SA-TIED

Southern Africa – Towards Inclusive Economic Development

WORKING PAPER 238

Income inequality in South Africa

Evidence from individual-level administrative tax data

Chandré Jacobs,[1],* Amina Ebrahim,[2] Murray Leibbrandt,[1,2],* Jukka Pirttilä,[2,3,4] and Marlies Piek[5]

September 2024

UNU WIDER
national treasury
Department: National Treasury
REPUBLIC OF SOUTH AFRICA
SARS
Co-funded by the European Union
British High Commission Pretoria

About the project

Southern Africa – Towards Inclusive Economic Development (SA-TIED)

SA-TIED is a unique collaboration between local and international research institutes and the government of South Africa. Its primary goal is to improve the interface between research and policy by producing cutting-edge research for inclusive growth and economic transformation in the southern African region. It is hoped that the SA-TIED programme will lead to greater institutional and individual capacities, improve database management and data analysis, and provide research outputs that assist in the formulation of evidence-based economic policy.

The collaboration is between the United Nations University World Institute for Development Economics Research (UNU-WIDER), theNational Treasury of South Africa, the South African Revenue Services, and other universities and institutes. It is funded by the National Treasury of South Africa, the Delegation of the European Union to South Africa and UNU-WIDER through the Institute's contributions from Finland, Sweden, and the United Kingdom to its research programme.

[1] SALDRU, University of Cape Town, South Africa; [2] UNU-WIDER, Helsinki, Finland; [3] University of Helsinki, Finland;
[4] VATT Institute for Economics, Helsinki, Finland ;[5] National Treasury, Pretoria, South Africa ;*corresponding authors: C.Jacobs,JCBCHA015@myuct.ac.za; M.Leibbrandt, murray.leibbrandt@uct.ac.za

WIDER Working Paper 2024/55

Income inequality in South Africa

Evidence from individual-level administrative tax data

Chandré Jacobs,[1,*] Amina Ebrahim,[2] Murray Leibbrandt,[1,2,*] Jukka Pirttilä,[2,3,4] and Marlies Piek[5]

September 2024

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: We use individual-level tax administrative data to estimate personal income inequality among the tax-compliant population in South Africa over the period 2011–21. Our results indicate that inequality of this population rose slightly over the period, with the Gini coefficient increasing from 0.64 to 0.66. The aggregate inequality metrics mask some notable changes in real incomes driven by relatively stronger real income growth at both the top and bottom ends of the distribution and very sluggish real income growth in the middle deciles. We also exploit the longitudinal nature of the data to conduct a preliminary examination of income mobility across this income distribution, particularly within the top income groups. We find low levels of upward mobility, driven by very little upward mobility into the top decile and low downward mobility out of this decile, particularly for those in the top percentile in 2011.

Key words: top income, income inequality, mobility, South Africa, administrative tax data

JEL classification: D3, N37, 015

Acknowledgements: Chandré Jacobs acknowledges support from UNU-WIDER to conduct this research. We thank participants of the 2022 WIDER Development Conference in Bogotá, Colombia, the 2023 Nordic Conference in Development Economics in Gothenburg, Sweden, the 79th Annual Congress of the International Institute of Public Finance at Utah State University, Logan, USA, and the Tenth Meeting of the Society for the Study of Economic Inequality (ECINEQ) at Aix-Marseille University, France, for their comments and feedback. We acknowledge considerable interaction and help from Bruce McDougall, Michael Kilumelume, and others managing the National Treasury Secure Data Facility as part of the SA-TIED partnership between the National Treasury and UNU-WIDER.

[1] SALDRU, University of Cape Town, South Africa; [2] UNU-WIDER, Helsinki, Finland; [3] University of Helsinki, Finland; [4] VATT Institute for Economics, Helsinki, Finland; [5] National Treasury, Pretoria, South Africa; * corresponding authors: C. Jacobs, JCBCHA015@myuct.ac.za; M. Leibbrandt, murray.leibbrandt@uct.ac.za

This study has been prepared within the UNU-WIDER project Southern Africa—Towards Inclusive Economic Development (SA-TIED).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-517-2

https://doi.org/10.35188/UNU-WIDER/2024/517-2

Typescript prepared by Ayesha Chari.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction and background

Given that inequality is associated with various social ills (Wilkinson and Pickett 2011), the persistently high levels of income inequality in South Africa continue to generate much public discussion and concern. Reliable estimates of the trends and drivers of inequality are fundamental to formulating effective strategies to mitigate its adverse effects.

There is extensive empirical work on income inequality in South Africa, mostly using publicly available household survey data (see review in Leibbrandt and Díaz Pabón 2022). Using 1993 data from the Project for Statistics on Living Standards and Development (PSLSD), the 2005 Income and Expenditure Survey and the 2008 base wave of the National Income Dynamics Study (NIDS), Leibbrandt et al. (2010) show that income in South Africa had become increasingly concentrated in the top decile between 1993 and 2008. The authors also show a rise in the Gini coefficient for per capita income from 0.66 to 0.70 over the same period. Hundenborn et al. (2018) use data from PSLSD and Waves 1 and 4 of NIDS and find similar results: the Gini coefficient increased between 1993 and 2008 but decreased slightly to 0.66 in 2014. Both Leibbrandt et al. (2010) and Hundenborn et al. (2018) show that overall income inequality tends to be driven mainly by earnings inequality in the labour market.

Wittenberg (2017b) uses earnings data from the Post-Apartheid Labour Market Series (PALMS)[1] and finds that while aggregate wage inequality increased in the 1990s, it remained roughly constant from 2000 to 2011. However, he shows changes within the wage income distribution over this period: the top of the distribution moved away from the median, whereas the bottom moved closer to the median. Finn and Leibbrandt (2018) report similar results to Wittenberg's (2017b).[2]

Though it is encouraging that the general conclusions from the literature on inequality using survey data are reasonably consistent, household survey data often suffer from under-sampling, non-response, and measurement error. These issues are of particular concern when measuring inequality, as high-income households may misreport incomes and are less likely to participate in surveys (Wittenberg 2017a). Re-weighting and imputations cannot always rectify these issues.

Administrative tax data provide better coverage of top incomes and are gaining traction as a source of information for empirical research (Card et al. 2010). There is a growing trend of studying income inequality using tax administrative data in both developed and developing countries (see, for example, Drechsel-Grau et al. 2022; Jäntti et al. 2022). Tax data are a valuable complement and cross-check to inequality estimates from survey data. However, access to tax data is fairly recent, and the opportunities provided by individual-level tax data to examine income inequality in South Africa have not yet been fully exploited (Ebrahim and Axelson 2019).

Most researchers using tax data to study income inequality in South Africa have used payroll data on pay-as-you-earn (IRP5) or final, personal tax returns (ITR12) data.[3] For example, Kerr (2021) describes earnings inequality using survey and payroll data, while Wittenberg (2017a) compares

[1] PALMS is a stacked cross-sectional dataset consisting of 69 household surveys (Kerr and Wittenberg 2019).

[2] These analyses also raise concerns over the consistency of the earnings series in the Quarterly Labour Force Survey data underlying PALMS from 2012 onwards, making the availability of the tax data important and useful.

[3] While IRP5s are submitted by employers on behalf of employees, ITR12 returns are completed by the individual. Employee income tax certificates are also issued by pension funds in respect of annuity income. Employed and retired individuals are recipients of these certificates.

earnings in the Quarterly Labour Force Survey to a sample of assessed personal tax records. Kerr (2021) finds that earnings inequality remained stable over the 2011–17 period, with a Gini coefficient of earnings hovering around 0.6. However, like Wittenberg (2017b), Kerr (2021) finds that this stability masks a situation in which the top of the earnings distribution moved away from the median, the median earnings remained roughly constant, and the bottom end moved closer to the median. Using combined IRP5 and income tax returns data, Ebrahim and Axelson (2019) calculate income ratios for taxable income and present results congruent with the findings of Wittenberg (2017b) and Kerr (2021).

Bassier and Woolard (2020) use aggregated personal income statistics from the final income tax returns to investigate the growth in income at the top end of the distribution in South Africa. Hundenborn et al. (2019) combine a 20 per cent sample of the ITR12 data for each of the 2011 and 2014 tax years and data from Waves 2 and 4 of NIDS to investigate the effect of top incomes on inequality across the entire income distribution in South Africa. They motivate the combination of these two data sources by the assumption that tax data capture incomes at the top end of the distribution more accurately, while household survey data provide better information at the bottom of the income distribution. They show significant differences in taxable income between the two data sources, especially at the ends of the income distribution. Focusing only on those with incomes above the compulsory tax filing threshold in the tax data, they find that the Gini coefficient for taxable income decreased from 0.367 in 2011 to 0.349 in 2014.[4] This subgroup's level of income differences is much lower than overall income inequality. Combining the NIDS and tax data, they find that the Gini coefficient decreased from 0.83 in 2011 to 0.79 in 2014.

Other researchers have also used South African tax data in their work. Like Hundenborn et al. (2019), Chatterjee et al. (2022) brings different sources of data together, but their focus is on wealth (rather than income) inequality. Chatterjee et al. (2021) examines the redistributive impacts of the tax-benefit system in South Africa for an extended period (1993–2019) using the distributional national income approach.

Although it would be optimal to study income inequality using a dataset that consistently captures incomes from different sources at the top and bottom of the distribution, we do not attempt to merge the tax data with other data in this paper. Merging survey and tax data is a substantial and important task that must be carefully considered. We leave this work for a future project building off initial work by Hundenborn et al. (2019) and Chatterjee et al. (2021). For now, we briefly discuss what proportion of the South African personal income distribution we capture in the tax data (Section 2) and use only data from the Individual Panel between 2011 and 2021[5] to provide an overview of baseline income inequality for formal sector workers, those who receive annuity income, and those who submit their own income tax returns (who we define as the 'tax-compliant population'). The Individual Panel combines information from the employee income tax certificates (IRP5s) and final personal income tax returns (ITR12s) at the tax-payer level, allowing for a more comprehensive picture of income sources and the income disparities of those covered by the tax data. The unit of analysis in this paper is the individual and not the household or the household's pooled income normalized by some equivalence scale. In the South African literature, per capita incomes are usually reported. Our approach is similar to many other studies that investigate (top) income inequality using tax data.

---

[4] The filing threshold was 67,111 South African rands (ZAR) for 2014 and changes each year. Details on the SARS website: https://www.sars.gov.za/tax-rates/income-tax/rates-of-tax-for-individuals/

[5] The tax year in South Africa runs from 1 March to 28 February. In the data the year 2021 refers to the 2020/21 tax year, which runs from 1 March 2020 to 28 February 2021.

Since the filing threshold for the payroll tax certificates in South Africa is very low (ZAR2,000[6] a year), we capture almost all those working in the formal sector and those who earn annuity income. On the other hand, the income threshold for mandatory filing for the personal income tax is higher (close to ZAR80,000 for the final years in the data), implying that the total incomes are captured for this population. However, since few persons in lower income groups have other forms of taxable income than labour income, we can capture almost all incomes reported to the South African tax authority for the population. According to our calculations, close to 80 per cent of the population in the upper half of the distribution works in the formal sector, and therefore we have a very good picture of incomes of the upper half of income earners. In the top 10 per cent group, the formal sector share rises to about 95 per cent, and we can capture extremely well the reported incomes for the top income groups. Consequently, we pay special attention to income earners in the top groups (top 10, top 1, and smaller groups) in this paper.

This paper makes four main contributions to the income inequality literature. First, we carefully prepare the tax data to provide a comprehensive picture of the individual-level income distribution for the tax-compliant population in South Africa. One major omission of the data is that dividends are not well captured since dividend income is subject to a separate withholding tax. Nonetheless, the underlying data are stronger than in the existing literature and we describe it carefully in Section 2 and Appendix 1 and 2 before presenting our results. Second, we disaggregate labour and capital income sources and show the development of different income components over the 2011–21 period. Third, we describe the top income shares and income composition. Finally, we make use of the panel element of the data to conduct a preliminary examination of income dynamics and the stability of the income distribution, again focusing on the top end. A thorough examination of income dynamics at the top end is beyond the scope of this paper and will be considered in future work. Overall, this paper serves as a foundation for research on individual-level income inequality using the tax data available in South Africa.

Our results show that personal income inequality remained stable over the period, with a Gini coefficient remaining in the 0.64–0.66 range. This estimate is very close to the World Income Inequality Database (WIID) estimate of 0.65 for individual gross incomes Gini in 2017 (UNU-WIDER 2022). Taxes reduce inequality, but even after tax, inequality remains high. Decompositions reveal that labour income is the primary driver of inequality, as previous research with household survey data suggests (see Hundenborn et al. 2018; Leibbrandt et al. 2010). We find that the top decile receives almost half of the total income. Income from capital sources become more important as incomes increase. Income mobility for the tax-compliant population in South Africa is relatively low: 32 per cent of the bottom decile in 2011 remained there in 2021, while almost 55 per cent of those in the top decile in 2011 were in the top decile in 2021. We find that 56 per cent of the top 0.1 per cent in 2011 were in the top 1 per cent in 2021. Our results are described in Sections 3–5, while Section 6 concludes.

2 Data description and methods

The tax data we use are described in detail in Ebrahim and Axelson (2019). To summarize, the anonymized Individual Panel combines the information contained in the IRP5 (payroll) and ITR12 (personal income tax returns) tax records. It consists of four linked panel datasets: the ID Panel, the Employment Panel, the Source of Income Panel, and the Income Panel. The ID Panel contains (anonymized) identification variables for each IRP5 and ITR12 tax submission and does not

[6] Approximately USD100. In contrast, the minimum wage was about ZAR22 an hour in 2021.

contain any income information. The Employment Panel contains all the information provided in each IRP5 certificate only. We use the Source of Income Panel and the Income Panel.[7] The Source of Income Panel contains information on the different types of income an individual receives in a tax year (included in their IRP5 or ITR12). At the same time, the Income Panel provides aggregated information on tax liability and other types of income for each person each year. We use the 2023 version of the Individual Panel, which includes IRP5 certificates and ITR12 returns for the 2010/11 to 2020/21 tax years.[8] The tax year runs from 1 March until end February. Hence, the tax year ending in 2021 mostly captures incomes from 2020.

By combining the IRP5 and ITR12 tax records, the Individual Panel provides a near complete picture of the income distribution of formal sector workers in South Africa, as well as those who submit their tax returns (like the self-employed). The tax data typically cover the incomes reported to the tax authority for the entire population, since the filing threshold for the employers is ZAR2,000. Of course, not all work in the formal sector, and Section 2.2 examines the formal sector employment shares by income groups in survey data. Also, some other incomes fall below the mandatory filing threshold for personal income tax returns—which stood at ZAR75,750 in 2018 for persons up to 65 years of age, and could be missing. We regard this as a minor issue since other forms of taxable income are rare for individuals with lower income. The tax data do not capture those who are informally employed or rely on earnings from non-taxable sources (such as government grants). Consequently, when using only the tax data, we exclude individuals at the bottom of the income distribution, and our inequality estimates may be too low.

Though we expect the tax data to provide better information at the top end of the income distribution, some individuals may hide incomes illegally (tax evasion) or underreport incomes in their ITR12 returns. Furthermore, the Individual Panel does not fully account for payments received through unit trusts or investment funds and data improvements in this regard are ongoing. Other forms of capital income, such as rental income, interest income, and dividend income, are captured but may also be underreported. Chatterjee et al. (2022) show that reported capital incomes in the tax microdata fall significantly below the totals in the national accounts.[9] Nonetheless, these issues would be more severe when using household survey data.

While survey data cover a wide range of demographic variables, this is not the case with administrative tax data. There is no race or education information in the Individual Panel, which would be particularly interesting for decomposing inequality into within-group and between-group contributions. However, the panel does include derived age and gender variables.[10]

We aim to provide a descriptive overview of income inequality using these tax data. In line with this, we use mostly graphs and tables to report our findings in the sections that follow. To measure inequality, we rely on standard measures such as the Gini coefficient, Theil indices, and Atkinson

[7] We merge the two panels by derived ID and tax year. For more details, see Appendix B on 'data'.

[8] The version of these data that we use in this working paper is an intermediate version that was available to researchers in November 2023. Changes will be made in the coming months before the Individual Panel is finalized. The version of the Individual Panel we use contains IRP5 certificates for the 2021/22 tax year, but there is a shortfall in the incomes that are recorded in the ITR12s for this tax year (i.e. 2021/22). This is because not all ITR12 submissions were submitted before the data were extracted. Consequently, we drop the 2021/22 tax year from our analysis, and focus only on the 2010/11–2020/21 tax period (i.e. 2011–21).

[9] Interest income reaches 25–30 per cent of total interest income received by households in the national accounts, whereas rental income and dividends cover between 7–13 per cent and 2–25 per cent of the national accounts totals (see Chatterjee et al. 2022: Table S4.7).

[10] Gender is a derived variable from the seventh digit of an individual's ID number. It is important to note that when we examine top incomes by gender, the analysis is skewed towards individuals with South African ID numbers.

measures. We make use of transition matrices and exploit the longitudinal nature of the data to describe mobility among the tax-compliant population.

2.1 Income variable

The key measure we use in this paper is gross income. We construct an income variable directly from the source codes[11] in the Source of Income Panel. We categorize the source codes into three broad categories: labour income, capital income, and other income. Labour income is made up of four sub-categories i.e., business income, earnings, pension income and other labour income. Capital income comprises interest income, rental income, capital gains income, dividend income, and other investment income. Other income is a sundries category for income from activities such as gambling and animal showing. These income sources and the respective tax forms submitted to the South African Revenue Service (SARS) are summarized in Table 1. Appendix Table B1 provides details on the source codes included under each sub-category.

