An assessment of inequality estimates for the case of South Africa

Muna Shifa, Rejoice Mabhena, Vimal Ranchhod, and Murray Leibbrandt*

July 2023

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: A substantial amount of research has been conducted examining inequality in South Africa using multiple data sources. We provide an overview of this research in this paper. Furthermore, we use nationally representative survey data to estimate income and consumption inequality in South Africa between 1993 and 2017, and we discuss the comparability of income and consumption measures across surveys. We compare our estimates of inequality to those in the World Income Inequality Database (WIID), which provides adjusted inequality estimates to ensure consistency across time and countries. Despite these adjustments, there are some large fluctuations in income inequality estimates from the WIID, which can be partly attributed to data comparability problems. The overall trend in post-apartheid South Africa indicates an increase in income and wealth inequality. However, estimates of inequality levels over time are influenced in part by methodological changes implemented across multiple surveys over time, making definitive conclusions about inequality levels and trends difficult.

Key words: income distribution, inequality, surveys, South Africa

JEL classification: D21, C83

---

* ACEIR, University of Cape Town, Cape Town, South Africa; corresponding author: Muna Shifa, muna.shifa@uct.ac.za

1   Introduction

Analysing inequality trends over time using survey data is a difficult task, since micro-level survey datasets that have detailed income and consumption modules are not readily available over long time frames, particularly in developing countries. Additionally, most existing surveys are not comparable due to changes in data collection methods. Over the years, there have been substantial improvements in the methodology used to capture income and expenditure in household surveys in South Africa. While these changes reflect the country's evolving statistical advances, they impact on the comparability of income and consumption estimates over time.

Several efforts have been made to create comparable inequality estimates that may be used to examine patterns in inequality over time and across nations. The World Income Inequality Database (WIID) is the most comprehensive compilation of income inequality statistics in the world. In recent years, UNU-WIDER has produced a WIID Companion (UNU-WIDER 2021), which provides inequality estimates for each country as well as estimates of global inequality levels over time, in addition to the WIID dataset. The WIID Companion provides adjusted inequality estimates to make the measure of inequality comparable across time and between countries (Gradín 2021b). In the context of South Africa, the WIID Companion series is useful for generating estimates of inequality dating back to the early 1960s, when representative household surveys were unavailable. Since additional survey data became accessible in the early 1990s, extensive research has been conducted which analyses the patterns and dynamics of inequality in post-apartheid South Africa.

In this paper, we provide an overview of the existing narrative on South African income inequality trends. We also provide income inequality estimates for South Africa between 1993 and 2017 using household survey data. Then, we compare our inequality estimates to those in the WIID Companion for South Africa, evaluating how well the WIID Companion's adjusted inequality data convey the same narrative as the original reported inequality estimates and the reality on the ground. We recognize the trade-off between having perfectly comparable data and being able to say something about inequality trends over time and compare across countries. Therefore, the main goal of our study is to provide context and highlight key methodological issues that may affect the comparability of inequality estimates in South Africa. Although some estimates of inequality have been available since the 1960s, our estimates of inequality and discussion of survey methodology are based primarily on surveys conducted after 1993, when nationally representative household surveys became more available.

2   An overview of the levels and trends of income and wealth inequality in South Africa

There is a substantial body of research that examines inequality in South Africa. In this section, we provide a brief overview of the main dimensions and narratives of inequality in the country. Previous work provides more detailed discussions (e.g., Bhorat et al. 2020a; Leibbrandt et al. 2009; Leibbrandt et al. 2018; Stats SA 2019).

2.1   Contemporary income and wealth inequality

There is consensus that contemporary South Africa has a very high level of income and wealth inequality. According to a recent World Bank assessment, South Africa is the most unequal country in the world, ranking first out of 164 countries; this based on the Gini coefficient of consumption

(or income) per capita (World Bank 2022: 12). This holds true if we look at income and wealth share data from the World Inequality Database (WID). To put this in perspective, the top 10 per cent income share in 2021 for South Africa was 65 per cent, while the estimate ranged from 31 per cent in France to 46 per cent in the United States (Figure 1). At the time, the bottom 50 per cent income share was only 6 per cent in South Africa, while it was estimated to be 23 per cent in France and 14 per cent in the United States and China. Similarly, in South Africa, the top 10 per cent received 86 per cent of net personal wealth, while the bottom 50 per cent had negative average wealth.

