is where it would be most likely to appear if it was due to international migration) and the North West. The comparisons also highlight inconsistencies between the more recent censuses and, particularly, the 1996 and, to a lesser extent, 2001 censuses.

In summary, census 2022 numbers in total are probably reasonable approximations for the Free State, Gauteng, North West and Western Cape, possibly less so for Limpopo and Mpumalanga, and poor estimates for the Eastern Cape (it is improbable that there has been net immigration in the 20-39 age group) and KwaZulu-Natal (overstated for all adult ages), and unknown for the Northern Cape (which has proven difficult to estimate in the past as well).

### **Issues at a sub-provincial level**

The problems identified above manifest even more clearly when one evaluates the data at both district and local municipality levels. The

SAMRC Technical Report identifies several significant anomalies in the sub-provincial data, when compared with

- the district-level results from the 2011 census (which reveals implausible levels of population growth in some of the remotest and least developed parts of the country);
- Stats SA's district-level MYPE<sup>4</sup>; and
- the numbers of adults over 18 registered on the voters' roll at the time of the 2011 and 2021 Local Government Elections.

As before, discrepancies between data from these other sources and the census results at a district level cannot be taken as *prima facie* evidence of errors in the census; nevertheless, significant differences require investigation in order to understand why or how those differences might have occurred.

The Technical Report explores each of these in greater detail. However, as a single example, the 2022 census population of the Central Karoo District (DC4, in the Western Cape) increased by more than 40% between the 2011 and 2022 censuses; and is more than 35% higher than the MYPE of this district's population at the census date. Further, within this district, the population aged 18 and over of the Beaufort West local municipality (WC053), which accounts for around 70% of the district's population, increased by nearly 55% between the censuses, while the number of registered voters increased by only 13% between the two local government elections. Finally, examination of satellite imagery of the growth of the town of Beaufort West over this period indicates that the population growth implied between the two censuses is implausible.

### **The Post-Enumeration Survey**

The United Nations Statistics Division<sup>12</sup> recommends that a PES be carried out shortly after a census to estimate the extent of (and, often, to adjust for) a census undercount. Unfortunately, delays in completing the fieldwork meant that the 2022 South African PES was run several months (rather than weeks) after the census date.

With the 31% undercount, the PES-derived weights are doing much more heavy lifting in producing an estimate of the population in 2022. In other words, in the 2022 census, the population estimates are far more affected by the adjustment factors applied than in previous censuses. And, while it is the case that all PES-adjusted population results are estimates rather than counts of the population, the extent of the adjustment required in the 2022 census makes this observation all the more salient.

Our key concern in this regard is illustrated by a comparison of data from the PES reports from the 2011 and 2022 censuses.<sup>2,13</sup>

From Table 2, we note that the undercount standard errors (UC SEs) from the PES data are, as might be expected, markedly higher in 2022 than they were in 2011. Based on the PES, the 95% confidence interval for the national undercount in 2011 was narrow (14.34%; 14.86%). In 2022, the equivalent interval was much wider (27.99%; 31.21%). However, when it comes to the standard errors of the population corrected for the undercount, the uncertainty surrounding the census estimate for 2022 all but disappears, nationally, and is only higher (relative to 2011) in the 2022 census in the Western Cape and comparable (but still lower) in KwaZulu-Natal and Gauteng.

In short, it is difficult to understand how it can be that the two censuses (2011 and 2022) had broadly equivalent PES sampling fractions, but with the undercount in 2022 about twice as high, one can be 95% certain that the true size of the population is within 235 000 of the estimated population size, whereas in 2011, one could only be 95% certain that the true population was within 1.955 million either side of the estimated population size. In the absence of any other explanation, this suggests a computational error in the derivation of the standard errors for the population in the more recent census.