Table 1: Income sources included in gross income

Broad category	Sub-category	Tax form
Labour income	Business income	ITR12
	Earnings	IRP5
	Other labour income	IRP5
	Pension income	IRP5
Capital income	Interest income	ITR12
	Rental income	ITR12
	Dividends	ITR12
	Capital gains	ITR12
	Other investment income	ITR12
Other income	Activity income	ITR12

Source: authors' compilation based on study data.

We define disposable income as the after-tax income, calculated as an individual's gross income minus their tax liability. All incomes reflect December 2021 prices. We exclude individuals with negative gross income for the annual analysis but include them in the long-term analysis. We also consider only individuals aged 16 years or older.

While the tax data provide detailed information on individual taxable incomes at the source code level, source codes can change over time and new codes may be added by SARS. The tax forms may also change. There is one such instance that is worth mentioning here: before 2017, employer contributions to pension funds were not included on the IRP5 certificates and hence not available in the data (Redonda and Axelson 2021). From 2017 onwards, these contributions are present in the data. We have included these codes in our income variable as these would form part of labour income (Kerr 2021).

We identify outliers and potential data entry errors in the data. We also corrected some other minor errors, all of which are detailed in Appendix B.

[11] A source code is a four-digit number that the South African Revenue Service (SARS) uses to capture information on the income tax return.

2.2 What can the Individual Panel tell us about the distribution of personal income in South Africa?

To understand how the distribution of individual incomes in the Individual Panel corresponds to the distribution of personal income in South Africa, we start by examining the proportion of individuals in different income groups who are formally employed using data from NIDS (Southern Africa Labour and Development Research Unit 2018). Formally employed individuals are most likely to appear in the Individual Panel because their employers submit their IRP5s.[12] In 2017, around 77 per cent of employed individuals were employed in the formal sector.[13] Most of the individuals who are formally employed are in the top half of the (monthly) gross income distribution (see Table 2). As income increases, the proportion of individuals in the formal sector increases and hence they are more likely to be included in the Individual Panel.[14]

Table 2: Proportion of employed individuals in the formal and informal sectors, 2017

Income group	Proportion employed in informal sector	Proportion employed in formal sector
Bottom 50	45.77	54.23
Top 50	22.82	77.18
Top 10	4.09	95.91
Top 5	4.23	95.77
Top 1	2.68	97.32

Note: data are weighted.

Source: authors' calculations based on NIDS Wave 5 (2017).

We now look at the percentiles in the personal total (annual) income distribution in Waves 2, 4, and 5 of NIDS that correspond to various filing thresholds that would be relevant for individuals who need to file ITR12s. There are different thresholds for individuals in different age groups (see Appendix Table C1).

Table 3 shows the first, 25th, 50th, 75th, 90th, and 99th percentiles; the first non-zero income percentiles; and the percentiles closest to the tax thresholds in each of the survey years.[15]

[12] They may not be in the panel if the income received in the tax year was less than ZAR2,000.

[13] The proportion of workers in the formal sector is around 75 per cent in NIDS Waves 2 and 4. The method we use to determine formality status in NIDS, and all other NIDS work, is described in Appendix C.

[14] We observe a similar trend in the other survey years (NIDS Waves 2 and 4).

[15] The tax thresholds for the 2010/11 tax year are used for Wave 2, the thresholds for the 2014/15 tax year for Wave 4, and the 2017/18 thresholds for Wave 5.

Table 3: Total income percentiles in NIDS

Percentile	2010/11	2014/15	2017/18
Panel A: Threshold (%) for individuals below the age of 65 years			
1	0	0	0
25	0	0	0
29.1		70	
32.3			243
44.8	21		
50	5,226	16,500	16,106
75	32,746	58,496	58,608
82.5			**90,253**
83			**93,932**
83.6		**98,630**	
83.7		**100,000**	
87.6	**98,601**		
87.7	**99,668**		
90	123,776	175,732	184,435
95	239,373	309,192	329,670
99	639,389	893,036	823,317
Panel B: Threshold (%) for individuals between 65 and 74 years			
1	0	0	0
6.1		279	
10.5	1,748		
11.6			716
25	22,448	21,757	21,845
50	22,657	22,594	23,105
75	25,385	33,426	33,700
90	120,507	100,279	115,802
92			**131,214**
92.5			**144,404**
94.2		**147,448**	
94.3		**156,084**	
94.8	**146,853**		
95	**166,957**	172,614	219,029
99	461,539	404,313	586,081
Panel C: Threshold (%) for individuals 75 years and older			
1	0	0	0
9.6		3,008	
16.4	2,517		
16.8			672
25	20,979	21,727	21,845
50	22,618	22,563	23,105
75	23,449	35,532	30,298
90	83,916	116,667	106,311
91.5			**131,068**
92			**157,895**
93.7	**146,161**		
93.8		**169,923**	
93.9		**184,367**	
95	**160,563**	232,673	187,114
99	293,706	889,488	428,987

Note: data are weighted; the percentiles closest to the tax thresholds in each year are in bold.

Source: authors' calculations based on data from NIDS Wave 2 (2010/11), Wave 4 (2014/15), and Wave 5 (2017).

It appears that over time the tax thresholds for the different age groups have moved slightly down the personal total income distribution. The tax threshold for individuals below the age of 65 years was between the 83rd and 88th percentile between 2010/2011 and 2017/18.[16] For individuals in the age group 65–74 years, the tax threshold was between the 92nd and 95th percentile.[17] A similar trend can be seen for individuals aged 75 years and older. However, even if individuals earn below the respective tax thresholds, a large proportion of them would be included in the Individual Panel if they worked in the formal sector. Table 4 shows that among the working age group, roughly two-thirds of individuals are formally employed and would thus be included in the Individual Panel, despite earning below the tax threshold.

Table 4: Proportion of individuals (in %) with total income below tax threshold employed in the formal sector, by age group and year

	Year		
Age group	2011	2015	2018
Below 65 years	66.29	63.34	66.75
Between 65 and 74 years	25.39	26.22	40.14
75 years and above	—*	4.1	10.87

Note: data are weighted. *In 2011, there were no individuals 75 years or above employed in the formal sector.

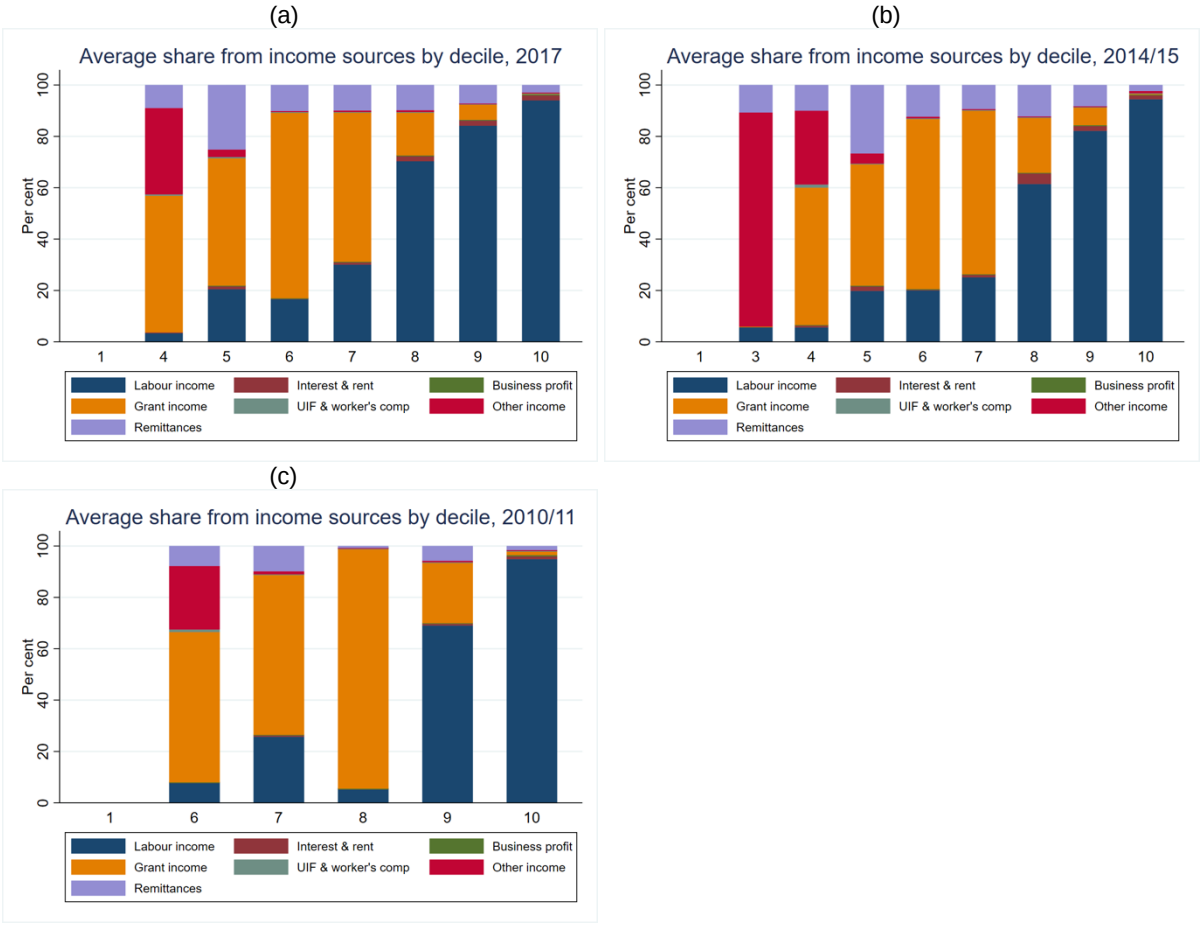
Source: authors' calculations based on data from NIDS Wave 2 (2010/11), Wave 4 (2014/15) and Wave 5 (2017).

We now look at the different sources of total income across the deciles of the South African personal income distribution in the NIDS data. In Figure 1 we observe that non-zero income for individuals starts at the fourth decile in 2017. In 2014/15 non-zero income starts at the third decile, while in 2010/11 income is first observed in the sixth decile. Income from grants, remittances, and other sources are most important for the lower deciles. Labour income becomes more important as incomes increase. While investment income and business profit contribute to total income for the top two deciles (and the eighth decile in 2017 and 2014/15), the average shares of these income sources are quite small. When we restrict the sample to those who are formally employed, a similar picture emerges, though labour income makes up a larger share in each decile across years (see Appendix Figure C1).

[16] In 2017 the threshold for individuals below 65 years was ZAR90,718 (December 2021 prices). In 2014/15 it was ZAR 99,158 and in 2010/2011 it was ZAR99,476.

[17] The threshold for this age group was ZAR140,479 in 2017, ZAR154,558 in 2014/2015, and ZAR154,499 in 2010/11.

Figure 1: Average share from income sources for (a) 2017, (b) 2014/15, and (c) 2010/11



Note: UIF, unemployment insurance fund. Based on total personal (annual) income; data are weighted.

Source: authors' illustration based on data from NIDS Wave 2 (2010/11), Wave 4 (2014/15), and Wave 5 (2017).

Taken together, these results suggest that the Individual Panel, which includes individuals who have IRP5s and/or ITR12s, captures about two-thirds of individuals who earn below the tax thresholds and the population of tax-compliant taxpayers who earn above the tax thresholds. Of course, these individuals can live in households lower down in the income distribution.

3 Annual changes in the income distribution

In this section, we describe overall changes in the distribution of personal gross income in the tax data and the sources of income for the top percentiles over time. We also discuss developments in income shares and before- and after-tax inequality over the 2011–21 period.

3.1 Income summary statistics

Real income changes

Table 5 presents summary statistics of the real gross income of the tax-compliant population over the 2011–21 period. The figures for mean income display an upward trend, with average income increasing from ZAR211,141 in 2011 to ZAR242,426 in 2021. The medians (which are less sensitive to very high incomes and potential outliers) were in the ZAR94,000– ZAR98,000 range over most of the period. The median in 2021 is lower compared with all other years—just below

ZAR92,000. While the number of individuals appearing in the Individual Panel has generally increased over time, there were fewer individuals in 2021 (the 2020/21 tax year) compared with the previous 4 years. This decrease in the number of individuals could be related to individuals' income falling below the tax threshold or out of formal employment during the COVID-19 pandemic.

Table 5: Mean and median real gross income, 2011–21

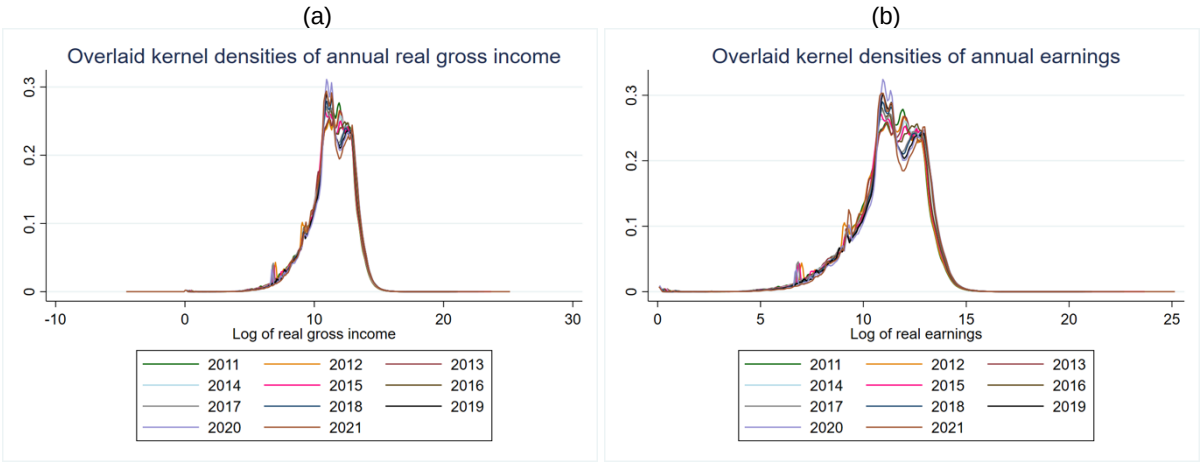
Year	Mean	Median	Observations
2011	211,141	96,258	12,769,345
2012	213,308	94,497	13,227,358
2013	218,475	95,540	13,449,739
2014	223,549	96,119	13,654,176
2015	229,437	93,955	14,001,074
2016	231,008	96,474	13,994,032
2017	241,492	96,867	14,165,017
2018	242,204	98,181	14,213,622
2019	241,246	97,446	14,376,224
2020	243,802	97,249	14,432,721
2021	242,426	91,631	13,998,352

Note: mean and median amounts expressed in December 2021 rands.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Figure 2 plots the income distribution of real gross income and real earnings for individuals in the panel each year. The figure shows a very slight rightward shift in real gross income over the period.[18] There have been some changes in the middle of the distribution, with the slight trough becoming more exaggerated over time. The gross income distribution mirrors the distribution of (real) earnings. The spikes in the lower end of the distribution may be due to bunching below a threshold, but since earnings would be captured on the IRP5 (submitted by employers and not individuals themselves), it is unclear whether this is the case.

Figure 2: Overlaid kernel densities of (a) annual real gross income and (b) annual earnings, 2011–21

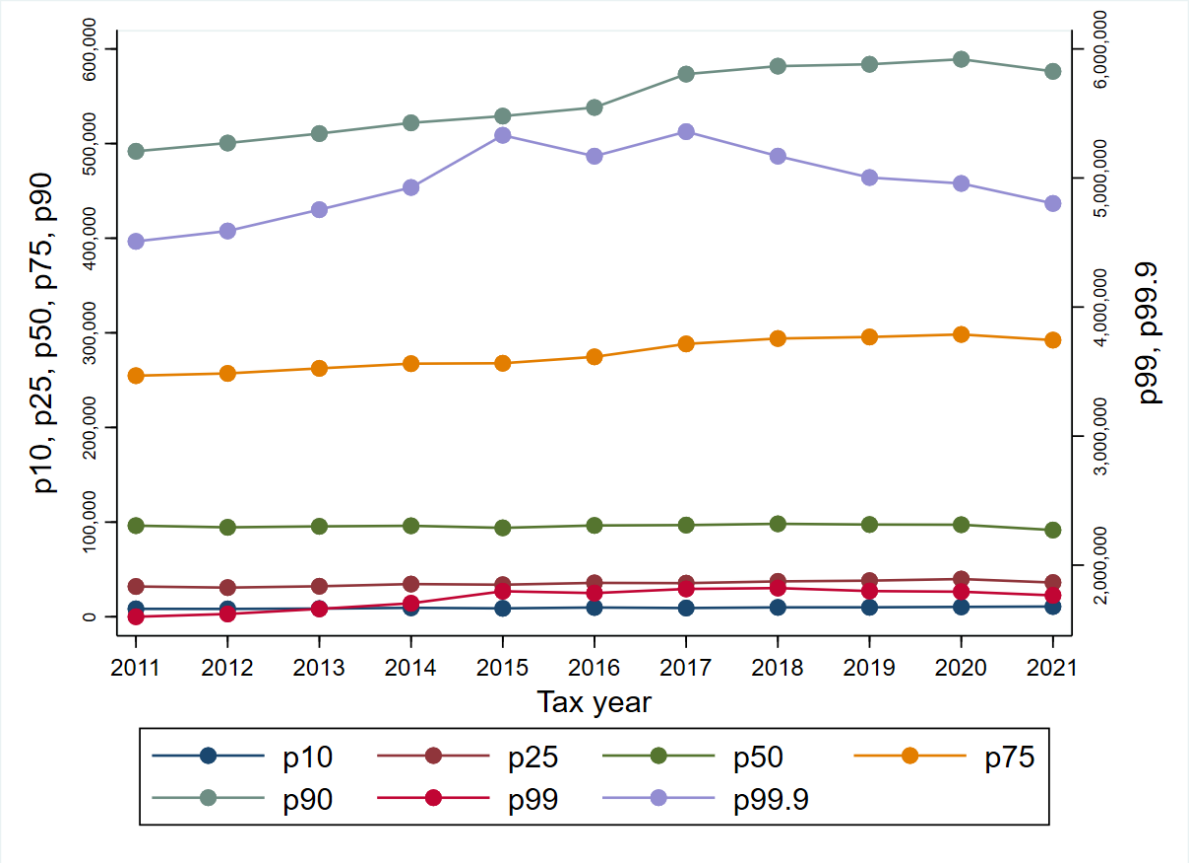


Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

[18] There is a much clearer rightward shift in the distribution of nominal income (see Appendix Figure A1).