Figure 1: Top 10 per cent and bottom 50 per cent income and wealth shares

Source: authors' illustration based on data from WID (2022).

2.2   Trends in income and wealth inequality

Prior to the early 1990s, nationwide analysis of the levels and trends of income and wealth inequality in South Africa was limited due to a lack of appropriate and representative survey data. Appendix A lists the data sources used for analysing inequality and poverty in South Africa. Most research that examined income inequality used population censuses and specialized surveys. (Ardington et al. 2006; Lachman and Bercuson 1992; Leibbrandt et al. 2006; McGrath 1984; Simkins 2004; Treiman et al. 1996; Whiteford and Seventer 2000). Others utilized income tax return data to study the long-run distribution of top incomes (see Alvaredo and Atkinson 2022). According to the evidence from these studies, income inequality in South Africa was high by international standards before and during the early 1990s, with the Gini coefficient ranging between 0.68 and 0.73. For example, Whiteford and van Seventer (2000) use census data to show that the national Gini coefficient for income remained stable between 1975 and 1991 (i.e., 0.68) then increased slightly to 0.69 in 1996. According to the WIID Companion series (see Figure 10), the Gini coefficient for income inequality ranged between 0.72 and 0.76 in the 1960s and early 1970s, then fell to 0.68 in 1975 and continued to fall until the late 1980s. Thus, the overall evidence

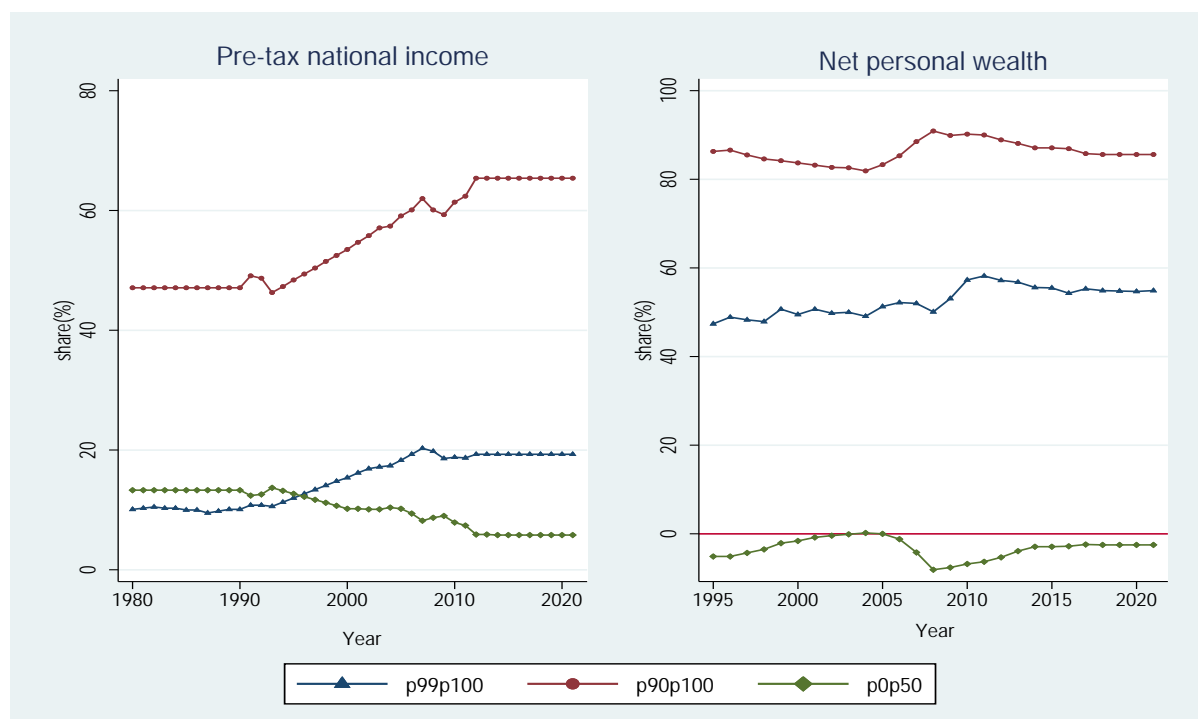suggests that income inequality remained stable or even declined between the 1960s and the late 1980s.