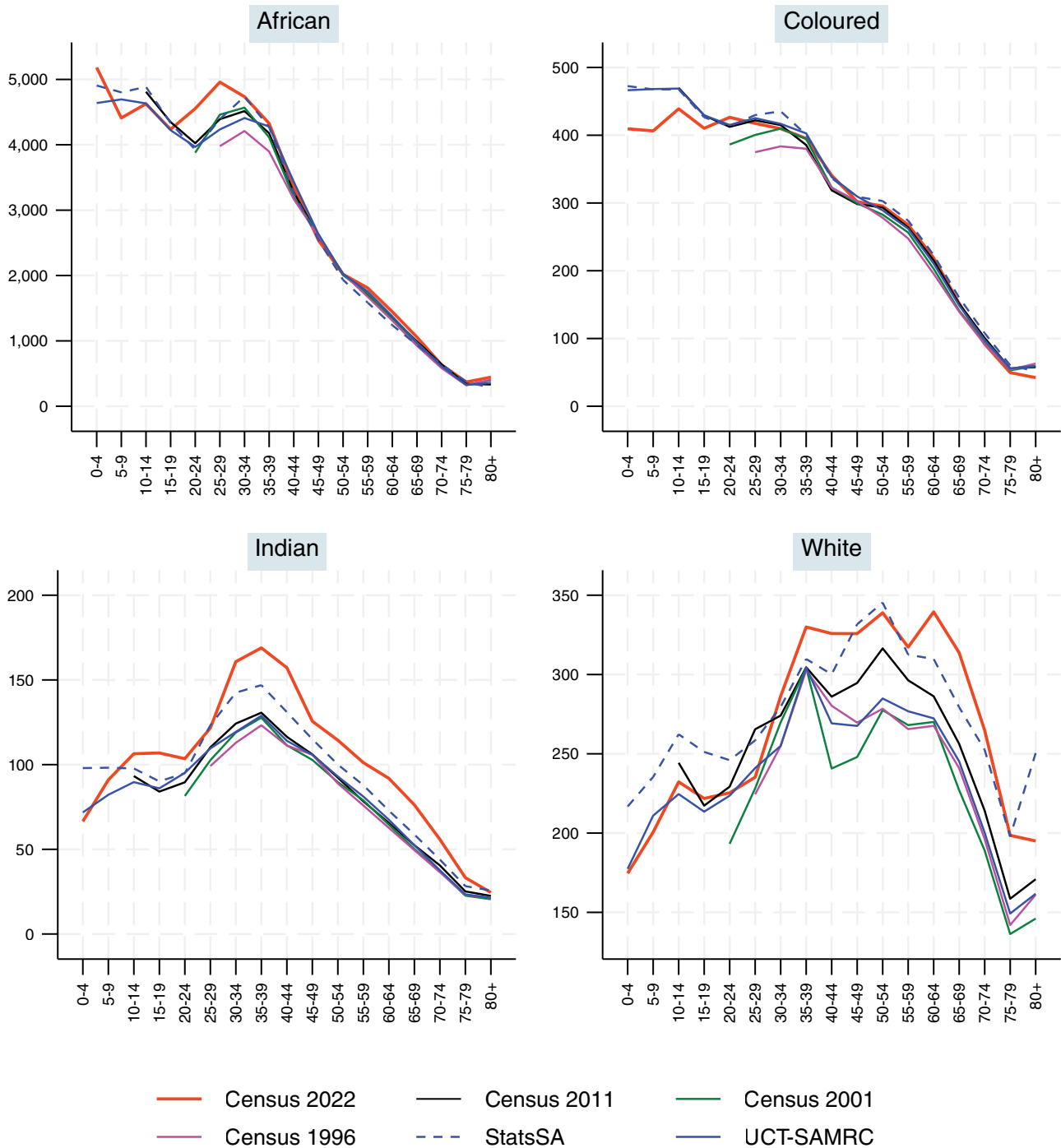


Figure 2: Census 2022 numbers (in thousands) by age compared to numbers projected from past censuses and models, by population group.