Figure 3 illustrates how levels of annual real gross income at selected percentiles changed over time. Explicit attention is given to the top 1 and 0.1 per cent. The left axis indicates income levels for the 10th to 90th percentiles and the axis on the right shows the income levels for the 99th and 99.9th percentiles. The 90th, 99th, and 99.9th percentiles in the tax data increased slightly over the period while the 10th, 25th, 50th, and 75th percentiles remained fairly stable. Bassier and Woolard (2020) report that the incomes of the top 1 per cent increased substantially between 2003 and 2017. This upward trend in top incomes is noticeable in Figure 3, especially at the 99.9th percentile, which peaked in 2017. However, our results indicate a decrease in the 99.9th percentile over the 2017–21 period, with income at the 99.9th percentile in 2021 being roughly the same as in 2013.

Figure 3: Change in real gross income by percentile, 2011–21



Note: lower percentiles measured on the left *y*-axis and p99 and p99.9 on the right *y*-axis.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Table 6 reports the average real income within each decile in the sample over the period. Average income appears to have decreased for the fifth and sixth deciles over time and increased for all other deciles. The increase in absolute income was smallest for the bottom decile and largest for the richest decile.

Table 6: Average real gross income by decile, 2011–21

	Decile									
Year	1	2	3	4	5	6	7	8	9	10
2011	3,569	15,091	32,191	52,273	78,834	117,382	170,853	255,902	390,101	995,225
2012	3,493	14,254	31,044	51,037	77,265	116,036	171,679	258,119	396,787	1,013,412
2013	3,721	15,561	32,483	52,485	78,753	118,168	175,291	263,242	404,110	1,040,960
2014	4,115	16,568	34,406	53,703	79,375	119,153	178,829	268,535	411,903	1,068,914
2015	3,760	16,008	33,827	52,674	77,704	116,751	177,149	269,126	414,819	1,132,561
2016	4,522	17,397	35,782	54,879	80,127	120,239	183,825	276,154	422,906	1,114,544
2017	3,973	16,568	35,481	55,240	80,787	121,625	190,535	290,940	450,994	1,168,854
2018	4,597	17,923	37,235	56,908	82,287	123,055	193,731	296,719	458,235	1,151,359
2019	4,593	18,112	37,863	57,284	81,937	121,563	192,372	297,988	460,476	1,140,281
2020	4,646	18,791	39,502	58,635	82,292	121,027	192,801	300,513	464,936	1,154,880
2021	5,073	17,990	36,045	54,192	77,386	113,980	185,584	294,296	456,212	1,183,503

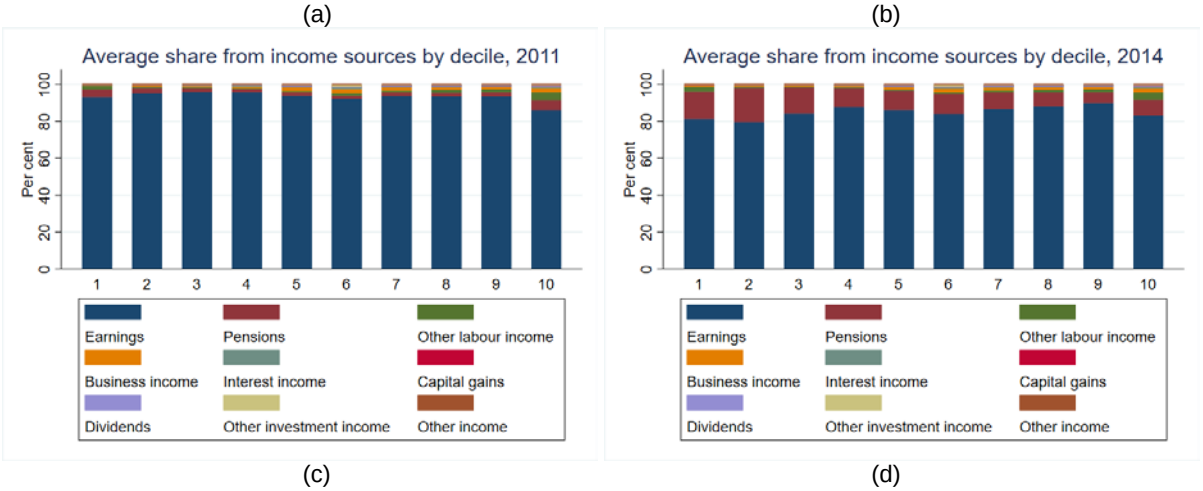
Note: all amounts expressed in December 2021 rands.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

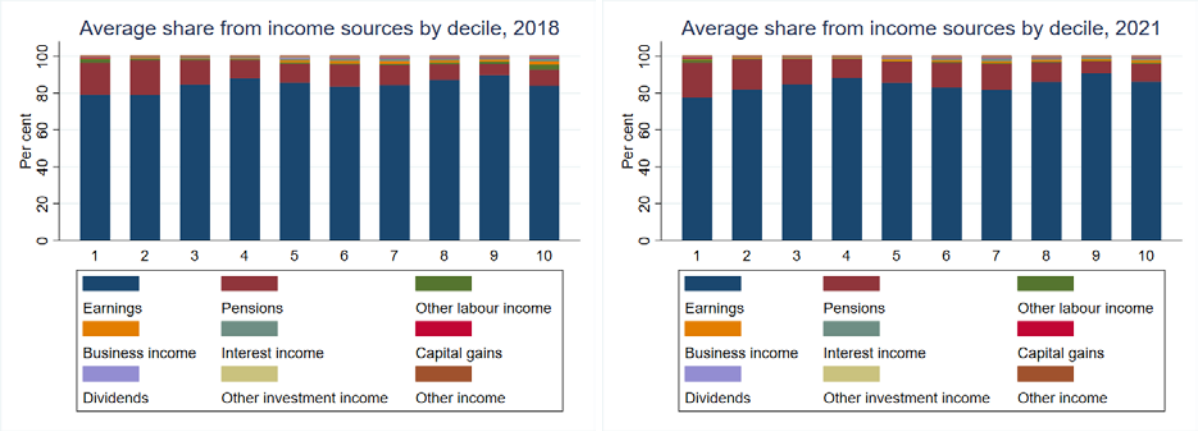
Sources of income

Figure 4 shows the average share of income from nine sub-categories of labour, capital, and other income. The graphs cover each decile in 2011, 2014, 2018, and 2021, respectively. Across all deciles and years, the predominant source of income is earnings followed by pension income.[19] As expected it is the higher deciles that, on average, receive income from more varied sources compared to the lower deciles. Given their prominence in debates over changes to their taxation the contributions from pensions and capital gains are particularly noteworthy. The dominance of earnings even in the top decile except in the top percentile is equally noteworthy.

Figure 4: Average share from income sources in (a) 2011, (b) 2014, (c) 2018, and (d) 2021



[19] One noticeable change is the increase in pension income and the decrease in 'earnings' for the lower deciles in 2014, 2018, and 2021 compared with 2011. This change may be due to the classification of pension income by SARS in 2011 and 2012: pension income was included in the 3601 code (which we classify as earnings), and it is difficult to separate these sources of income in these years. As such, the share of pension income in 2011 is probably too low.



Note: based on before-tax values.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

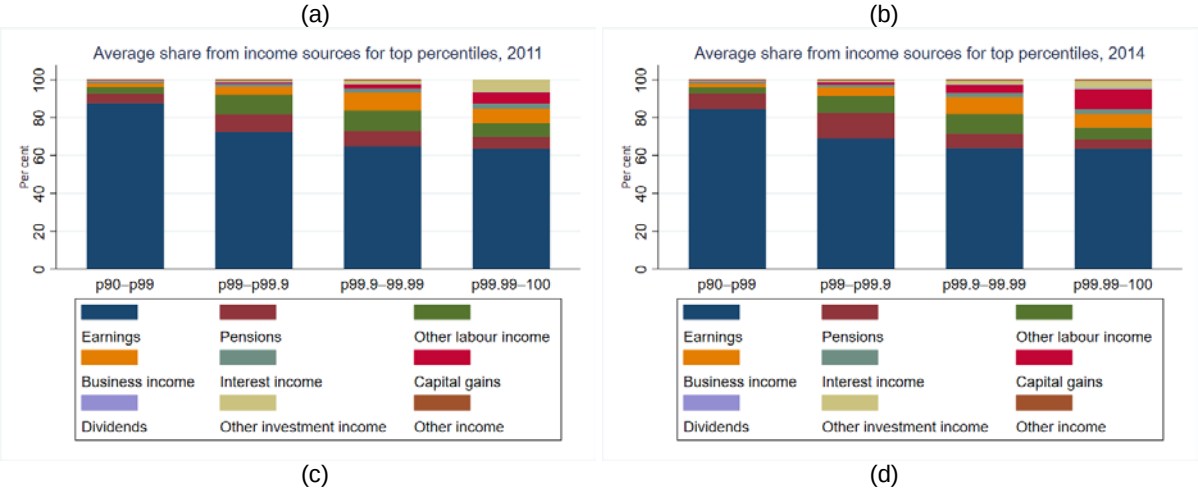
Sources of income for top-income groups

Figure 5 takes a closer look at the average shares of different income sources for the top income groups. The groups were examined using the following percentiles: 90–99, 99–99.9 (top 1 per cent), 99.9–99.99 (top 1 permille), and 99.99–100 (the highest income 10 per cent of the top per-mille).

Capital incomes contribute progressively more to total income for the wealthiest percentiles. Like Bassier and Woolard (2020), we find that earnings contribute over 80 per cent to income until the 99th percentile,[20] after which income from capital gains, interest and rent, other investment income, and business income become more important. In 2011 and 2014, the average share of earnings was above 60 per cent for the highest income group (the top 0.01 per cent); in 2018 and 2021 this dropped below 60 per cent. Dividends that would be captured on the ITR12 make up a small proportion of total income over the period, while the share of income from capital gains increased from 2011 to 2018 and decreased again in 2021.

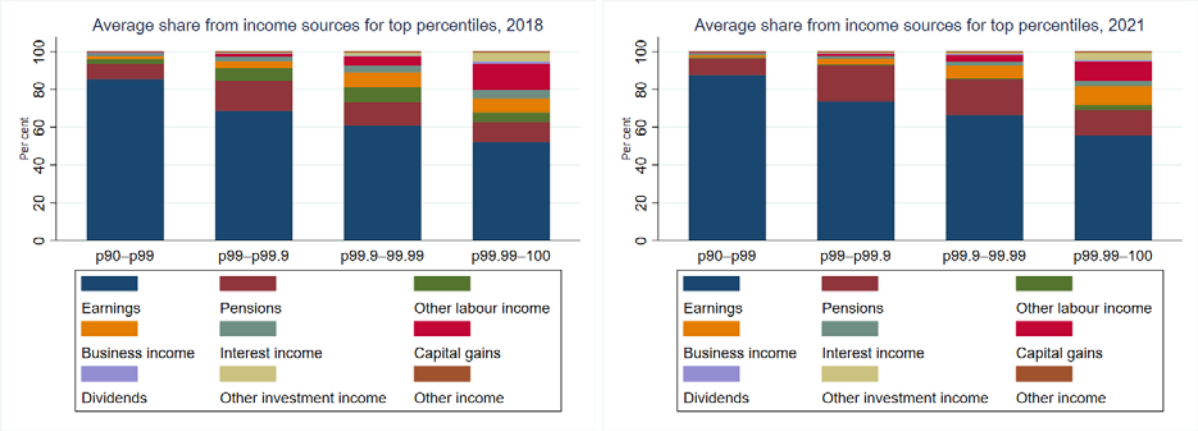
Figure 5: Average share from income source for top percentiles in (a) 2011, (b) 2014, (c) 2018, and (d) 2021

(a) (b)



(c) (d)

[20] Bassier and Woolard (2020) refer to 'salaries and bonuses', however, and not to 'earnings'.



Note: based on before-tax values.

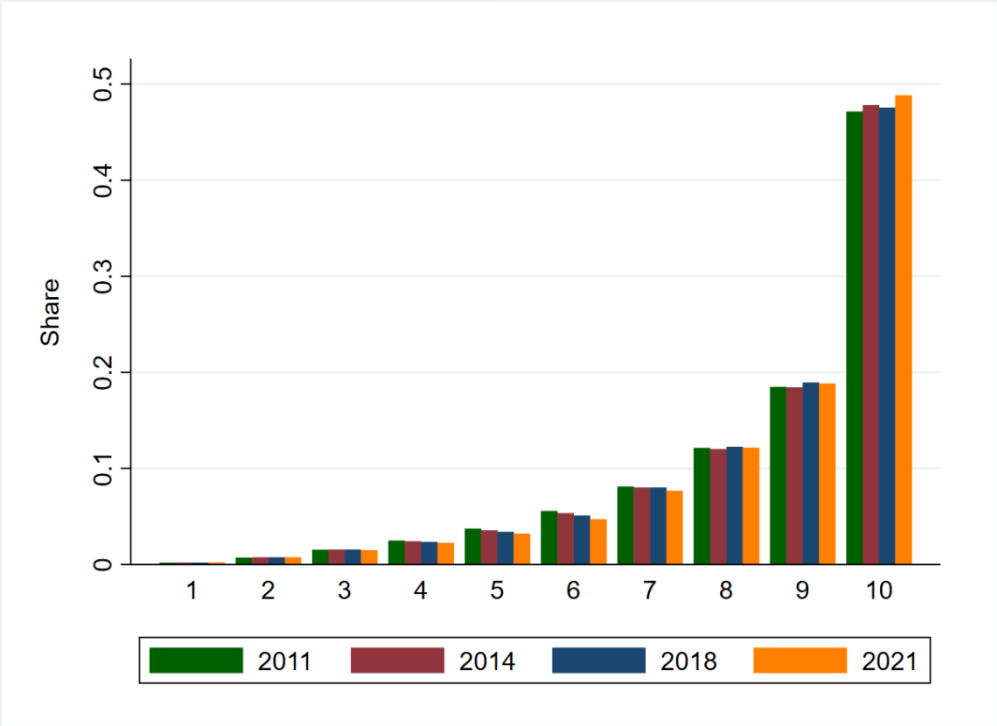
Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

3.2 Income shares and inequality measures

Income shares

Figure 6 illustrates the share of real gross income that went to each decile in 2011, 2014, 2018, and 2021. The top decile earned nearly half of all income across all years and is discussed in more detail below. The income shares of the lower deciles remained constant, whereas the income shares of the fifth, sixth, and seventh deciles appear to have decreased slightly from 2011 to 2021.

Figure 6: Income shares by decile for tax years 2011, 2014, 2018, and 2021



Note: based on before-tax values.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Leibbrandt et al. (2010) used survey data to show that per capita income had become increasingly concentrated in the top decile between 1993 and 2008. Our results from the tax data suggest a similar trend. The share of income received by the top decile in the tax data was around 47 per cent in 2011 and 49 per cent in 2021 (Table 7). While the income shares of the top 1, 0.1, and 0.01 per cent remained relatively constant between 2011 and 2020, our results suggest an increase in the shares of these income groups in 2021.