With the increased availability of representative household surveys after 1994, income inequality analyses proliferated (see Hundenborn et al. 2018; Leibbrandt et al. 2006; Leibbrandt et al. 2012; Seekings et al. 2004; Simkins 2005; Van der Berg et al. 2011). According to the evidence from this literature, income inequality increased over the 1990s and early 2000s. Seekings et al. (2004), for example, use the Income and Expenditure Survey (IES) and show an increase in the Gini coefficient for income inequality from 0.65 to 0.70 between 1995 and 2000. According to Leibbrandt et al. (2006), the income Gini coefficient increased from 0.68 in 1996 to 0.73 in 2001, based on the 1996 and 2001 censuses. Using the National Income Dynamics Study (NIDS) dataset, Leibbrandt et al. (2012) also show that the income Gini coefficient increased from 0.66 in 1993 to 0.70 in 2008. Overall, while there is agreement on the rising trend of income inequality during the 1990s and early 2000s, there are differences in the level of the Gini estimates across the different studies, which may be attributable to data comparability issues (see Appendix A).

More recent studies using survey data suggest that income inequality has remained stable or even declined since 2008 (Checchi et al. 2018; Hundenborn et al. 2018; Stats SA 2019). For example, Hundenborn et al. (2018) report that the income Gini coefficient rose from 0.68 in 1993 to 0.69 in 2008 before decreasing to 0.66 in 2014.

In this paper, we estimate income and consumption inequality using household surveys. According to our estimates of income inequality based on various household surveys from 1993 onwards, income inequality increased in the early 1990s and 2000s then remained stable or even fell in recent years (more on this later). However, evidence from research combining tax and survey data suggests that income and wealth inequality in South Africa continued to rise after 2008 because of dynamics at the top end of the income and wealth distribution (Alvaredo and Atkinson 2022; Assouad et al. 2018; Bassier and Woolard 2021; Chatterjee et al. 2022). Given that household surveys under-report the income of high-income earners, recent research has increasingly used a combination of tax data and survey data to analyse top income shares.

According to Bassier and Woolard (2021), the share of top 1 per cent real income in South Africa nearly doubled between 2003 and 2016. Similarly, data from WID show that the share of income going to the top 10 per cent increased from 47.31 per cent in 1994 to 60.1 per cent in 2008 and 65.4 per cent in 2017, while the share going to the bottom 50 per cent fell from 13.2 per cent in 1994 to 8.7 per cent in 2008 and 5.8 per cent in 2017 (Figure 2). Similarly, the top 10 per cent's share of wealth increased from 86 per cent in 1995 to 91 per cent in 2008 before declining slightly, while the share of the top 1 per cent's wealth increased from 47.4 per cent in 1995 to 50.1 per cent in 2008 and 55.5 per cent in 2017. On the other hand, the poorest 50 per cent had a negative average net worth during the same time period. Thus, the evolution of income and wealth shares suggests rising income and wealth inequality in post-apartheid South Africa.

Figure 2: Evolution of income and net personal wealth

Income share estimates prior to 1994 demonstrate that income shares remained stable with few fluctuations, implying that income inequality remained stable between 1980 and 1990 (Figure 2). Overall, what can be concluded from the existing body of inequality research is that income inequality, as measured by the Gini coefficient and income shares, increased in the early 1990s and 2000s then remained stable or perhaps increased thereafter.

2.3   Non-income dimensions of inequality

In contrast to income and wealth disparities, inequalities in access to basic services have been declining since apartheid's demise (Bhorat and Van der Westhuizen 2013; Wittenberg and Leibbrandt 2017; Shifa and Leibbrandt 2022). For instance, Wittenberg and Leibbrandt (2017) find that the Gini index for asset inequality fell from 0.47 to 0.29 between 1993 and 2008.[1] Similarly, Shifa and Leibbrandt (2022) show that inequalities in education levels and access to basic services significantly declined between 1998 and 2016. The decline in asset and basic services inequality is consistent with the large expansion in the provision of access to basic services and increases in ownership of household durables over time (Figure 3).