## Conclusions

Despite official pronouncements to the contrary, there is sufficient doubt about the quality and content of the data produced by Stats SA from the 2022 census to conclude that – as they currently stand – the data, at least in terms of numbers of people, may not be fit for purpose to assist with fiscal allocation, or for national, provincial, or local government planning and resource allocation. In addition, other important demographic data, *inter alia* on fertility, mortality, and migration, have not yet been released.

Although the investigation is at this stage, *perforce*, preliminary, it would appear that the population of South Africa may have been overestimated by as much as a million people, with about half of that number accounted for by significant overestimates of the white and Indian/Asian populations, particularly. Furthermore, this overall excess is concentrated in the age groups aged 50 and over – an age range

unlikely to have been affected by substantial in-migration and appears to be a little higher for men than women.

While a full analysis of all the factors that may have contributed to the poor execution of the 2022 census is beyond the scope of the current work, and would require access to internal documents that are not in the public domain, it is nonetheless possible to identify some of the factors that impeded the successful conduct of the census.

SARS-CoV-19 was first detected in South Africa in March 2020, which resulted in lockdowns of variable severity and repeated waves of infection and excess mortality that lasted until mid-2022. Although it is not clear to what extent preparations for a census in October 2021 were on track in March 2020, there is no doubt that the outbreak of the COVID-19 pandemic was a major interruption to the process, forcing Stats SA to delay the census.