Table 7: Before-tax income shares (%) for top percentiles

Year	Top 10%	Top 1%	Top 0.1%	Top 0.01%
2011	47	14	4	1
2012	48	14	4	1
2013	48	14	4	1
2014	48	14	4	1
2015	49	16	5	2
2016	48	14	4	1
2017	48	15	5	2
2018	48	14	4	1
2019	47	13	4	1
2020	47	14	4	2
2021	49	16	6	4

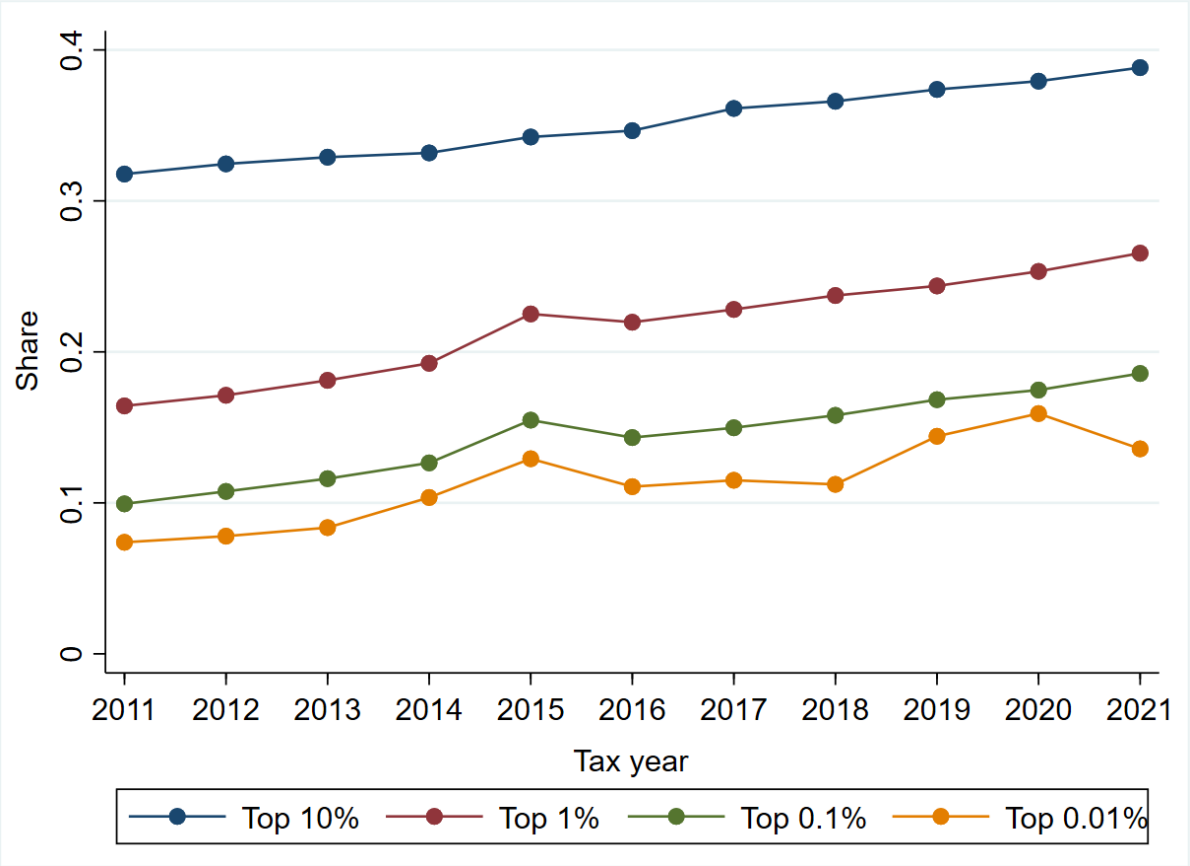
Note: income shares calculated using real gross income.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

The tax data also allow gendered distributional analysis, and we turn to this analysis now with particular attention to the top decile. The share of women in the top income groups has increased over time, as Figure 7 shows. The share of women in the top 10 per cent increased from 30 per cent in 2011 to 37 per cent in 2021. At the same time, the share of women in the top 1 per cent increased from just over 14 per cent to 25 per cent. The share of women in the top 0.01 per cent doubled from about 8 to 17 per cent over the period. Each of these are increases off a very low base.

In 2011, men outnumbered women by four to one in the top 1 per cent, by seven to one in the top 0.1 per cent, and by nine to one in the top 0.01 per cent. In 2021, this changed to three to one, four to one, and six to one in the top 1, 0.1, and 0.01 percentiles, respectively.

Figure 7: Share of women in top income groups, 2011–21



Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Looking at the trend of top income shares for men and women separately, Table 8 shows that the income shares of the women's top-end were lower than the men's. That said, the income shares for the top 10 per cent of both women and men increased over the period, despite some fluctuations between 2011 and 2021.

Table 8: Top income shares for men and women, 2011–21

	Men				Women			
Year	Top 10	Top 1	Top 0.1	Top 0.01	Top 10	Top 1	Top 0.1	Top 0.01
2011	0.48	0.14	0.04	0.01	0.41	0.10	0.02	0.01
2012	0.49	0.14	0.04	0.01	0.42	0.10	0.02	0.01
2013	0.49	0.15	0.04	0.01	0.42	0.10	0.02	0.01
2014	0.49	0.15	0.04	0.01	0.43	0.11	0.03	0.01
2015	0.51	0.16	0.06	0.02	0.44	0.12	0.03	0.01
2016	0.49	0.15	0.05	0.02	0.43	0.11	0.03	0.01
2017	0.49	0.15	0.04	0.01	0.44	0.11	0.03	0.01
2018	0.48	0.14	0.04	0.01	0.43	0.11	0.03	0.01
2019	0.48	0.14	0.04	0.01	0.44	0.11	0.03	0.01
2020	0.48	0.15	0.05	0.02	0.44	0.11	0.03	0.01
2021	0.50	0.18	0.08	0.06	0.44	0.11	0.03	0.01

Note: income shares calculated using real gross income.

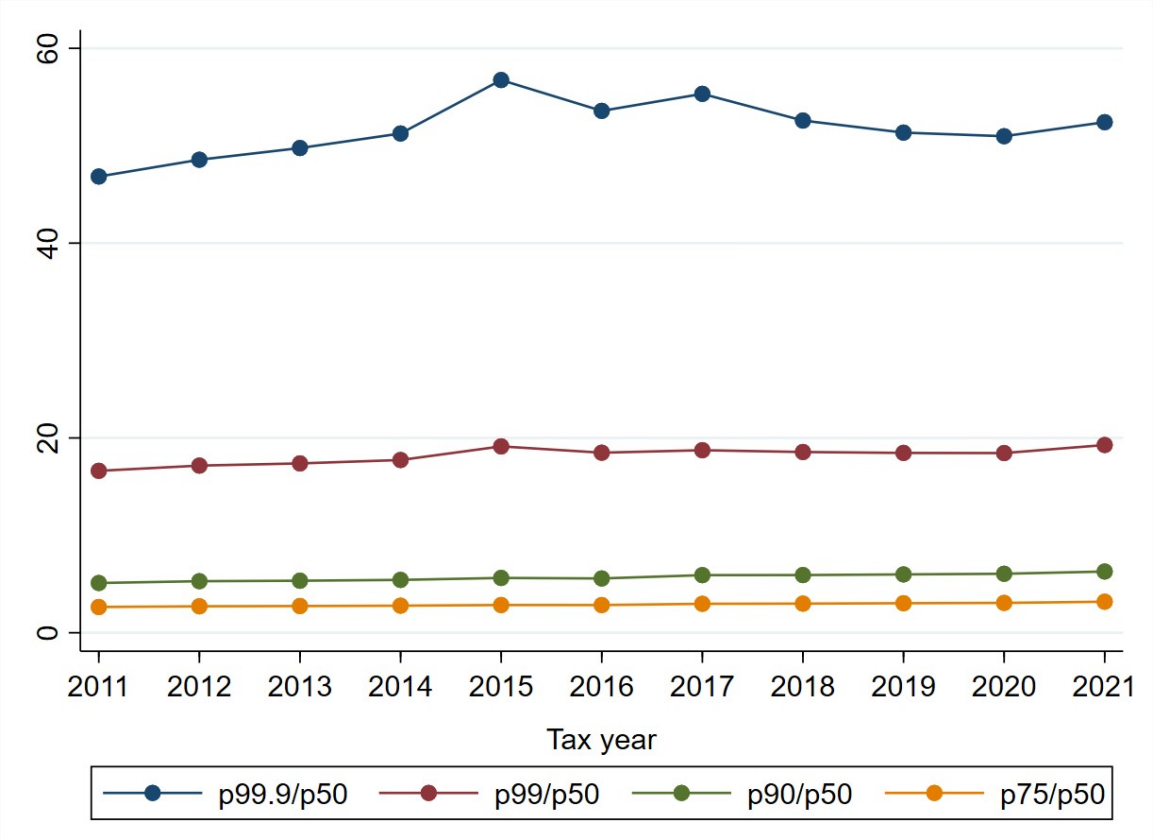
Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

We also find that the top-end income shares of those aged between 16–30 and 31–45 years increased over time, whereas the top-end income shares of those aged between 46–60 and above 61 years remained fairly constant (see Appendix Table A2).

Inequality measures

Figure 8 plots changes in the percentile ratios for incomes in the tax data over time. Income at the 99th percentile is over 15 times the income at the 50th percentile every year, while income at the 99.9th percentile is over 46 times the income at the 50th percentile. Although all percentile ratios increased over the period, the 90/50 and 75/50 ratios have increased only slightly. The ratios between the top percentile and the 50th percentile increase more sharply.

Figure 8: Changes in key percentile ratios (p99.9/p50, p99/p50, p90/p50 and p75/p50), 2011–21



Note: percentile ratios calculated using real gross income.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Table 9 presents various inequality measures for before-tax (gross) income for those with IRP5s and individuals who submit their own tax returns. The 90/50 ratio shows the biggest increase in inequality, reinforcing the finding above that the top end of the distribution moved away from the middle. The indices that look across the entire distribution do not report uniformly increasing inequality across the whole distribution. The Gini coefficients suggest that income inequality remained relatively stable, despite slight increases between 2015 and 2017 and in 2021. In addition, we report generalized entropy indices and Atkinson measures of inequality. The generalized entropy index with the parameter set to zero (GE(0)), also known as mean log deviation, is more sensitive to the changes in the distribution at the bottom than GE(1), Theil's *T* index. These measures also suggest an increase in inequality during the middle and end of the analysis period. The Atkinson index also reveals this with inequality aversion set to 0.5 or 1, A(0.5) and A(1). Those

measures showing the highest inequality aversion (GE(0) and A(2)) show the most stability in measured inequality showing higher increases in relative real incomes at the bottom of the distribution than those in the middle sections of the distribution. It would be useful to see data from 2022 and 2023 before making too much of the uniform increase in inequality between 2020 and 2021.

Table 9: Before-tax measures of inequality

Year	90/50 ratio	Gini	GE(0)	GE(1)	A(0.5)	A(1)	A(2)
2011	5.11	0.64	1.01	0.81	0.35	0.63	0.98
2012	5.30	0.64	1.03	0.82	0.36	0.64	0.98
2013	5.34	0.64	1.02	0.83	0.36	0.64	0.99
2014	5.43	0.64	1.01	0.83	0.36	0.63	0.99
2015	5.63	0.65	1.05	0.93	0.37	0.65	0.99
2016	5.58	0.65	1.01	0.85	0.36	0.64	0.99
2017	5.92	0.65	1.05	0.90	0.37	0.65	0.99
2018	5.93	0.64	1.01	0.83	0.36	0.64	0.99
2019	5.99	0.64	1.00	0.82	0.36	0.63	0.99
2020	6.06	0.64	1.00	0.86	0.36	0.63	0.99
2021	6.29	0.66	1.03	1.11	0.38	0.64	0.99

Note: measures calculated using real gross income. GE(0) and GE(1) refer to the generalized entropy indices, with the parameter (often α) set to 0 or 1, respectively. A(0.5), A(1), and A(2) refer to the Atkinson index, with inequality aversion set to 0.5, 1, and 2, respectively.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Decomposing inequality in gross income by income source

Tables 10 and 11 show overall decompositions of income inequality by income sources. The main sources we consider are labour, capital, and other income.[21] The share of labour income is over 97 per cent of total income. While capital income is more unevenly distributed than labour income (Table 11), its importance for overall inequality is limited because of its small share across the distribution. When we decompose income inequality using the GE(1) index, we find that about 89 per cent of income inequality is due to labour income (relative contribution). Based on the Gini index decomposition, 97 per cent of inequality is due to labour income. The results reported here are broadly consistent with those reported by Leibbrandt et al. (2010) and Hundenborn et al. (2018), who use household survey data in their analyses. However, as these studies reach the poorest deciles, they capture the importance of state social grants in the lowest three deciles and therefore lower the contribution of labour income to inequality somewhat.

[21] See Appendix Table A1 for more details on these categories, their sub-categories, and source codes.

Table 10: Decomposition by income source using the GE(1) index

Sources	Income share	Absolute contribution	Relative contribution
Labour income	0.974	0.866	0.893
Capital income	0.026	0.074	0.076
Other income	0.000	0.030	0.031
Total	1.000	0.970	1.000

Note: the decomposition of real gross income inequality is done using the Shapley approach. The absolute contribution reflects how much each income source adds to the overall index, while the relative contribution divides this by the overall index.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Table 11: Decomposition of Gini index by income sources

	Income shares	Gini coefficient (g)	Gini correlation (r)	Contribution ($s \times g \times r$)	% Contribution ($s \times g \times r$/G)
Labour income	0.976	0.649	0.994	0.630	0.971
Capital income	0.024	0.991	0.781	0.019	0.029
Other income	0.000	1.000	0.719	0.000	0.000
Total	1.000	0.649	1.000	0.649	1.000

Note: Gini's for each income group are given by g, and the correlation of the income source with overall income is represented by r. Contribution to overall inequality is given by $s \times g \times r$ and the relative contribution by dividing this by overall Gini.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

4 Redistribution via the tax system

In this section, we discuss tax rules in South Africa and how the average tax rate develops over the income distribution. This is useful in exploring the differences between inequality before and after taxes.

South Africa has a progressive income tax schedule with unchanged rates until 2014–15, with only inflation adjustments to the tax brackets (Table 12). In 2015–16, all but the first marginal tax rates were increased by one percentage point, and the highest marginal tax rate was raised from 41 to 45 per cent in 2017/18 for incomes exceeding 1.5 million a year. In addition, there is a tax rebate for those below 65 years of age, which stood at about ZAR14,000 in 2018. The rebate implies that incomes below approximately ZAR75,000 remain tax-free. The rebate has been in place throughout the years, with greater values for those above 65 and 75 years.

Table 12: Tax rates for selected years

2014–15		2015–16		2017–18	
Bracket	MTR (%)	Bracket	MTR (%)	Bracket	MTR (%)
1–174,550	18	1–181,900	18	1–189,880	18
174,551–272,700	25	181,901–284,100	26	189,881–296,540	26
272,701–377,450	30	284,101–393,200	31	296,541–410,460	31
377,451–528,000	35	393,201–550,100	36	410,461–555,600	36
528,001–673,100	38	550,101–701,300	39	555,601–708,310	39
673,101–	40	701,301–	41	708,311–1,500,000	41
				1,500,001–	45

Source: data from South African Revenue Service (2024).

Table 13 shows that taxes have the effect of reducing inequality but even after taxes, income inequality remains high, with the after-tax Gini between 0.60 and 0.62 across years. There is some indication of a very small impact of the increased progressivity on inequality; for example, the A(0.5) measure for after-tax income declines in 2018 while the corresponding measure for pre-tax income does not.

Table 13: After-tax measures of inequality, 2011–21

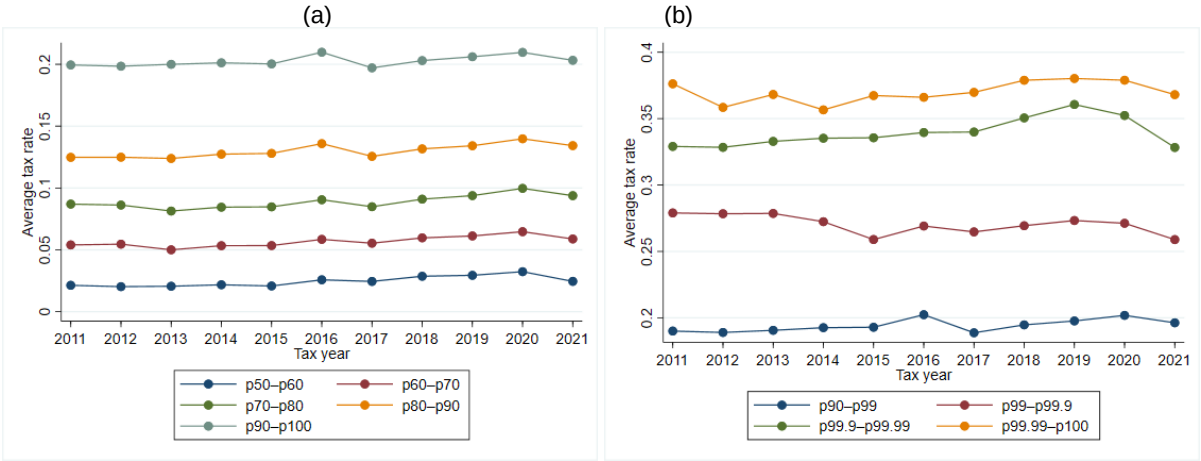
Year	90/50 ratio	Gini	GE(0)	GE(1)	A(0.5)	A(1)	A(2)
2011	4.37	0.60	0.90	0.69	0.32	0.59	0.96
2012	4.53	0.61	0.92	0.70	0.32	0.60	0.97
2013	4.57	0.61	0.91	0.70	0.32	0.60	0.98
2014	4.62	0.61	0.90	0.71	0.32	0.59	0.98
2015	4.78	0.62	0.94	0.79	0.33	0.61	0.98
2016	4.68	0.61	0.90	0.72	0.32	0.59	0.98
2017	5.06	0.62	0.94	0.75	0.33	0.61	0.98
2018	5.05	0.61	0.90	0.71	0.32	0.59	0.98
2019	5.09	0.61	0.90	0.70	0.32	0.59	0.98
2020	5.12	0.61	0.90	0.75	0.32	0.59	0.99
2021	5.33	0.62	0.93	1.06	0.34	0.60	0.99

Note: measures calculated using real disposable income (ZAR).