---

[1] Wittenberg and Leibbrandt (2017) used asset indices constructed using access to basic services such as water and electricity and ownership of household assets such as TV, radio, fridge, car, and livestock.

Figure 3: Distribution of access to basic services
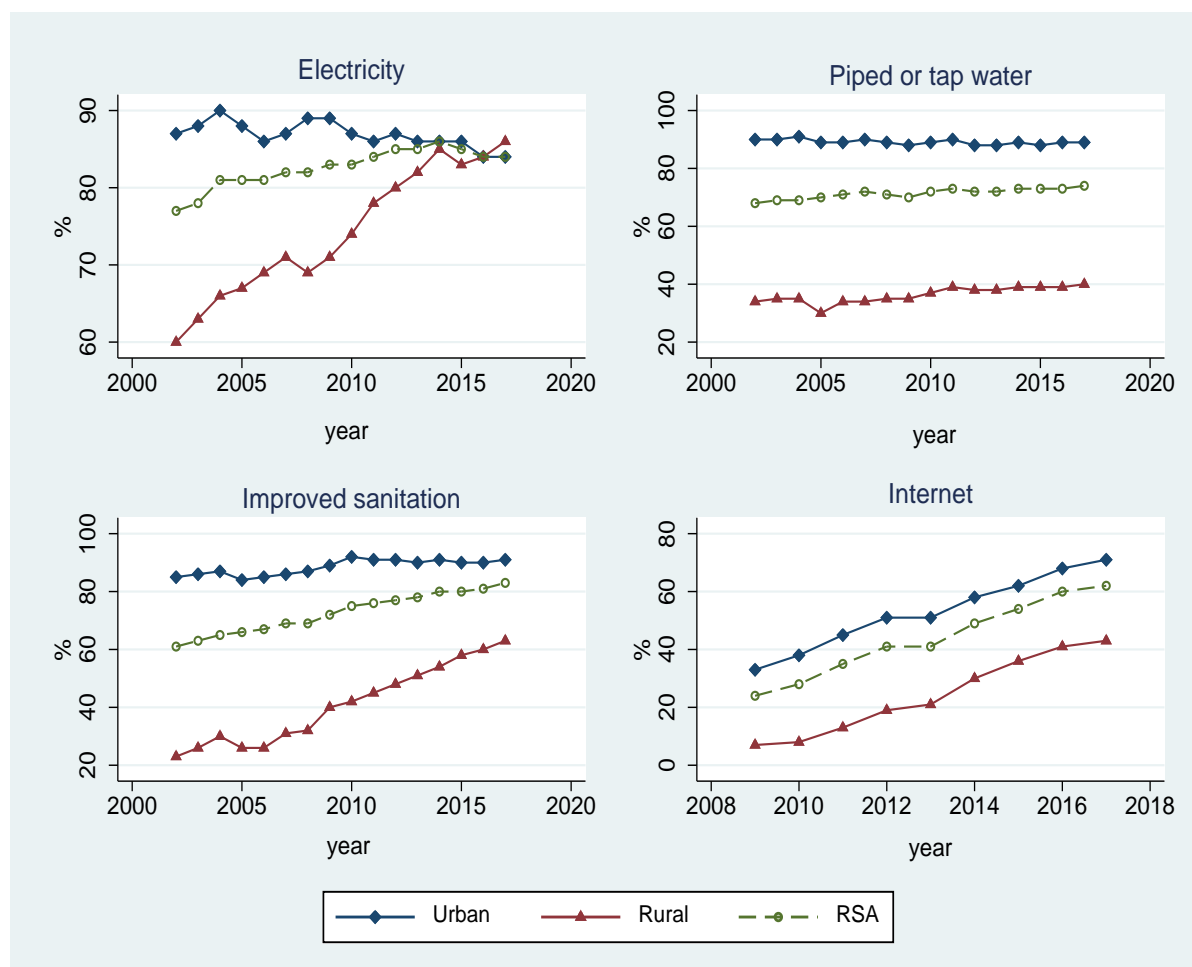
3 Key drivers of income inequality