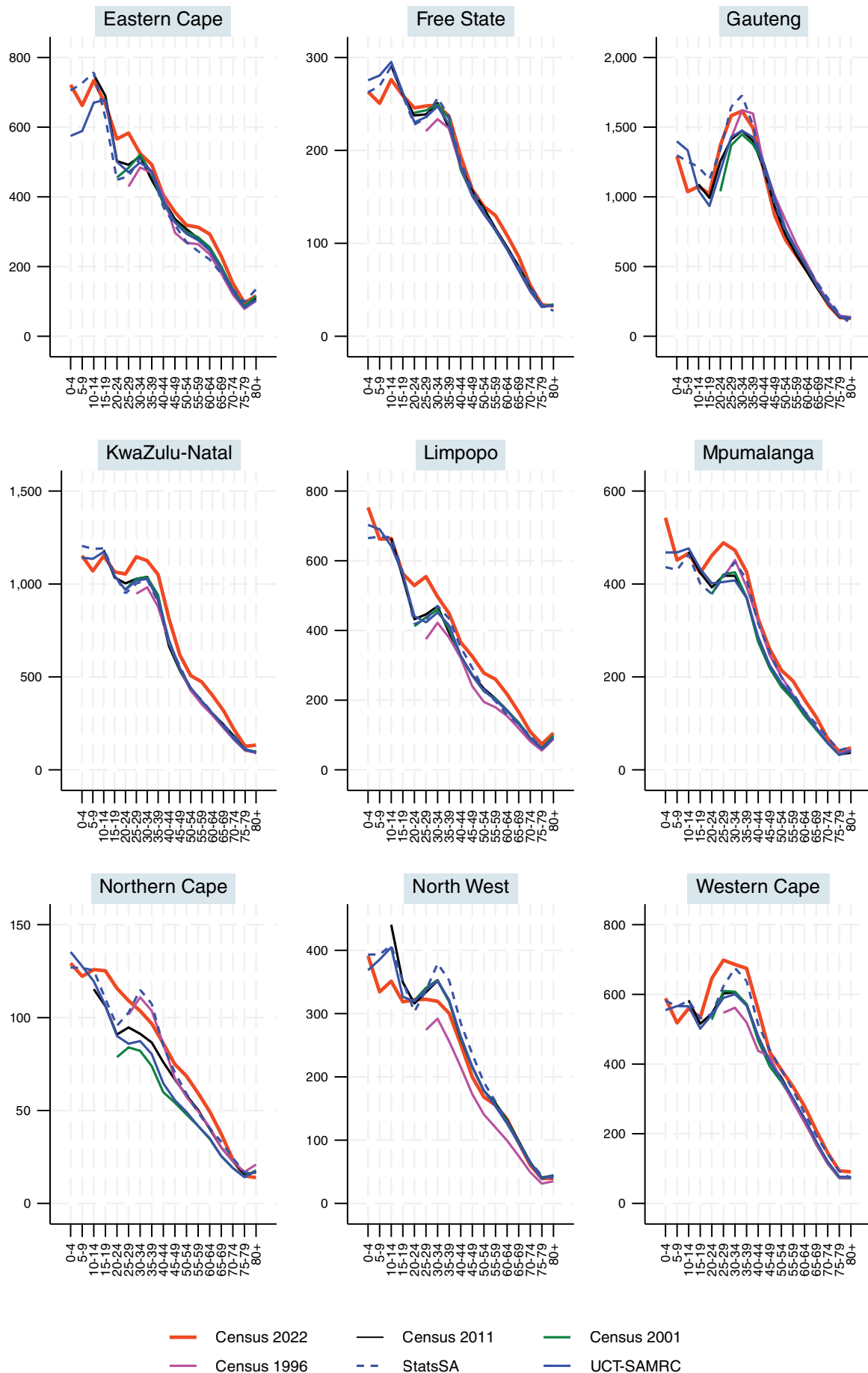


Figure 3: Census 2022 numbers (in thousands) by age compared to numbers projected from past censuses and models, by province.

**Table 2:** Estimates of the undercount (UC) and their standard errors (UC SE), and estimates of the population and their standard errors (SE), for South Africa and its provinces, for 2011 and 2022

	2011		2022		2011		2022	
	UC (%)	UC SE (%)	UC (%)	UC SE (%)	Population	SE	Population	SE
Western Cape	18.5	0.542	35.56	1.56	5 822 734	40 830	7 430 000	70 000
Eastern Cape	12.9	0.196	30.56	0.59	6 562 053	125 810	7 230 000	60 000
Northern Cape	13.4	0.318	29.11	2.27	1 145 861	82 466	1 350 000	30 000
Free State	10.1	0.362	18.52	2.62	2 745 590	117 567	2 960 000	40 000
KwaZulu-Natal	16.7	0.379	31.19	1.08	10 267 300	109 994	12 390 000	100 000
North West	14.9	0.532	17.92	5.39	3 509 953	166 754	3 800 000	40 000
Gauteng	14.7	0.174	30.79	2.07	12 272 263	106 023	15 120 000	90 000
Mpumalanga	15.5	0.473	34.00	4.51	4 039 939	219 299	5 160 000	60 000
Limpopo	10.0	0.135	23.77	1.02	5 404 868	251 244	6 570 000	60 000
South Africa	14.6	0.132	29.6	0.82	51 770 560	997 560	62 030 000	120 000

Sources: 2011<sup>13</sup> Undercount, Table 10; Population, Table 14; 2022<sup>2</sup> Undercount, Table 1; Population, Table 5

Had they been consulted, most demographers would have strongly motivated for the census to be deferred to 10 October 2022, or even 10 October 2023, to ensure that processes and implementation were not rushed, that it was on an anniversary of previous censuses, and that the census took part when the population was most stable (in terms of migration and other interruptions).

Unfortunately, the threat of the withdrawal of funding by National Treasury<sup>6,14</sup> if the census was not undertaken in the financial year ending March 2022, offered Stats SA Hobson's choice – either carry out the census in that financial year or lose the funding for the census, possibly to 2031. Thus, Stats SA was forced to undertake the census while not ready to do so, and entered the enumeration period in a state of significant unreadiness.

Both public and private sectors should exercise caution in drawing policy conclusions or making long-term plans based on these data. We are particularly concerned that – as they currently stand – the estimates of the population (by age, sex, province, population group) may lead to significant misallocation of resources through (for example) the Equitable Share Formula, or of education and health resources at national, provincial, or local government levels.