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

We define the average tax rate from the data as real tax liability divided by real gross income for different income groups. Figure 9 depicts the average tax rates for different groups. Figure 9a examines the top five deciles. The average tax rate for those between the 50th and 60th, 60th and 70th, and 70th and 80th percentiles is strikingly low. This is because people in these income groups have low incomes (e.g., incomes in the fifth decile were close to 80,000 in 2021, whereas the tax threshold stood at 83,000 for people below 65 years of age).

Figure 9: Average tax rates for (a) the top five deciles and (b) the top decile, 2011–21



Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Figure 9b shows the average tax rate for the top income groups. We observe that average tax rates increased slightly for those between the 90th and 99th percentiles, the 99.9th and 99.99th percentiles, and those at the very top. Average tax rates for the top 1 per cent appear to have

decreased in the final years in the data. This is flagged for further detailed interrogation in a future study.

5 Long-term inequality and income mobility

Income inequality at a given time does not tell us about the extent of mobility within society. In this section, we exploit the panel dimension of the tax administrative data and compare the Gini coefficients of annual real gross income to the Gini coefficient based on income averaged over several years. We also present transition matrices to measure mobility for the tax-compliant population in South Africa.

5.1 Annual and long-term Gini coefficients

If there is considerable mobility, individuals do not remain in the same place in the income distribution over time. In that case, the level of inequality in a given year may overstate the extent of long-run inequality of averaged incomes.

Table 14 presents inequality indices for a balanced sample that is restricted to individuals who are observed in each year. Gini coefficients and Theil indices are shown for each year over the 2011–21 period, the averages of these, and Gini and Theil indices for average incomes for each individual. The results suggest that income inequality based on average income is less than the average of annual inequality indices would indicate, but not very significantly so. This implies that the extent of income mobility in the balanced panel is low.

Table 14 Inequality in average incomes

	Gini	Theil
2011	0.555	0.593
2012	0.541	0.565
2013	0.536	0.560
2014	0.532	0.553
2015	0.537	0.579
2016	0.531	0.558
2017	0.531	0.563
2018	0.529	0.541
2019	0.526	0.526
2020	0.529	0.533
2021	0.546	0.563
Mean of annual indices	0.536	0.558
Inequality using average income	0.504	0.482

Note: coefficients calculated using real gross income (ZAR).

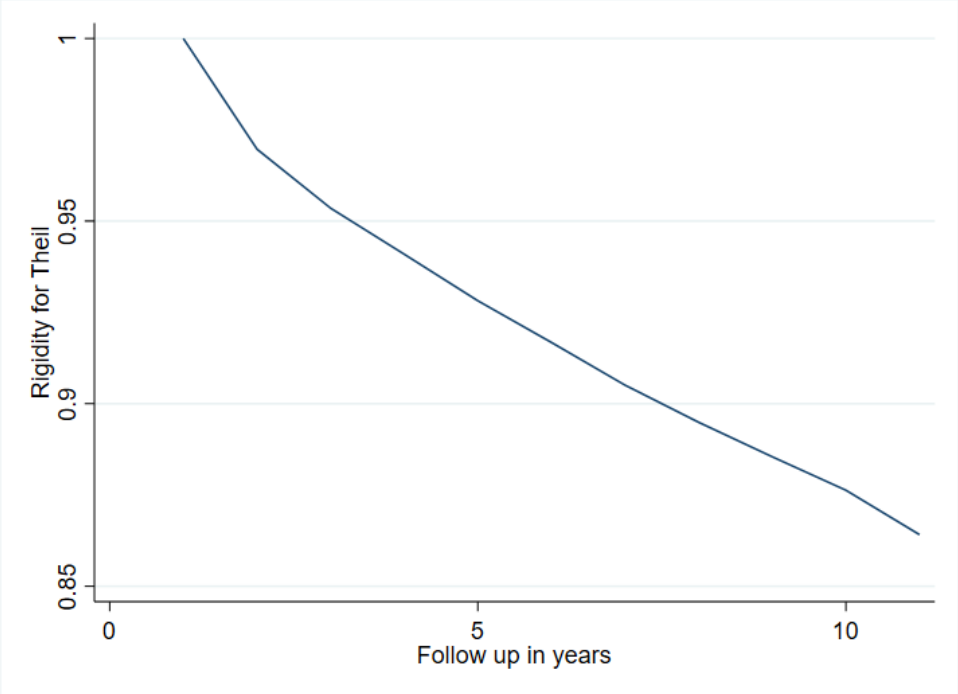
Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

This is further corroborated by the results of Figure 10, which depicts the Shorrocks (1978) rigidity index, calculated as

$$R = \frac{I(y_m)}{\sum_t w_t I(y_t)},$$

where $I(y_m)$ is an inequality index (Theil in our case) calculated using the average income for the individual across years, whereas $I(y_t)$ represents the same index calculated using annual incomes and w_t are weights (mean income in year t / total mean income over the years) that sum to one.

Figure 10: Shorrocks rigidity index for gross income over different follow-up periods, 2011–21



Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

The rigidity index compares the Theil index calculated for 2011 incomes to inequality based on mean incomes for different periods (comparing with 2011–12 mean income inequality for follow-up of 2 years, 2011–13 mean income inequality for follow-up of 3 years, etc.). The index is equal to unity if there is no difference in long-run versus short-run inequality. Figure 10 illustrates the relative reduction in inequality when the observation period is extended. Income inequality using the Theil index for mean incomes across 8 years is still close to 90 per cent of the inequality of the mean of the annual Theil indices. This rigidity is at the same level as in the United States, a very immobile society (Jäntti and Jenkins 2015: Fig. 10.8).

To reiterate, this analysis is restricted to those who were in the panel in 2011 and remained in the panel for each year through 2021. It ignores those who left the panel, implying potentially large negative income changes, and new entrants, implying potentially large positive income changes for these individuals. To assess the sensitivity of the mobility picture to these dynamics, the next sub-section looks at a set of transition matrices covering different periods between 2011 and 2021.

5.2 Transition matrices

Table 15 reports a transition matrix for individuals present in the tax data at the start and end of the period (i.e. in 2011 and 2021). They are included even if they dropped out of the panel at some point between these 2 years as long as they had transitioned back into the panel by 2021.

Apart from the top decile, between 22 and 36 per cent of individuals remained in the same decile 10 years later. More than half the number of individuals who were in the top decile remained in the top decile. In the middle of the distribution (deciles 4–7) as many individuals moved to a higher

decile than to a lower decile. This is a long-run view and a lot can happen in-between that would shape the average income of each person. As before, if a person drops out of the panel and does not return by 2021, they will be lost to the panel. Also, no post-2011 entrants will be reflected in the panel.

Table 15: Decile mobility matrix, 2021 compared with 2011

	Decile in 2021										
Decile in 2011	1	2	3	4	5	6	7	8	9	10	Total
1	32	17	14	10	7	6	6	4	2	2	100
2	22	28	16	10	7	6	5	3	2	1	100
3	12	21	25	16	9	6	4	3	2	1	100
4	9	10	19	25	15	9	7	4	2	1	100
5	7	7	8	16	22	17	11	6	4	2	100
6	5	6	6	7	15	22	20	10	6	3	100
7	4	4	4	5	8	14	23	24	10	5	100
8	3	3	3	4	6	7	11	30	25	8	100
9	3	2	2	3	6	8	8	11	36	21	100
10	3	2	2	3	4	5	6	7	13	55	100

Note: unbalanced panel; rows may not add up to 100 per cent due to rounding.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

To gain a sense of the changes in mobility between 2011 and 2021, Table 16 shows the decile mobility matrices considering 3-year lags. For comparability, we restrict the analysis to those who are in the tax data throughout the period (i.e. the balanced panel that we used earlier for the Shorrocks rigidity index). We start with the most recent year first with Panel A reporting the transition matrix for 2018 as the base period and 2021 as the end period. Panel B shows the transition matrix for 2015 as the base period and 2018 as the end period. The transition matrix for 2014 compared with 2011 is reported in Panel C. The annual transition matrices are presented in Appendix Tables A3–A12.

Table 16: Decile mobility matrix

Panel A: 2021 versus 2018											
	Decile in 2021										
Decile in 2018	1	2	3	4	5	6	7	8	9	10	Total
1	65	24	5	3	1	1	1	0	0	0	100
2	13	50	28	5	2	1	1	0	0	0	100
3	7	11	45	27	5	2	1	1	0	0	100
4	5	5	10	43	27	6	2	1	1	1	100
5	3	3	4	10	42	27	5	2	1	1	100
6	2	2	2	4	10	44	25	5	3	1	100
7	1	1	2	3	4	9	48	24	5	2	100
8	1	1	1	2	3	4	10	53	21	5	100
9	1	1	1	2	3	3	4	10	58	17	100
10	1	1	1	2	2	3	3	4	11	73	100
Panel B: 2018 versus 2015											
	Decile in 2018										
Decile in 2015	1	2	3	4	5	6	7	8	9	10	Total
1	72	14	6	3	2	1	1	0	0	0	100
2	12	61	17	5	2	1	1	0	0	0	100
3	4	14	54	18	5	2	1	1	0	0	100

4	3	3	13	48	21	7	3	1	1	0	100
5	2	2	3	15	46	22	6	2	1	0	100
6	2	2	2	3	14	45	24	5	2	1	100
7	1	1	1	2	3	13	46	24	5	2	100
8	1	1	1	2	2	3	12	50	23	4	100
9	1	1	1	2	2	3	4	12	56	18	100
10	1	1	1	2	2	2	3	4	11	74	100

Panel C: 2014 versus 2011

	Decile in 2014										
Decile in 2011	1	2	3	4	5	6	7	8	9	10	Total
1	54	16	10	7	5	3	2	1	1	0	100
2	28	42	12	7	4	3	2	1	1	0	100
3	6	31	38	11	6	3	2	1	1	0	100
4	3	4	30	37	12	7	3	2	1	1	100
5	2	2	4	28	38	14	6	3	2	1	100
6	2	1	2	5	27	40	14	5	3	1	100
7	1	1	1	2	4	23	47	14	6	2	100
8	1	1	1	1	2	3	20	53	14	4	100
9	1	0	1	1	1	2	3	18	59	14	100
10	1	0	1	1	1	1	2	2	13	78	100

Note: balanced panel; rows may not add up to 100 per cent due to rounding.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

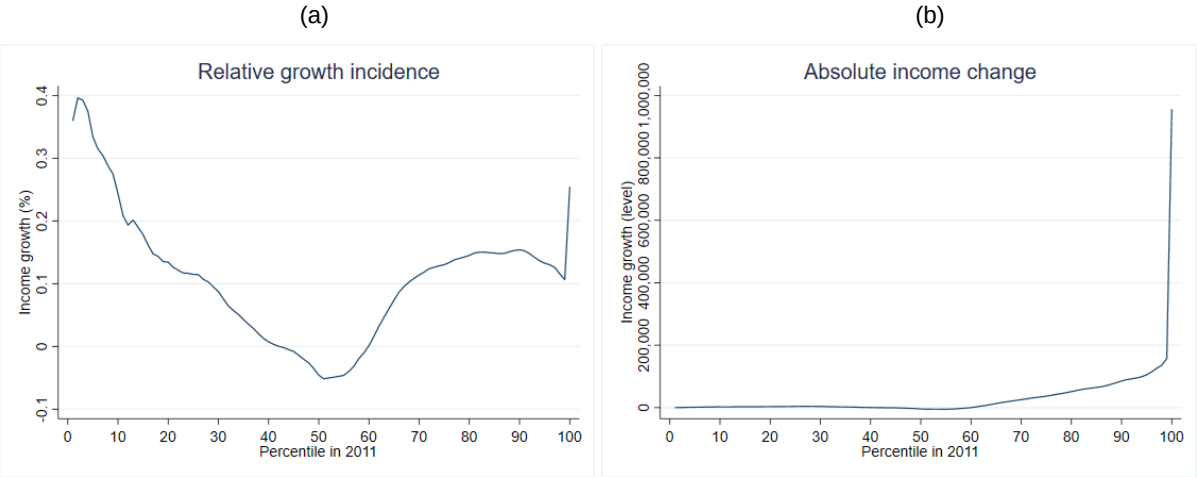
Reflecting low mobility out of the bottom and top deciles, 65 per cent of individuals in the bottom decile in 2018 remained there in 2021, while 73 per cent of individuals in the top decile in 2018 were in the top decile in 2021. This immobility is somewhat higher for the period 2015–18. Almost 74 per cent of individuals in the wealthiest decile in 2015 were in the top decile in 2018. At the same time, 72 per cent of those in the poorest decile in 2015 were there in 2018. There is even higher immobility at the top end, with about 78 per cent of those in the top decile in 2014 also in the top decile in 2011. However, there is higher mobility out of the bottom decile in 2011–14 than in the two later periods, with about 54 per cent of those in the bottom decile in 2011 being in the bottom decile in 2014.

On aggregate, this suggests that immobility at the top decreased slightly over the period under consideration, while immobility at the bottom increased in the middle of the period. In each sub-period there is notably more mobility up and down in the middle deciles of the income distribution than in the tails. This sub-period analysis and the annual transition matrices in Appendix A indicate that the tax data are picking up the changes in relative mobility over the period that give texture and timing to the higher net mobility reflected in the longer-run transition. The Shorrocks mobility index makes the key point that these relative mobility matrices reflect low mobility of real incomes.

5.3 Growth incidence

Income developments over a longer period can also be illustrated with the help of growth incidence curves. Figure 11 depicts the relative and absolute income growth for people in different income percentiles. These are anonymous growth incidence curves, implying that different cross sections are compared in 2011 and 2021. The relative income growth has a U shape, but the absolute gains concentrate to the top groups.

Figure 11: Relative (a) and absolute (b) income growth by percentile, 2011–2021



Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

5.4 Mobility at the top

Table 17 presents a mobility matrix for individuals in the top income groups at the start and end of the period. Ninety-five per cent of those whose incomes were below the 90th percentile in 2011 remained below the top 10 per cent also in 2021. More than a quarter of individuals who were at the very top of the income distribution (in the top 0.01 per cent) in 2011 stayed at the very top in 2021. Most downward mobility for this group was into the top 1 per cent. Thus, 55 per cent of this group remained in the top percentile and less than a quarter dropped below the 10th decile. Less than a third of those in the top percentile in 2011 had moved down below the top 10th decile in 2021. Roughly half of those who were in the 90th to 99th percentile remained there.

Thus, there is very low upward mobility into the top decile and very low downward mobility out of the top decile for those who were at the top end in 2021. Nonetheless, there is some mobility within the 10th decile that can be and might be important to interrogate in the light of the impact of specific economic events or policy proposals.

The annual transition matrices for the top income groups are shown in Appendix Tables A13–A22.

Table 17: Mobility in top income groups, 2021 versus 2011

	Income group in 2021				
Income group in 2011	p0–p90	p90–p99	p99–p99.9	p99.9–p100	Total
p0–p90	95	5	0	0	100
p90–p99	46	49	4	0	100
p99–p99.9	31	34	30	4	100
p99.9–p100	22	21	29	27	100

Note: unbalanced panel; rows may not add up to 100 per cent due to rounding.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

To give a sense of these changes over time, between 2011 and 2021, we consider the mobility of those in the top income groups in 3-year lags in Table 18. This reveals more movement in the most recent period: 43 per cent of those at the very top in 2018 were still at the very top in 2021 compared with 53 per cent of those in the top income group in 2014 and 2011. Importantly, the

sub-periods reveal that there is more mobility within the 10th decile than the longer-run picture of 2011–21 showed. But the very low mobility into and out of the top decile is consistent.

Table 18: Mobility in top income groups

Panel A: 2021 versus 2018

	Income group in 2021				
Income group in 2018	p0–p90	p90–p99	p99–p99.9	p99.9–p100	Total
p0–p90	97	3	0	0	100
p90–p99	28	68	3	0	100
p99–p99.9	19	30	47	4	100
p99.9–p100	14	12	31	43	100

Panel B: 2018 versus 2015

	Income group in 2018				
Income group in 2015	p0–p90	p90–p99	p99–p99.9	p99.9–p100	Total
p0–p90	97	3	0	0	100
p90–p99	27	69	4	0	100
p99–p99.9	17	26	52	4	100
p99.9–p100	10	12	30	48	100

Panel C: 2014 versus 2011

	Income group in 2014				
Income group in 2011	p0–p90	p90–p99	p99–p99.9	p99.9–p100	Total
p0–p90	98	2	0	0	100
p90–p99	24	73	3	0	100
p99–p99.9	9	29	58	4	100
p99.9–p100	6	9	32	53	100

Note: balanced panel; rows may not add up to 100 per cent due to rounding.

Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

6 Conclusion

In this paper, we make use of individual-level tax administrative data and describe changes in the income distribution of formal sector workers and those who file their own tax returns (the tax-compliant population) in South Africa for the period 2011–21. The tax data do not cover individuals at the bottom of the South African income distribution who do not have any taxable income. This needs to be kept in mind when interpreting the results. Nonetheless, the data reach further down the personal income distribution than the personal income tax filing threshold because they include information that employers are required to submit to SARS authority for each of their formal sector employees who were paid more than ZAR2,000 in a tax year.