3.1 Structural factors

Understanding the fundamental causes of inequality in South Africa necessitates a look at the country's history (see Bhorat et al. 2020a; Feinstein 2005; Wilson 2011). Since the 1960s, the appropriation of land, water, and minerals by a few led to the accumulation of wealth and power by a few individuals and groups, the majority of whom were white (Wilson 2011). This was worsened by discriminatory labour and education policies that prevented the majority of non-whites from obtaining quality education and skills. This institutionalized discrimination resulted in a high level of racial income and wealth inequality. Although long-run examination of income and wealth inequalities in South Africa is sparse, recent evidence suggests that in the early 1900s, racial income inequality in South Africa was very high compared with other similar settler colonies (Alvaredo and Atkinson 2022). According to Alvaredo and Atkinson (2022), in 1956, 98.4 per cent of the top 5 per cent income earners were white, while Africans accounted for only 0.05 per cent. Over the next three decades, the estimate for whites fell to 90.6 per cent in 1987, while it increased to 2.5 per cent for Africans.

Beginning in the 1960s, changes in the political, social, and economic landscapes were related to shifting income inequality dynamics in South Africa. The South African economy enjoyed high

and sustained economic growth and manufacturing-led structural transformation up until the early 1970s, with both manufacturing value added and employment shares increasing (see Bhorat et al. 2020a; Lachman and Bercuson 1992). While such development first benefited highly skilled white workers, with African/white income disparities increasing in the 1960s, the continued demand for labour increased the employment of Africans, reducing the racial income gaps in the early 1970s (see Bhorat et al. 2020a; Lachman and Bercuson 1992; Leibbrandt et al. 2009). However, the decline in the GDP growth rate and increased political instability which started in the mid-1970s were accompanied by a rise in unemployment and de-industrialization. According to Lachman and Bercuson (1992: 3), formal sector unemployment increased from 25 per cent in 1974 to over 42 per cent by 1989. Such developments were associated with a slowing of progress in reducing racial income inequality that started in the 1980s.

3.2 Racial inequality

Since the 1970s, the role of racial disparities in explaining income inequalities has been diminishing (Leibbrandt et al. 2006). For example, the contribution of inter-racial inequality in explaining overall income inequality decreased from 55 per cent in 1993 to 28.3 per cent in 2017 (see Figure 4). This is primarily due to an increase in within-race income disparities, particularly among Africans. In 1993, the average per capita income for whites was ten times that of Africans, but by 2017 this had dropped to six times. According to Chatterjee et al. (2021: 34), the average factor income for whites was seven times that of Africans in the early 1990s but since fell to four in 2015.[2] However, when all Africans in the top 10 per cent of the income distribution are excluded from the analysis, the white–African income ratio appears to have risen significantly, from 10.5 in the early 1990s to 11.5 in 2019. This indicates that racial inequalities in South Africa have decreased, but that this decrease is primarily due to the emergence of a new African elite, who have taken an increasing share of the top 10 per cent of the income distribution.