There is an urgent need to produce alternative population estimates that better describe the South African population in the mid-2020s than those in the census 2022 data. Those estimates might then be used to inform evidence-based resource allocation and planning in order to benefit the lives of all South Africans.

## Declarations

We did not use AI in the writing of this article. Both authors read and approved the final manuscript.

## Competing interests

We have no competing interests to declare.

## References

1. Statistics South Africa. Census 2022 statistical release. Report P0301.4. Pretoria: Statistics South Africa; 2023. Available from: [https://census.statssa.gov.za/assets/documents/2022/P03014\\_Census\\_2022\\_Statistical\\_Release.pdf](https://census.statssa.gov.za/assets/documents/2022/P03014_Census_2022_Statistical_Release.pdf)
2. Statistics South Africa. Post-enumeration survey statistical release. Report P0301.5. Pretoria: Statistics South Africa; 2023. Available from: <https://census.statssa.gov.za/assets/documents/2022/P030152022.pdf>
3. GBD 2021 Demographics Collaborators. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950–2021, and the impact of the COVID-19 pandemic: A comprehensive demographic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2024;403(10440):1989–2056. [https://doi.org/10.1016/S0140-6736\(24\)00476-8](https://doi.org/10.1016/S0140-6736(24)00476-8)
4. Statistics South Africa. Mid-year population estimates, 2022. Report P0302. Pretoria: Statistics South Africa; 2022. Available from: <https://www.statssa.gov.za/publications/P0302/P03022022.pdf>
5. US National Research Council. The 2000 census: Counting under adversity. Washington DC: National Academies Press; 2004. <https://doi.org/10.17226/10907>
6. Parliamentary Monitoring Group. Briefing by Statistician-General on the census 2022 results [webpage on the Internet]. c2023 [cited 2024 May 15]. Available from: <https://pmg.org.za/committee-meeting/37818/>
7. Moultrie TA, Dorrington R. The 2022 South African census. Cape Town: South African Medical Research Council; 2024. Available from: <https://www.samrc.ac.za/research-reports/2022-south-african-census>
8. Bradshaw D, Dorrington R, Laubscher R, Groenewald P, Moultrie T. COVID-19 and all-cause mortality in South Africa – the hidden deaths in the first four waves. *S Afr J Sci*. 2022;118(5/6), Art. #13300. <https://doi.org/10.17159/sajs.2022/13300>
9. Bradshaw D, Dorrington RE, Laubscher R, Moultrie TA, Groenewald P. Tracking mortality in near to real time provides essential information about the impact of the COVID-19 pandemic in South Africa in 2020. *S Afr Med J*. 2021;111(8):732–740. <https://doi.org/10.7196/SAMJ.2021.v111i8.15809>
10. Statistics South Africa. Migration profile report for South Africa. Report 03-09-17. Pretoria: Statistics South Africa; 2024. Available from: <https://www.statssa.gov.za/publications/03-09-17/03-09-172023.pdf>
11. Dorrington RE. Who was counted in? Some possible deficiencies with the 1996 South African census results. Paper presented at: Actuarial Society of South Africa Convention; 1999 November 2–3; Johannesburg, South Africa.
12. UN Statistics Division. Post-enumeration surveys: Operational guidelines [Technical Report]. New York: United Nations Statistics Division; 2010.
13. Statistics South Africa. Post-enumeration survey. Report 03-01-46. Pretoria: Statistics South Africa; 2012. Available from: [http://www.statssa.gov.za/census/census\\_2011/census\\_products/Census\\_2011\\_PES.pdf](http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_PES.pdf)
14. Statistics South Africa. Annual report 2022/23 (Book 1). Pretoria: Statistics South Africa; 2023. Available from: [https://www.statssa.gov.za/publications/AnnualReport/StatisticsSouthAfricaAnnualReport202223\\_Book1.pdf](https://www.statssa.gov.za/publications/AnnualReport/StatisticsSouthAfricaAnnualReport202223_Book1.pdf)