We present baseline inequality measures and find that income inequality among our sample remained stable at a very high level over the period. Our estimates are in line with estimates from survey data. Taxes have the effect of reducing inequality, but even after tax, inequality remains high. Those with higher incomes tend to have larger tax liabilities, and the average tax rate for individuals in the top 1 per cent is between 30 and 40 per cent. This average rate appears to have fallen marginally in the last few years.

The aggregate metric masks some notable changes as it is driven by relatively stronger real income growth at both the top and the bottom ends of the distribution with very sluggish real income growth in the middle deciles. The real income growth at the bottom comes off very low base incomes and these real incomes are still very low. The strong growth real incomes at the top end see the share of top decile in total income rising to just below 50 per cent by 2021. With the share of the ninth decile rising to just under 20 per cent in 2021, the limiting texture of South Africa's very high inequality is made clear.

Labour income is the overwhelmingly dominant income source and source of income inequality, accounting for over 97 per cent of total income on average and 90 per cent of income inequality. The average share of income received from capital sources increases as incomes increase. But only in the top percentile do these other income sources (particularly pension income and capital gains) constitute up to 40 per cent of total income.

These tax administrative data are a panel of tax-compliant individuals. We exploit the longitudinal nature of the data to examine income mobility within these top income groups. We find low levels of upward mobility, driven by very little upward mobility into the top decile and little downward mobility out of this decile, particularly from those in the top percentile in 2011. There are also low levels of upward mobility out of the bottom deciles. There is more upward and downward mobility in the middle deciles, mirroring the sluggish and somewhat volatile real income growth for those in the middle of the income distribution.

This paper reflects much hard work by the UNU-WIDER team to strengthen the underlying data. There is more work to be done. Going forward, it would be important to be able to utilize tax administrative data covering all dividends and interest received and also incorporate, if possible, the beneficiaries of income held in trusts. These data sources would enable researchers and users of income distribution analysis to obtain a full picture of the distribution of capital incomes, which would also be influential for the top-income shares.

References

Araar, A., and J.Y. Duclos (2022). 'User manual DASP version 3.03'. Université Laval, Partnership for Economic Policy and World Bank. Available at: https://www.pep-net.org/news/new-version-distributive-analysis-stata-package-launched (accessed in September 2024).

Bassier, I., and I. Woolard (2020). 'Exclusive Growth?: Rapidly Increasing Top Incomes Amidst Low National Growth in South Africa'. WIDER Working Paper 2020/53. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/810-8

Card, D., R. Chetty, M.S. Feldstein, and E. Saez (2010). 'Expanding Access to Administrative Data for Research in the United States'. American Economic Association, Ten Years and Beyond: Economists Answer NSF's Call for Long-Term Research Agendas. https://doi.org/10.2139/ssrn.1888586

Cichello, P., and M. Rogan (2017). 'Informal Sector Employment and Poverty in South Africa: Identifying the Contributions of "Informal" Sources of Income on Aggregate Poverty Measure'. RED1 3x3 Working Papers 34. Cape Town: Southern Africa Labour and Development Research Unit.

Chatterjee, A., L. Czajka, and A. Gethin (2020). 'Estimating the Distribution of Household Wealth in South Africa'. WIDER Working Paper 2020/45. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/802-3

Chatterjee, A., L. Czajka, and A. Gethin (2021). 'Can Redistribution Keep Up with Inequality? Evidence from South Africa, 1993–2019'. World Inequality Lab Working Paper 2021/20. Available at: https://wid.world/news-article/can-redistribution-keep-up-with-inequality/ (accessed August 2024).

Chatterjee, A., L. Czajka, and A. Gethin (2022). 'Wealth Inequality in South Africa, 1993–2017'. *World Bank Economic Review*, 36(1): 19–36. https://doi.org/10.1093/wber/lhab012

Drechsel-Grau, M., A. Peichl, J.F. Schmieder, K.D. Schmid, H. Walz, and S. Wolter (2022). 'Inequality and Income Dynamics in Germany'. National Bureau of Economic Research Working Paper 29818. https://doi.org/10.3386/w29818

Ebrahim, A., and C. Axelson (2019). 'The Creation of an Individual Panel Using Administrative Tax Microdata in South Africa'. WIDER Working Paper 2019/27. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2019/661-6

Finn, A., and M. Leibbrandt (2018). The evolution and determination of earnings inequality in post-apartheid South Africa. WIDER Working Paper 2018/83. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2018/525-1

Hundenborn, J., M. Leibbrandt, and I. Woolard (2018). 'Drivers of Inequality in South Africa'. WIDER Working Paper 2018/162. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2018/604-3

Hundenborn, J., I. Woolard, and J. Jellema (2019). 'The Effect of Top Incomes on Inequality in South Africa'. *International Tax and Public Finance*, 26(5): 1018–47. https://doi.org/10.1007/s10797-018-9529-9

Jann, B. (2016). 'Assessing Inequality Using Percentile Shares'. *The Stata Journal*, 16(2): 264–300. https://doi.org/10.1177/1536867X1601600202

Jäntti, M., and S.P. Jenkins (2015). 'Income Mobility'. In A.B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution* (Vol. 2, Chapter 10, pp. 807–935). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-59428-0.00011-4

Jäntti, M., M.I. Nalukwago, and R. Waiswa (2022). 'Income Distribution in Uganda based on Tax Registers: What Do Top Incomes Say?'. WIDER Working Paper 2022/6. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2022/137-2

Jenkins, S. (2021). 'INEQDEC0: Stata Module to Calculate Inequality Indices with Decomposition by Subgroup'. Statistical Software Components from Boston College Department of Economics. https://econpapers.repec.org/software/bocbocode/s366002.htm

Kerr, A., and M. Wittenberg (2019). 'A Guide to PALMS version 3.3'. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/434/download/10286 (accessed 26 June 2022).

Kerr, A. (2020). 'Earnings in the South African Revenue Service IRP5 Data'. WIDER Working Paper 2020/62. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/819-1

Kerr, A. (2021). 'Measuring Earnings Inequality in South Africa Using Household Survey and Administrative Tax Microdata'. WIDER Working Paper 2021/82. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/020-7

Leibbrandt, M., and F. Díaz Pabón (2022). 'Inequality in South Africa'. In A. Oqubay, F. Tregenna, and V. Valodia (eds), *The Oxford Handbook of the South African Economy* (Chapter 9, pp. 175–92). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780192894199.013.9

Leibbrandt, M., I. Woolard, A. Finn, and J. Argent (2010). 'Trends in South African Income Distribution and Poverty Since the Fall of Apartheid'. OECD Social, Employment and Migration Working Papers 101. Paris: OECD Publishing. https://doi.org/10.1787/5kmms0t7p1ms-en

National Treasury and UNU-WIDER (2023). 'Individual Panel 2011–2022 [dataset]. Version 2023'. Pretoria: South African Revenue Service [producer of the original data], 2022. Pretoria: National Treasury and UNU-WIDER [producer and distributor of the harmonized dataset], 2023.

Redonda, A., and C. Axelson (2021). 'Assessing Pension-Related Tax Expenditures in South Africa: Evidence from the 2016 Retirement Reform'. WIDER Working Paper 2021/54. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/992-1

Shorrocks, A.F. (1978). 'Income Inequality and Income mobility'. *Journal of Economic Theory*, 19(2): 376–93.

South African Revenue Service (2024). Archive—Tax Rates. Available at: https://www.sars.gov.za/tax-rates/archive-tax-rates/ (accessed August 2024).

Southern Africa Labour and Development Research Unit (2018) National Income Dynamics Study 2017, Wave 5 [dataset]. Version 1.0.0 Pretoria: Department of Planning, Monitoring, and Evaluation [funding agency]. Cape Town: Southern Africa Labour and Development Research Unit [implementer], 2018. Cape Town: DataFirst [distributor], 2018. https://doi.org/10.25828/fw3h-v708

UNU-WIDER (2022). World Income Inequality Database (WIID). Version 30 June 2022. https://doi.org/10.35188/UNU-WIDER/WIID-300622

Van Kerm, P. (2020). 'SGINI: Stata Module to Compute Generalized Gini and Concentration Coefficients, Gini Correlations and Fractional Ranks'. Statistical Software Components from Boston College Department of Economics. https://econpapers.repec.org/software/bocbocode/s458778.htm

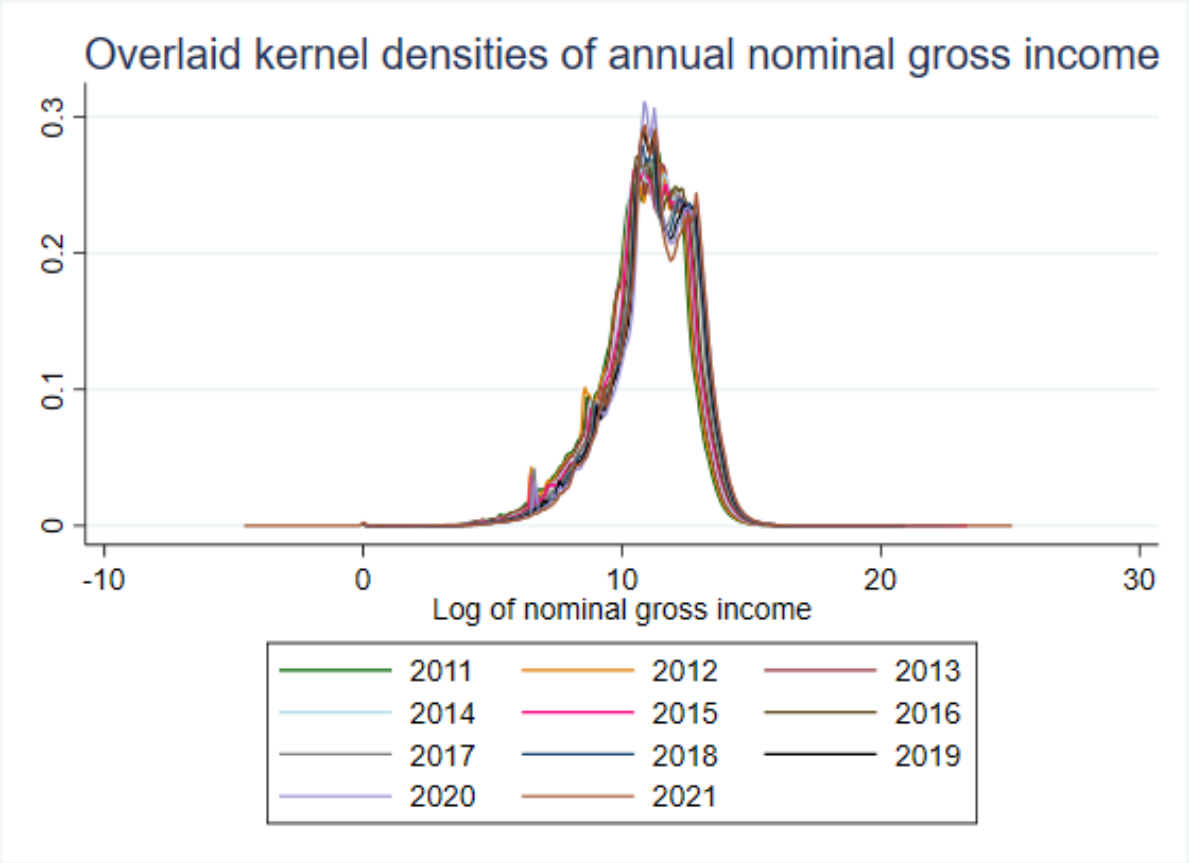
Wilkinson, R., and K. Pickett (2011). *The Spirit Level: Why Greater Equality Makes Societies Stronger*. New York: Bloomsbury Publishing.

Wittenberg, M. (2017a). 'Measurement of Earnings: Comparing South African Tax and Survey Data'. REDI3x3 Working Paper 41. Cape Town: SALDRU, University of Cape Town. Available at: https://www.siopsa.org.za/wp-content/uploads/2020/06/Wittenberg-2017-Measuring-earnings-tax-and-survey-data1.pdf (accessed August 2024).

Wittenberg, M. (2017b). 'Wages and Wage Inequality in South Africa 1994–2011: Part 2–Inequality Measurement and Trends'. *South African Journal of Economics*, 85(2): 298–318. https://doi.org/10.1111/saje.12147

Appendix A: Additional figures and tables

Figure A1: Overlaid kernel densities of nominal gross income over time



Source: authors' calculations based on the 2023 version of the Individual Panel (National Treasury and UNU-WIDER 2023).

Table A1: Income shares by decile, 2011–21

	Year										
Decile	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
3	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01
4	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
5	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
6	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
7	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
8	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
9	0.18	0.19	0.18	0.18	0.18	0.18	0.19	0.19	0.19	0.19	0.19
10	0.47	0.48	0.48	0.48	0.49	0.48	0.48	0.48	0.47	0.47	0.49

Notes: shares are calculated using real gross income (ZAR).

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A2: Income shares by age group

	Age 16–30				Age 31–45				Age 46–60					Age 61 +		
Year	Top 10	Top 1	Top 0.1	Top 0.01	Top 10	Top 1	Top 0.1	Top 0.01	Top 10	Top 1	Top 0.1	Top 0.01	Top 10	Top 1	Top 0.1	Top 0.01
2011	0.45	0.10	0.02	0.01	0.44	0.11	0.03	0.01	0.43	0.13	0.04	0.01	0.51	0.16	0.05	0.01
2012	0.45	0.10	0.02	0.00	0.44	0.12	0.03	0.01	0.43	0.13	0.04	0.01	0.52	0.16	0.05	0.01
2013	0.45	0.10	0.02	0.00	0.44	0.12	0.03	0.01	0.44	0.13	0.04	0.01	0.52	0.16	0.05	0.01
2014	0.45	0.10	0.02	0.01	0.44	0.12	0.03	0.01	0.44	0.13	0.04	0.01	0.52	0.17	0.05	0.02
2015	0.46	0.10	0.02	0.01	0.45	0.12	0.03	0.01	0.45	0.14	0.04	0.01	0.56	0.22	0.10	0.07
2016	0.45	0.10	0.02	0.01	0.44	0.12	0.03	0.01	0.45	0.14	0.04	0.01	0.53	0.17	0.05	0.02
2017	0.47	0.12	0.04	0.03	0.44	0.12	0.04	0.01	0.44	0.14	0.04	0.02	0.53	0.17	0.06	0.02
2018	0.46	0.10	0.02	0.01	0.43	0.11	0.03	0.01	0.43	0.13	0.03	0.01	0.53	0.17	0.05	0.02
2019	0.46	0.10	0.02	0.01	0.43	0.11	0.03	0.01	0.43	0.12	0.03	0.01	0.52	0.16	0.05	0.01
2020	0.49	0.14	0.07	0.05	0.43	0.11	0.03	0.01	0.43	0.12	0.03	0.01	0.52	0.16	0.05	0.02
2021	0.49	0.13	0.05	0.03	0.47	0.16	0.09	0.07	0.43	0.12	0.03	0.01	0.52	0.16	0.04	0.01

Notes: shares are calculated using real gross income (ZAR).

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A3: Decile mobility matrix, 2012 compared to 2011

	Decile in 2012										
Decile in 2011	1	2	3	4	5	6	7	8	9	10	Total
1	65.55	16.78	8.13	5.09	2.23	1.1	0.6	0.27	0.16	0.1	100
2	24.05	53.6	11.68	5.16	2.6	1.37	0.9	0.37	0.2	0.08	100
3	4.65	23.44	53.68	10.55	3.93	2.09	0.81	0.47	0.26	0.11	100
4	2.32	2.95	21.39	54.17	11.48	4.62	1.62	0.74	0.43	0.29	100
5	1.21	1.17	2.33	20.1	57.17	12.4	3.56	1.24	0.58	0.24	100
6	0.78	0.75	1.12	2.42	19	58.93	12.29	3.04	1.26	0.41	100
7	0.45	0.47	0.59	0.98	1.81	16.2	63.39	12.21	3.05	0.85	100
8	0.28	0.31	0.36	0.56	0.76	1.76	14.36	67.65	11.99	1.96	100
9	0.25	0.23	0.34	0.48	0.56	0.88	1.63	12.71	72.85	10.07	100
10	0.3	0.22	0.32	0.46	0.5	0.71	0.88	1.34	9.3	85.97	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A4: Decile mobility matrix, 2013 compared to 2012

	Decile in 2013										
Decile in 2012	1	2	3	4	5	6	7	8	9	10	Total
1	75.83	12.72	5.35	3.01	1.46	0.75	0.42	0.21	0.15	0.09	100
2	15.19	66	11.57	3.92	1.6	0.82	0.47	0.22	0.14	0.07	100
3	3.89	15.95	61.8	11.93	3.44	1.64	0.67	0.36	0.21	0.11	100
4	1.89	2.36	16.57	59.57	12.6	4.36	1.43	0.67	0.37	0.19	100
5	1.03	0.97	2	16.86	60.95	12.98	3.26	1.18	0.58	0.2	100
6	0.72	0.67	1.01	2.28	16.35	61.08	13.06	3.19	1.24	0.4	100
7	0.45	0.48	0.54	0.95	1.69	14.96	64.68	12.54	2.94	0.78	100
8	0.27	0.33	0.39	0.54	0.78	1.68	13.21	68.89	12.05	1.86	100
9	0.26	0.27	0.38	0.49	0.59	0.98	1.79	11.33	73.51	10.4	100
10	0.33	0.22	0.33	0.47	0.56	0.77	1.05	1.45	8.85	85.97	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A5: Decile mobility matrix, 2014 compared to 2013

	Decile in 2014										
Decile in 2013	1	2	3	4	5	6	7	8	9	10	Total
1	77.46	13.66	4.36	2.2	1.01	0.56	0.32	0.18	0.14	0.1	100
2	14.13	67.95	11.89	3.32	1.36	0.6	0.36	0.19	0.12	0.07	100
3	3.48	13.38	66.11	11.51	2.91	1.4	0.57	0.33	0.22	0.09	100
4	1.77	2.1	13	63.59	13.56	3.57	1.33	0.6	0.33	0.15	100
5	0.94	0.89	1.95	14.75	62.07	14.26	3.13	1.16	0.63	0.22	100
6	0.72	0.62	0.94	2.1	15.39	61.92	13.36	3.13	1.42	0.4	100
7	0.45	0.48	0.53	0.87	1.65	14.47	65.38	12.27	3.04	0.87	100
8	0.31	0.35	0.39	0.54	0.77	1.56	12.9	69.35	11.65	2.17	100
9	0.29	0.28	0.43	0.57	0.65	0.87	1.66	11.43	73.5	10.32	100
10	0.35	0.23	0.38	0.55	0.66	0.81	1.02	1.38	8.98	85.64	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual Panel (National Treasury and UNU-WIDER 2023).