Figure 4: Contribution of within-race and between-race inequalities
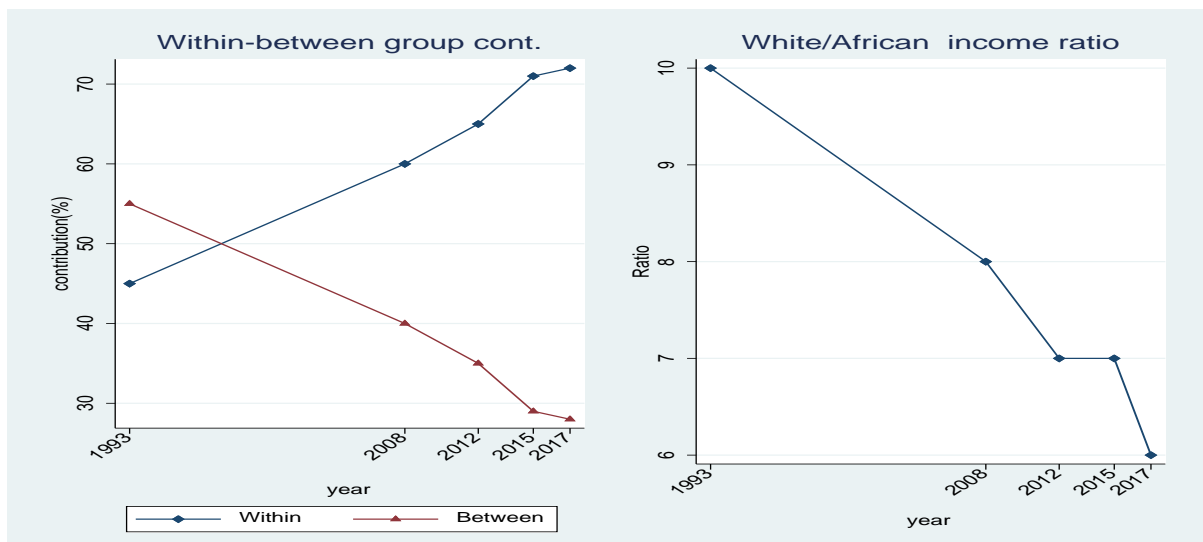
Source: authors' illustration based on own estimates from NIDS and PSLSD.

Although race is often associated with factors such as education, it remains a significant factor in itself in explaining income inequality even after controlling for such factors (Gradín 2013). Recent evidence shows that racial inequalities account for around 41 per cent of overall consumption

---

[2] Chatterjee et al. (2021) used survey, tax, and national accounts data.

inequality in South Africa, while education and labour market contribute about 29.9 per cent and 16.2 per cent respectively (World Bank 2022).

3.3 Labour market

Labour income is the primary source of income for the majority of South African households, and decomposition of income inequality by income sources shows that labour income is the primary contributor to overall income inequality (Leibbrandt, Finn et al. 2010), accounting for around 70 per cent of total income inequality in 2018 (World Bank 2022). As a result, evidence indicating persistently high and possibly increasing wage inequality in post-apartheid South Africa is consistent with evidence indicating persistently high and possibly increasing wage inequality over this period (Finn and Leibbrandt 2018; Finn et al. 2017; Kerr 2021; Leibbrandt et al. 2012; Wittenberg 2017). However, due to data comparability issues, the estimates of earnings inequality vary widely across data sources, with the Gini coefficient ranging from 0.5 to 0.7 (Kerr 2021).

Overall, the evidence suggests that wage inequality has remained persistently high and shows no signs of decline (see Kerr 2021). Gender and racial unemployment and wage disparities are particularly high, with women and Africans having relatively higher unemployment rates and being segregated in lower-paying jobs (see Gradín 2019, 2021a; Mosomi 2019; Ranchhod and Daniels 2021).

Several factors have been proposed to explain the continued high level of inequality in labour market outcomes, including sectoral composition shift, skills-biased technical change, labour market practices, and regulations (see Banerjee et al. 2008; Finn and Leibbrandt 2018; Bhorat et al. 2020b). For instance, the shift in sectoral composition from agriculture and mining to finance and services increased demand for more highly skilled labour. This shift in the labour market has been linked to an increase in the real wages of high-skilled workers and a fall in the wages of low-skilled workers between 1995 and 2005 (Banerjee et al. 2008). Finn and Leibbrandt (2018) show that returns to tertiary education is the main driver of earnings inequality in recent years through increasing the incomes of top income earners.

With the mounting challenge of creating more jobs in the formal sector, South Africa's labour market is characterized by an increasing demand for skilled labour while unskilled labour is experiencing significant unemployment. According to the broad definition, South Africa's unemployment rate was close to 30 per cent in 1994, with estimates ranging between 30 and 35 per cent between 2009 and 2017 (Heintz and Naidoo 2021). Despite widespread unemployment, the informal sector in South Africa has not been large enough to absorb growing labour force participation (Kingdon and Knight 2004; Rogan and Skinner 2018). Although there is debate about the definition and size of the informal sector, some estimates indicate that the share of the informal sector employment ranges between 16 per cent and 18 per cent between 2008 and 2014 (Rogan and Skinner 2018).

3.4 Spatial inequality

Because of the legacy of spatial injustice before and during the apartheid era, inequalities in income and other wellbeing outcomes in South Africa continue to have a strong spatial dimension. In particular, poverty and deprivation are concentrated mostly in areas formerly known as Bantustans or homelands areas. Homelands areas were created by the apartheid government following the Group Areas Act of 1950, which made residential segregation mandatory, and Africans were forcibly relocated to largely rural areas or townships within urban centres far from jobs. As a result of this policy, racial inequalities in employment access and commuting costs remain significant (Kerr 2017; Shilpi et al. 2018). With limited economic opportunities, segregated areas became

characterized by relatively high levels of poverty and deprivation. According to David et al. (2018), the rate of income poverty in the richest nine municipalities is less than 35 per cent, while that of the poorest 12 municipalities ranges between 83 per cent and 87 per cent. The poorest municipalities are located largely in homelands areas (Figure 5). Similarly, spatial inequalities are observed in access to basic services (see Shifa et al. 2021; Stats SA 2019).

Figure 5: Spatial distribution of income poverty (2011)
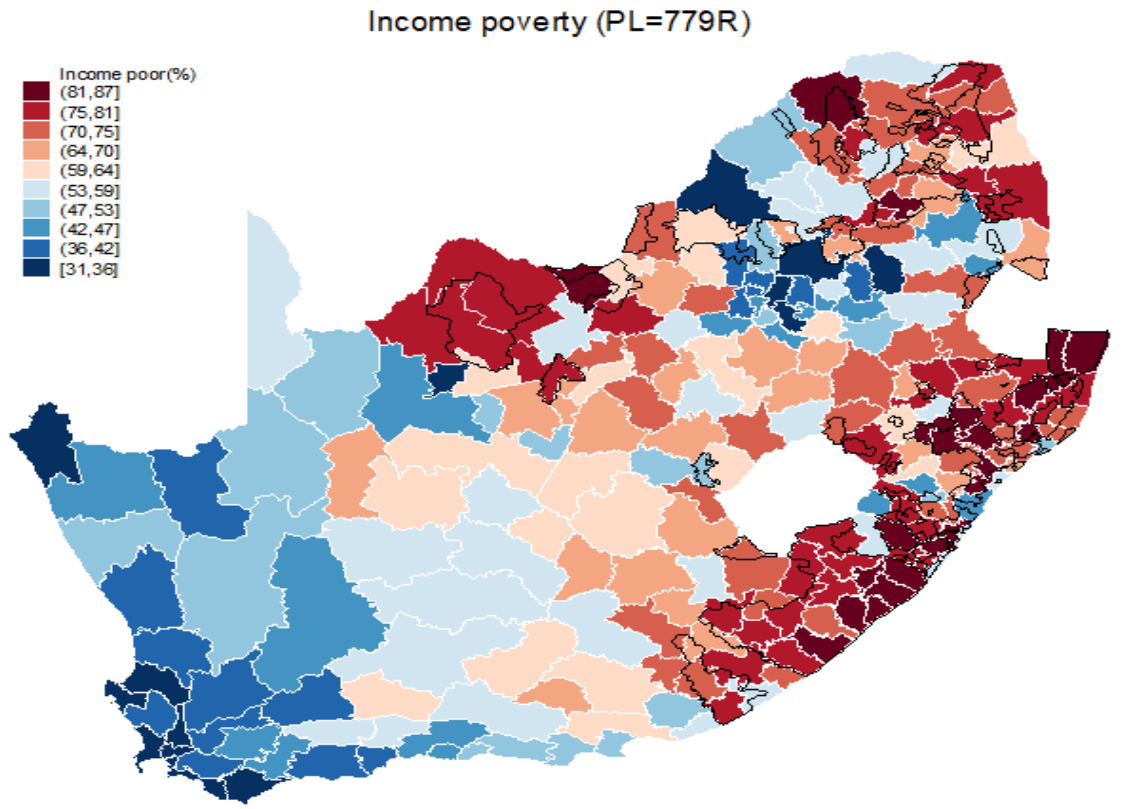
Source: reproduced with permission from David et al. (2018).