Table A6: Decile mobility matrix, 2015 compared to 2014

	Decile in 2015										
Decile in 2014	1	2	3	4	5	6	7	8	9	10	Total
1	80.38	11.84	3.85	1.85	0.92	0.48	0.29	0.17	0.12	0.1	100
2	11.28	71.65	11.97	2.85	1.14	0.51	0.28	0.15	0.1	0.07	100
3	3.21	11.37	68.06	11.88	3.09	1.29	0.56	0.26	0.18	0.11	100
4	1.69	2.11	11.49	65.01	13.83	3.34	1.39	0.6	0.36	0.17	100
5	0.99	0.89	1.83	13.72	63.62	13.39	3.35	1.3	0.63	0.28	100
6	0.74	0.62	0.9	1.95	13.15	64.22	13.65	2.88	1.36	0.54	100
7	0.49	0.51	0.51	0.88	1.75	13.13	65.69	12.81	3.02	1.21	100
8	0.34	0.38	0.45	0.6	0.81	1.64	11.83	69.23	11.8	2.92	100
9	0.34	0.33	0.49	0.67	0.83	1.03	1.78	10.99	72.75	10.78	100
10	0.47	0.29	0.43	0.65	0.82	0.98	1.2	1.6	9.69	83.86	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A7: Decile mobility matrix, 2016 compared to 2015

	Decile in 2016										
Decile in 2015	1	2	3	4	5	6	7	8	9	10	Total
1	81.29	11.42	3.58	1.77	0.86	0.42	0.27	0.17	0.13	0.1	100
2	9.8	73.02	12.19	2.78	1.1	0.49	0.28	0.16	0.11	0.07	100
3	3.1	10.02	68.9	12.72	2.89	1.11	0.54	0.43	0.19	0.1	100
4	1.73	1.91	10.19	65.09	14.71	3.86	1.3	0.64	0.4	0.18	100
5	1.09	0.92	1.85	12.43	63.44	14.47	3.47	1.31	0.71	0.31	100
6	0.78	0.68	0.99	1.94	12.39	63.97	14.39	2.95	1.4	0.52	100
7	0.58	0.64	0.56	0.93	1.72	11.74	65.68	13.84	3.07	1.23	100
8	0.42	0.51	0.47	0.61	0.88	1.58	10.66	68.82	13.16	2.9	100
9	0.46	0.42	0.62	0.77	0.8	1.01	1.82	9.66	72.41	12.03	100
10	0.77	0.5	0.68	0.98	1.23	1.32	1.56	1.98	8.39	82.59	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A8: Decile mobility matrix, 2017 compared to 2016

	Decile in 2017										
Decile in 2016	1	2	3	4	5	6	7	8	9	10	Total
1	80.5	11.5	3.59	1.92	0.99	0.55	0.35	0.22	0.2	0.18	100
2	10.46	71.3	12.96	3.13	0.98	0.51	0.29	0.17	0.12	0.09	100
3	3.15	11.44	67.1	13.36	2.77	1.16	0.48	0.26	0.17	0.12	100
4	1.81	2.05	11.02	63.14	16.95	2.89	1.14	0.52	0.3	0.19	100
5	1.08	0.97	1.84	12.52	60.23	18.46	2.82	1.18	0.61	0.28	100
6	0.78	0.69	0.92	2.07	12.85	58.34	20.2	2.53	1.16	0.47	100
7	0.6	0.63	0.61	1.08	2.02	13.39	58.79	19.21	2.64	1.03	100
8	0.46	0.53	0.62	0.78	0.99	2.01	12.13	62.45	17.56	2.48	100
9	0.49	0.47	0.65	0.86	0.88	1.27	2.18	11.28	68.36	13.56	100
10	0.81	0.51	0.73	1.11	1.3	1.37	1.56	2.12	8.85	81.65	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A9: Decile mobility matrix, 2018 compared to 2017

	Decile in 2018										
Decile in 2017	1	2	3	4	5	6	7	8	9	10	Total
1	82.71	10.5	3.27	1.7	0.79	0.41	0.24	0.16	0.13	0.11	100
2	8.94	74.06	12.45	2.58	0.99	0.44	0.26	0.13	0.1	0.07	100
3	2.88	9.16	70.5	12.85	2.61	0.99	0.5	0.23	0.16	0.12	100
4	1.65	2.67	8.6	68.33	13.69	2.92	1.14	0.52	0.3	0.19	100
5	0.98	0.95	2.06	9.3	67.25	14.56	2.82	1.22	0.55	0.29	100
6	0.71	0.68	0.9	1.94	10.1	67.15	14.15	2.68	1.2	0.49	100
7	0.49	0.63	0.56	0.97	1.71	9.65	68.58	13.76	2.65	1	100
8	0.37	0.47	0.46	0.62	0.86	1.64	9.19	70.67	13.3	2.42	100
9	0.38	0.4	0.59	0.72	0.79	1.03	1.78	8.88	73.72	11.7	100
10	0.59	0.39	0.57	1	1.25	1.27	1.4	1.81	7.97	83.75	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A10: Decile mobility matrix, 2019 compared to 2018

	Decile in 2019										
Decile in 2018	1	2	3	4	5	6	7	8	9	10	Total
1	83.98	9.78	3.03	1.56	0.74	0.36	0.22	0.13	0.11	0.1	100
2	7.88	75.43	12.01	2.72	1.01	0.44	0.23	0.11	0.09	0.07	100
3	2.73	9.19	71.44	12.38	2.42	0.93	0.44	0.21	0.15	0.1	100
4	1.71	2.06	8.73	69.21	13.69	2.65	1.06	0.47	0.27	0.15	100
5	1.05	0.96	1.71	8.97	67.83	14.92	2.68	1.09	0.53	0.27	100
6	0.74	0.63	0.85	1.91	9.61	67.89	14.37	2.41	1.13	0.46	100
7	0.52	0.59	0.57	0.95	1.76	8.88	69.02	14.27	2.48	0.97	100
8	0.4	0.47	0.49	0.61	0.92	1.63	8.85	71.08	13.25	2.3	100
9	0.4	0.42	0.56	0.72	0.73	0.98	1.75	8.42	74.28	11.74	100
10	0.59	0.42	0.58	0.99	1.29	1.31	1.4	1.82	7.71	83.89	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A11: Decile mobility matrix, 2020 compared to 2019

	Decile in 2020										
Decile in 2019	1	2	3	4	5	6	7	8	9	10	Total
1	82.88	11.18	2.82	1.41	0.73	0.37	0.22	0.15	0.12	0.11	100
2	7.98	73.94	13.79	2.41	0.95	0.41	0.21	0.13	0.1	0.08	100
3	3.12	8.68	69.87	14.28	2.36	0.86	0.38	0.19	0.14	0.11	100
4	2.04	2.4	8.19	68.15	14.88	2.47	0.98	0.43	0.27	0.19	100
5	1.16	1.13	2	8.14	67.4	15.67	2.5	1.08	0.59	0.33	100
6	0.79	0.67	0.97	2.14	8.69	67.56	15.09	2.37	1.16	0.57	100
7	0.54	0.6	0.59	1.03	1.91	8.52	68.79	14.39	2.49	1.13	100
8	0.43	0.51	0.49	0.62	0.95	1.71	8.62	71.06	13.15	2.46	100
9	0.43	0.44	0.61	0.76	0.74	1.02	1.79	8.46	74.37	11.39	100
10	0.58	0.41	0.58	1.09	1.41	1.41	1.43	1.77	7.62	83.7	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A12: Decile mobility matrix, 2021 compared to 2020

	Decile in 2021										
Decile in 2020	1	2	3	4	5	6	7	8	9	10	Total
1	72.52	20.7	3.16	1.6	0.81	0.45	0.26	0.18	0.16	0.16	100
2	12.36	58.42	23.98	2.94	1.09	0.55	0.28	0.14	0.13	0.1	100
3	5.82	11.07	54.92	22.79	2.91	1.23	0.59	0.28	0.21	0.18	100
4	3.52	4.18	10.22	54.54	21.55	3.07	1.3	0.66	0.56	0.41	100
5	1.9	1.9	3.2	10.19	56.06	20.7	2.83	1.42	0.97	0.83	100
6	1.19	1.19	1.48	3.29	10.47	58.62	17.85	2.75	1.88	1.27	100
7	0.76	0.88	0.91	1.59	2.95	9.6	62.35	16.24	2.77	1.96	100
8	0.57	0.65	0.78	0.93	1.47	2.54	10.12	66.04	13.41	3.5	100
9	0.58	0.52	0.73	1	1.06	1.43	2.54	9.93	70.15	12.05	100
10	0.98	0.56	0.64	1.12	1.59	1.75	1.82	2.29	9.7	79.54	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A13: Mobility in top income groups, 2012 compared to 2011

	Income group in 2012				
Income group in 2011	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.43	1.54	0.02	0	100
p90 - p99	15.03	82.48	2.45	0.05	100
p99 - p99.9	5.13	21.14	70.33	3.39	100
p99.9 - p100	3.94	5.12	27.33	63.61	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A14: Mobility in top income groups, 2013 compared to 2012

	Income group in 2013				
Income group in 2012	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.43	1.54	0.03	0	100
p90 - p99	15.08	82.49	2.39	0.04	100
p99 - p99.9	4.76	21.21	70.73	3.3	100
p99.9 - p100	3.25	5.16	26.38	65.21	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A15: Mobility in top income groups, 2014 compared to 2013

	Income group in 2014				
Income group in 2013	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.4	1.56	0.04	0	100
p90 - p99	15.4	82.21	2.34	0.05	100
p99 - p99.9	5.18	21.53	70.17	3.11	100
p99.9 - p100	2.76	4.99	26.1	66.15	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A16: Mobility in top income groups, 2015 compared to 2014

	Income group in 2015				
Income group in 2014	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.2	1.72	0.07	0	100
p90 - p99	17.18	80.34	2.4	0.09	100
p99 - p99.9	7.17	24.15	65.55	3.13	100
p99.9 - p100	3.41	5.73	30.32	60.53	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A17: Mobility in top income groups, 2016 compared to 2015

	Income group in 2016				
Income group in 2015	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.06	1.87	0.06	0	100
p90 - p99	18.08	79.08	2.76	0.07	100
p99 - p99.9	11.92	21.21	63.33	3.54	100
p99.9 - p100	6.11	8.44	25.74	59.71	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A18: Mobility in top income groups, 2017 compared to 2016

	Income group in 2017				
Income group in 2016	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	97.95	1.99	0.05	0	100
p90 - p99	19.17	78.01	2.74	0.08	100
p99 - p99.9	11.58	20.8	64.23	3.38	100
p99.9 - p100	4.87	6.34	28.66	60.13	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A19: Mobility in top income groups, 2018 compared to 2017

	Income group in 2018				
Income group in 2017	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.18	1.76	0.06	0	100
p90 - p99	17.01	80.3	2.61	0.08	100
p99 - p99.9	9.76	21.42	65.44	3.38	100
p99.9 - p100	5.45	7.79	27.29	59.47	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A20: Mobility in top income groups, 2019 compared to 2018

	Income group in 2019				
Income groups in 2018	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.2	1.74	0.05	0	100
p90 - p99	16.71	80.58	2.65	0.06	100
p99 - p99.9	11.06	19.75	65.81	3.37	100
p99.9 - p100	7.42	7	24.8	60.79	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A21: Mobility in top income groups, 2020 compared to 2019

	Income group in 2020				
Income group in 2019	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	98.18	1.75	0.07	0	100
p90 - p99	16.95	80.45	2.51	0.08	100
p99 - p99.9	10.81	20.46	65.62	3.11	100
p99.9 - p100	6.68	6.57	25.96	60.79	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Table A22: Mobility in top income groups, 2021 compared to 2020

	Income group in 2021				
Income group in 2020	p0 - p90	p90 - p99	p99 - p99.9	p99.9 - p100	Total
p0 - p90	97.72	2.12	0.15	0.01	100
p90 - p99	21.17	76.17	2.53	0.14	100
p99 - p99.9	14.55	25.88	56.43	3.14	100
p99.9 - p100	9.97	7.64	28.65	53.74	100

Note: balanced panel.

Source: authors' calculations based on the 2023 version of the individual panel (National Treasury and UNU-WIDER 2023).

Appendix B: Data

This section in the appendix reports the dataset as well as the variables used to generate the results in this paper. The data are restricted use access at the National Treasury Secure Data Facility in Pretoria.

Data access

The data was accessed under a non-disclosure agreement. The output was checked so that the anonymity of no individual or firm was compromised. The results do not represent any official statistics (National Treasury or SARS). The views expressed in our research do not necessarily represent the views of the National Treasury or SARS.

Data used come from the intermediate 2023 version of the Individual Panel. More specifically, we used the Income Panel and Source of Income Panel for our analyses. Date of first access: 6 May 2022. Last accessed: 8 December 2023.

Software

Our analysis was conducted using Stata 17. To calculate inequality measures we made use of the ineqdeco module (Jenkins 2021). Income shares were calculated using pshare (Jann 2016). Inequality decompositions were done using the DASP package (Araar and Duclos 2022) and the sgini command (van Kerm 2020). Lastly, the growth incidence curve was also generated with the DASP package (v3).

Variables

From the Income Panel: id_d, tax_year, tax_liability_d, age, gender

From the Source of Income Panel: id_d, tax_year, source_code, amount, final_d, irp5_id

We followed the approach described in Ebrahim and Axelson (2019) to create a gross income variable (see the appendix in their paper). We started by splitting the source codes ('source_code') into one of nine income categories and then took the sum of all the amounts ('amount' variable) for each category per individual per year. We calculated CPI deflators based on the numbers provided by Stats SA to create real gross income and real tax liability. We define disposable income as real gross income minus real tax liability.

We also created variables for each income category to create the source of income graphs and decompose inequality by income source. These income variables were also adjusted for price level changes.

Cleaning and notes

To merge the Income and Source of Income panels, we started by cleaning the income panel. We noticed what appeared to be duplicates: some individuals had two entries per tax year, where one entry had missing gender and the other had gender being either male or female. We dropped the observations with missing gender. We then merged the Source of Income Panel and Income Panel by id_d and tax_year (m:1). We dropped observations that could not be matched. This created a very large dataset which we later collapsed.

We dropped observations for the 2022 tax year and individuals with age < 16. We imputed gender for cases where an individual had 'male' or 'female' for at least one of the tax years and missing in others. Some individuals have what appear to be duplicate IRP5 certificates, or IRP5s with the same certificate number. Rather than drop all these individuals, we took the average of all the amounts for each source code on the certificates.

We dropped what appeared to be errors in the data (extremely large isolated values) and decided to top-code the income in the 3601 code at the 99.5th percentile. We decided on this cut-off as it allowed us to exclude what looked very much like errors but keep what looked like very large values that may or may not be errors. Other researchers are welcome to experiment with different cut-offs, as excluding/including very large values will affect estimates of income shares and inequality.

Our first inequality estimates showed an increase in inequality for 2018, and an unbelievable jump in the income shares of men between the ages of 16 and 30. By browsing the data, we noticed that some individuals had amounts of over ZAR1 billion for the 3809 and 3820 source codes, but zero tax liability. We believe these amounts are errors, and the individuals with these amounts are not included in our analysis.

After cleaning we created a collapsed dataset that had one observation per individual per tax year. This is the dataset we used to generate the results presented in this paper.

The Individual Panel does not include a variable to flag outliers. Consequently, we followed the following steps to identify potential errors in the data:

1. We listed the ten highest and lowest amounts for the source codes that contribute the most to total income each year
2. We listed the ten highest and lowest amounts for each of the source codes added after 2016
3. We checked the value of the tax liability for each individual that seemed to have implausibly large amounts for the source codes in steps 1 and 2 and dropped observations where the tax liability seemed too low.

Table B1: Source codes included in gross income

Main category	Subcategory	Source code	Description	Category
Labour	Business income	102 to 3499	Profit/Loss	Business income
Labour	Earnings	3601	Income - PAYE	Normal income
Labour	Earnings	3605	Annual Payment - PAYE	Normal income
Labour	Earnings	3606	Commission - PAYE	Normal income
Labour	Earnings	3607	Overtime - PAYE	Normal income
Labour	Earnings	3701	Travel allowance - PAYE	Allowance
Labour	Earnings	3702	Reimbursed travel allowance - PAYE	Allowance
Labour	Earnings	3703	Reimbursed travel allowance - Excl	Allowance
Labour	Earnings	3704	Subsistence allowance local travel - IT	Allowance
Labour	Earnings	3707	Share option exercised - PAYE	Allowance
Labour	Earnings	3708	Public office allowance - PAYE	Allowance
Labour	Earnings	3709	Uniform allowance	Allowance
Labour	Earnings	3710	Tool allowance	Allowance
Labour	Earnings	3711	Computer allowance	Allowance
Labour	Earnings	3712	Phone allowance	Allowance
Labour	Earnings	3713	Other allowances - PAYE	Allowance
Labour	Earnings	3714	Other allowances - Excl	Allowance
Labour	Earnings	3715	Subsistence allowance foreign travel - IT	Allowance
Labour	Earnings	3717	Broad-based employee share plan - PAYE	Allowance
Labour	Earnings	3718	Vesting of equity instruments - PAYE	Allowance
Labour	Earnings	3751	Travel allowance - foreign service income	Allowance
Labour	Earnings	3752	Reimbursed travel allowance - foreign service income	Allowance
Labour	Earnings	3753	Foreign reimbursive travel allowance	Allowance
Labour	Earnings	3757	Share option exercised - foreign service income	Allowance
Labour	Earnings	3763	Other allowances - foreign service income	Allowance
Labour	Earnings	3764	Other non-taxable allowances - foreign service income	Allowance
Labour	Earnings	3765	BBE share plan - foreign service income	Allowance
Labour	Earnings	3768	Vesting of equity instruments - foreign service income	Allowance
Labour	Earnings	3801	General fringe benefits - PAYE	Fringe benefit
Labour	Earnings	3802	Use of motor vehicle acquisition by employer, not lease - PAYE	Fringe benefit
Labour	Earnings	3803	Use of asset - PAYE	Fringe benefit

Labour	Earnings	3804	Meals, etc. - PAYE	Fringe benefit
Labour	Earnings	3805	Accommodation - PAYE	Fringe benefit
Labour	Earnings	3806	Free or cheap services	Fringe benefit
Labour	Earnings	3808	Employee’s debt - PAYE	Fringe benefit
Labour	Earnings	3809	Taxable bursaries or scholarships - PAYE	Fringe benefit
Labour	Earnings	3810	Medical aid contributions - PAYE	Fringe benefit
Labour	Earnings	3813	Medical service costs - PAYE	Fringe benefit
Labour	Earnings	3815	Non-taxable bursaries or scholarships - Exclu	Fringe benefit
Labour	Earnings	3816	Use of motor vehicle acquisition by employer by lease - PAYE	Fringe benefit
Labour	Earnings	3820	Taxable bursaries or scholarships (FE) - PAYE	Fringe benefit
Labour	Earnings	3821	Non-taxable bursaries or scholarships (FE) - Excl	Fringe benefit
Labour	Earnings	3852	Use of motor vehicle acquisition by employer, not lease - foreign income	Fringe benefit
Labour	Earnings	3855	Foreign accommodation	Fringe benefit
Labour	Earnings	3856	Foreign fee or cheap services	Fringe benefit
Labour	Earnings	3858	Foreign employee’s debt	Fringe benefit
Labour	Earnings	3851	General fringe benefits (Subject to PAYE) - Foreign Service Income	Fringe benefit
Labour	Earnings	3860	Medical aid contributions - foreign service income	Fringe benefit
Labour	Earnings	3863	Medical service costs (PAYE) - foreign service income	Fringe benefit
Labour	Earnings	3865	Non-taxable bursaries or scholarships - foreign services	Fringe benefit
Labour	Earnings	3817	Benefit employer pension fund contributions (subject to PAYE)	Fringe benefit
Labour	Earnings	3825	Benefit employer provident fund contributions (subject to PAYE)	Fringe benefit
Labour	Earnings	3828	Employer Retirement annuity Fund contributions	Fringe benefit
Labour	Earnings	4236	Remuneration from foreign employer for services rendered in South Africa	Investment income
Labour	Other labour income	3615	Director’s income	Normal income
Labour	Other labour income	3616	Independent contractors’ income	Normal income
Labour	Pension income	3603	Pension (subject to PAYE)	Normal income
Labour	Pension income	3604	Pension (non-taxable)	Normal income
Labour	Pension income	3610	Annuity from a RAF (subject to PAYE)	Normal income
Labour	Pension income	3614	Other retirement lump sums (subject to PAYE)	Normal income
Labour	Pension income	3653	Foreign pension (subject to paye)	Normal income
Labour	Pension income	3660	Foreign annuity from a raf (subject to paye)	Normal income
Labour	Pension income	3664	Foreign other retirement lump sums (subject to paye)	Normal income
Labour	Pension income	3902	Pension or RAF in respect of withdrawal (subject to PAYE)	Lump sum retirement
Labour	Pension income	3903	Pension or RAF in respect of retirement (subject to PAYE)	Lump sum retirement

Labour	Pension income	3904	Provident in respect of withdrawal (subject to PAYE)	Lump sum retirement
Labour	Pension income	3905	Provident in respect of retirement (subject to PAYE)	Lump sum retirement
Labour	Pension income	3908	Surplus apportionments and exempt policy proceeds (non-taxable)	Lump sum retirement/income
Labour	Pension income	3909	Unclaimed benefits	Lump sum retirement/income
Labour	Pension income	3915	Retirement or termination of employment lump sum benefits or commutation of annuities	Lump sum retirement
Labour	Pension income	3920	Lump sum withdrawal benefits (subject to PAYE)	Lump sum retirement
Labour	Pension income	3921	Living annuity and section 15C of the pension funds act, surplus apportionments (subject to PAYE)	Lump sum retirement
Labour	Pension income	3923	Transfer of unclaimed benefits	Lump sum retirement
Capital	Interest income	4201	Local interest	Investment income
Capital	Interest income	4218	Foreign interest	Investment income
Capital	Rental income	2532	Business income component: property letting income, residential accommodation	Business income
Capital	Rental income	2533	Business income component: property letting loss, residential accommodation	Business income
Capital	Rental income	4210	Local rental from letting of fixed property	Investment income
Capital	Rental income	4211	Local rental loss from letting of fixed property	Investment income
Capital	Rental income	4289	Foreign rental loss	Not categorised
Capital	Rental income	4288	Foreign rental gain	Not categorised
Capital	Dividends	3719	Dividends not exempt ito para (dd) of the proviso to s10(1)(k)(i) (PAYE)	Allowance
Capital	Dividends	3720	Dividends not exempt ito para (ii) of the proviso to s10(1)(k)(i) (PAYE)	Allowance
Capital	Dividends	3721	Dividends not exempt ito para (jj) of the proviso to s10(1)(k)(i) (PAYE)	Allowance
Capital	Dividends	3723	Dividends not exempt ito para (kk) of the proviso to s10(1)(k)(i) (PAYE)	Allowance
Capital	Dividends	3767	Foreign broad-based employee share plan (subject to PAYE)	Allowance
Capital	Dividends	3770	Foreign dividends not exempt ito para (ii) of the proviso to s10(1)(k)(i) (PAYE)	Not categorised
Capital	Dividends	4216	Foreign dividends	Investment income
Capital	Dividends	4230	Controlled foreign company share of profit	Investment income
Capital	Dividends	4238	Taxable local dividends, i.e. REIT	Investment income
Capital	Dividends	4292	Dividends deemed to be income in terms of s8E and s8EA	Investment income
Capital	Capital gains	4250	Local capital gain	Investment income
Capital	Capital gains	4251	Loss local capital	Investment income
Capital	Capital gains	4252	Foreign capital gain	Investment income
Capital	Capital gains	4253	Loss - foreign capital	Investment income
Capital	Other investment income	4212	Royalties	Investment income

Capital	Other investment income	4213	Loss - royalties	Investment income
Capital	Other investment income	4214	Other receipts and accruals	Investment income
Capital	Other investment income	4222	Foreign business or trading	Investment income
Capital	Other investment income	4223	Loss - foreign business or trading	Investment income
Capital	Other investment income	4228	Other foreign income	Investment income
Capital	Other investment income	4229	Loss - other foreign income	Investment income
Capital	Other investment income	4235	Income reflected on a South African IRP5 or IT3a that was subject to tax outside SA	Investment income
Capital	Other investment income	4249	Foreign tax credits refunded or discharged in terms of S6quat(1C)	Investment income
Capital	Other investment income	4291	Foreign income in terms of s6quat(1C)	Investment income
Other	Activity income	4278	Foreign royalties	Activity income
Other	Activity income	4279	Loss - foreign royalties	Activity income
Other	Activity income	4280	Sporting	Activity income
Other	Activity income	4281	Loss - sporting	Activity income
Other	Activity income	4282	Collectables	Activity income
Other	Activity income	4283	Loss - collectables	Activity income
Other	Activity income	4284	Animal showing	Activity income
Other	Activity income	4285	Loss - animal showing	Activity income
Other	Activity income	4286	Gambling	Activity income
Other	Activity income	4287	Loss - gambling	Activity income

Source: authors' compilation based on study data.

Appendix C: Estimates from NIDS

The definitions of informal and formal sector workers are based on those suggested by Cichello and Rogan (2017).

Informal sector workers can be defined as those who are:

- self-employed in enterprises not registered for income tax or VAT.
- employed with regular employment who do not receive both pensions and medical aid from their employers, do not pay UIF, and do not have written employment contracts.
- employed with work that is irregular and short-term or work done in addition to their first two jobs/self-employment businesses.

Formal sector workers are those who are:

- self-employed in VAT or tax-registered enterprises.
- employed with regular employment, who have written contracts or pay UIF or receive employer-based pensions and medical aid.

It is not possible to distinguish between those employed in the informal sector and those with informal employment (those employed in the formal sector but without contracts and/or benefits). This means some formal sector workers may still be captured in the 'informal sector' category.

Based on the above definitions, almost 23 per cent of individuals in South Africa were employed in the informal sector in 2017. This proportion is around 25 per cent using NIDS Wave 2 and Wave 4 data.

Most of the income questions in NIDS refer to income received in the last 30 days. We create four income variables in our analysis: monthly and annual versions of gross income and total income. The gross income variable includes the same categories included by Hundenborn et al. (2019), but we exclude the *em1proflm_a* variable and include the *spen* variable in a different income category. The total income category includes all sources from gross income, plus income from grants, other income from the government, remittances, and other income. The income variables are available in the *indderived* and *adult* NIDS data files. Table C2 below summarizes the categories and descriptions of the variables we use. The rules used to create a crude estimate of annual income values are also explained in the table. Table C1 reports the filing thresholds for different age groups across tax years.

Table C1: Personal income tax thresholds for different age groups, 2011–23

Tax year	<65	>65	>75
2011	57,000	88,528	
2012	59,750	93,150	104,261
2013	63,556	99,056	110,889
2014	67,111	104,611	117,111
2015	70,700	110,200	123,350
2016	73,650	114,800	128,500
2017	75,000	116,150	129,850
2018	75,750	117,300	131,150
2019	78,150	121,000	135,300
2020	79,000	122,300	136,750
2021	83,100	128,650	143,850
2022	87,300	135,150	151,100
2023	91,250	141,250	157,900

Source: authors' elaboration based on South African Revenue Service (SARS) tax rates (archive). See https://www.sars.gov.za/tax-rates/archive-tax-rates/.

Table C2: Income variables in NIDS Waves 2, 4, and 5

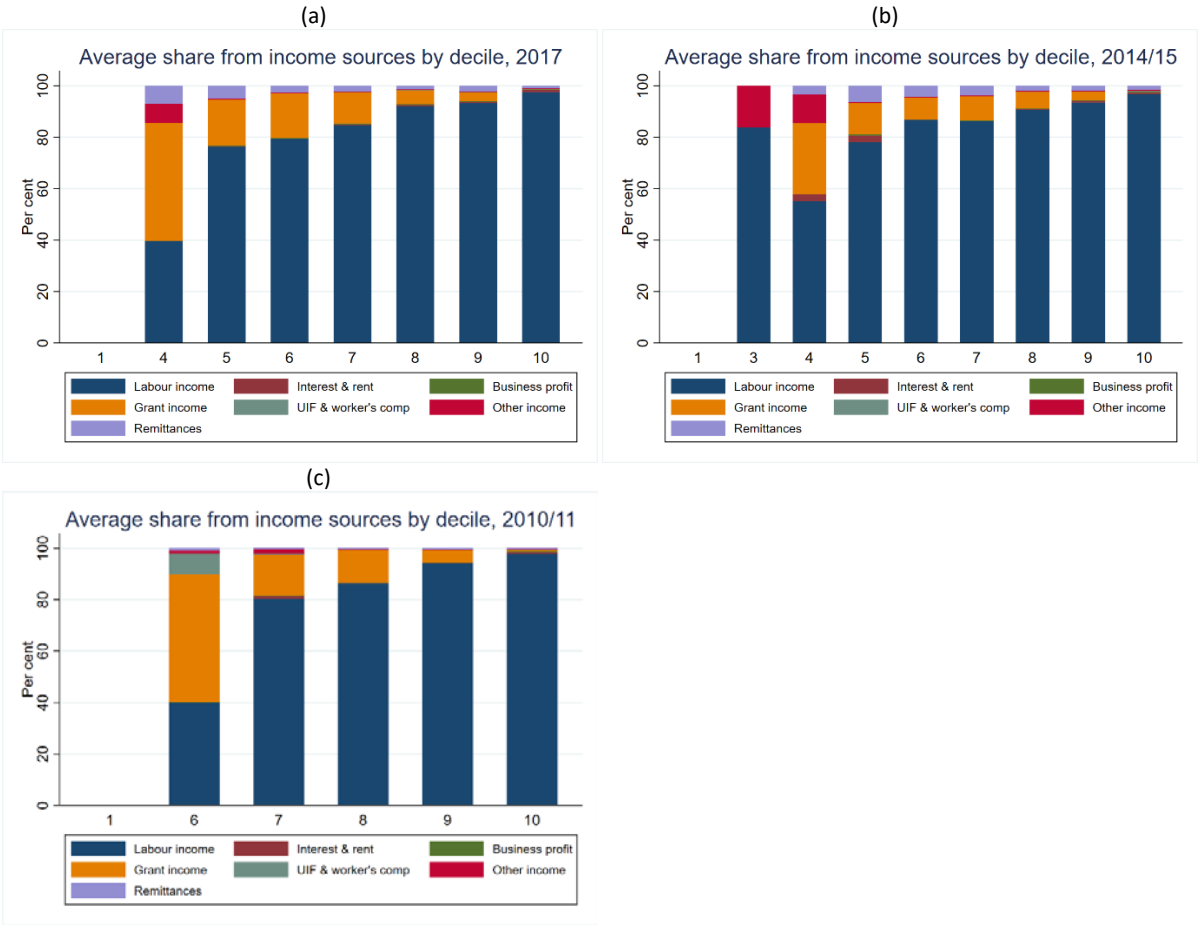
	Category	Variable	Description	Rule for annual values
	Investment income	indi	Income from interest and dividends	Multiply by 12
		rnt	Income from rentals	Multiply by 12
+	Business income	emsprof_a	Profit from own business	No adjustment
		emsloss_a	Loss from own business	No adjustment
+	Labour income	em1inc	Gross income from primary occupation	Multiply by 12
		em2inc	Gross income from secondary occupation	Multiply by 12
		swag	Income from self-employment	Multiply by 12
		prof	Income from profit share	*em1prf_a* variable (share of profit in last 12 months)*
		ppen	Income from private/foreign pension and retirement annuities	Multiply by 12
		extr	Income from extra payment on a piece-rate basis	No adjustment
		cheq	Income from 13th cheque	*em1cheq_a* variable (13th cheque in last 12 months)*
		bonu	Income from other bonuses from main job	*em1bon_a* variable (bonus in the last 12 months)*
=	*Gross income*			
+	Grant income	spen	Income from state old age pension	Multiply by 12
		chld	Income from child support grant	Multiply by 12
		dis	Income from disability grant payments	Multiply by 12
		fost	Income from foster care grant	Multiply by 12
		cdep	Income from care dependency grant	Multiply by 12
+	Other income from government	uif	Income from UIF payments	No adjustment
		comp	Income from workmen's compensation	Multiply by 3
+	Remittances	remt	Income from remittance payments	Multiply by 12
+	Other income	help	Income from helping friends	No adjustment
		emcinc	Income from paid casual employment	No adjustment
		retr	Income from retrenchment payments	No adjustment
		inhe	Income from inheritances	No adjustment

	brid	Income from lobola/bridal payments	No adjustment
	gift	Income from gifts	No adjustment
	loan	Income from loan repayments	No adjustment
	sale	Income from sale of household assets	No adjustment
	othe	Income from other sources	No adjustment
= *Total income*			

Note: em1cheq_a, em1prf_a, em1bon_a, and emsinc are not available in Wave 2.

Source: authors' compilation based on study data.

Figure C1: Average share from income sources – formally employed sample



Note: based on total personal (annual) income. Data are weighted.

Source: authors' illustration based on data from NIDS Wave 2 (2010/11), Wave 4 (2014/15), and Wave 5 (2017).