Overall, inequality in South Africa is rooted in the country's history, and political, social, and economic institutional factors have created multiple and intersecting dimensions of inequity. The various labour and social policies enacted to redress historical inequities helped to reduce income poverty and expanded access to basic services. Taxation and cash transfers are found to be progressive, contributing to poverty and inequality reduction (Chatterjee et al. 2021; Goldman et al. 2020; Maboshe and Woolard 2018). However, income and wealth inequality remain persistently high because the level of redistribution is unable to significantly reduce income and wealth disparities in post-apartheid South Africa (Chatterjee et al. 2021). Inherited factors like race, parental education, and location continue to drive income inequality and intergenerational mobility (Finn et al. 2017; Piraino 2015).

4 Trends in income inequality: Evidence from survey data

In this section, we compare income and expenditure inequality estimates that we calculate based on the publicly available datasets since 1993. For this purpose, we use data from the IES, LCS (Living Conditions Survey), PSLSD, and NIDS, which collect detailed information on both income and consumption data and are thus better for analysing inequality and poverty than other datasets which collect income data.[3] We make no adjustments to the income and expenditure variables provided in these datasets.

We then compare the inequality estimates we calculate with those reported in the WIID Companion. WIID provides inequality estimates for South Africa from the 1960s until 2017, spanning 23 survey years. The WIID Companion included in recent versions of WIID is useful for analysing changes in income distribution within nations, between countries, and globally. This paper's discussion is based on the recent version of WIID (May 2021) obtained from the UNU-WIDER website (UNU-WIDER 2021). WIID also reports consumption inequality measures in addition to income inequality estimates.

Both income and consumption inequality estimates are provided using different equivalent scale adjustments, resulting in more than one inequality estimate per year. Tables A1 and A2 in Appendix A provide the list of data sources used in WIID as well as other data sources available for inequality analysis in South Africa. In Appendix A, we also provide some discussion of the comparability of surveys used to estimate inequality in South Africa. Previous research has also examined the comparability of existing survey measures in South Africa (see Ardington et al. 2006; Kerr 2021; Van der Berg 2011; Van der Berg et al. 2008; Yu 2016, 2017).

4.1 The evolution of income shares (1993–2017)

Both income and consumption metrics have been used to examine income inequality in South Africa. The income concept in the IESs and LCSs refers to annual gross (pre-tax) income, including all benefits such as social grants, and imputed rental income from owner-occupied housing. In the case of PSLSD and NIDS, the income variable reflects regular monthly income, net of taxes, as well as imputed rental income from owner-occupied housing (Brophy et al. 2018). We calculate per capita values by dividing total household income and expenditure values by total household size.

Figures B1–B6 in Appendix B depict the distributions of per capita income and expenditure based on the IES, LCS, PSLSD, and NIDS datasets.[4] Overall, the income and expenditure distributions based on the IESs and the LCSs are similar throughout all survey years. The closeness in the distribution of income and expenditure data in the early IESs (1995 and 2000) may be attributable in part to the imputation process employed for missing and undeclared income and expenditure numbers (Appendix A). Given the imputation processes used before the 2005 IES, discussing income and consumption metrics as a different concept may make little sense.

Figure 6 shows the evolution of top and bottom income shares based on the two datasets. According to the IES and LCS, the share of income going to the bottom 50 per cent fell from 9.4 per cent in 1995 to 7.4 per cent in 2000, then remained stable before increasing slightly to 8.2 per

---

[3] The household surveys used in this paper are available from DataFirst: www.datafirst.uct.ac.za (accessed 29 April 2022).

[4] We use CPI data from Stats SA (2019) to adjust for prices: www.statssa.gov.za/publications/P0141/CPIHistory.pdf.

cent in 2015. Similarly, the income shares of the middle 40 per cent fell between 1995 and 2005, increased slightly in 2009, and have remained stable since. On the other hand, the income share of the top 10 per cent increased from 52.6 per cent in 1995 to 61.4 per cent in 2005, then declined, reaching 53.9 per cent in 2015.

Figure 6: The evolution of top and bottom income shares (1993–2017)
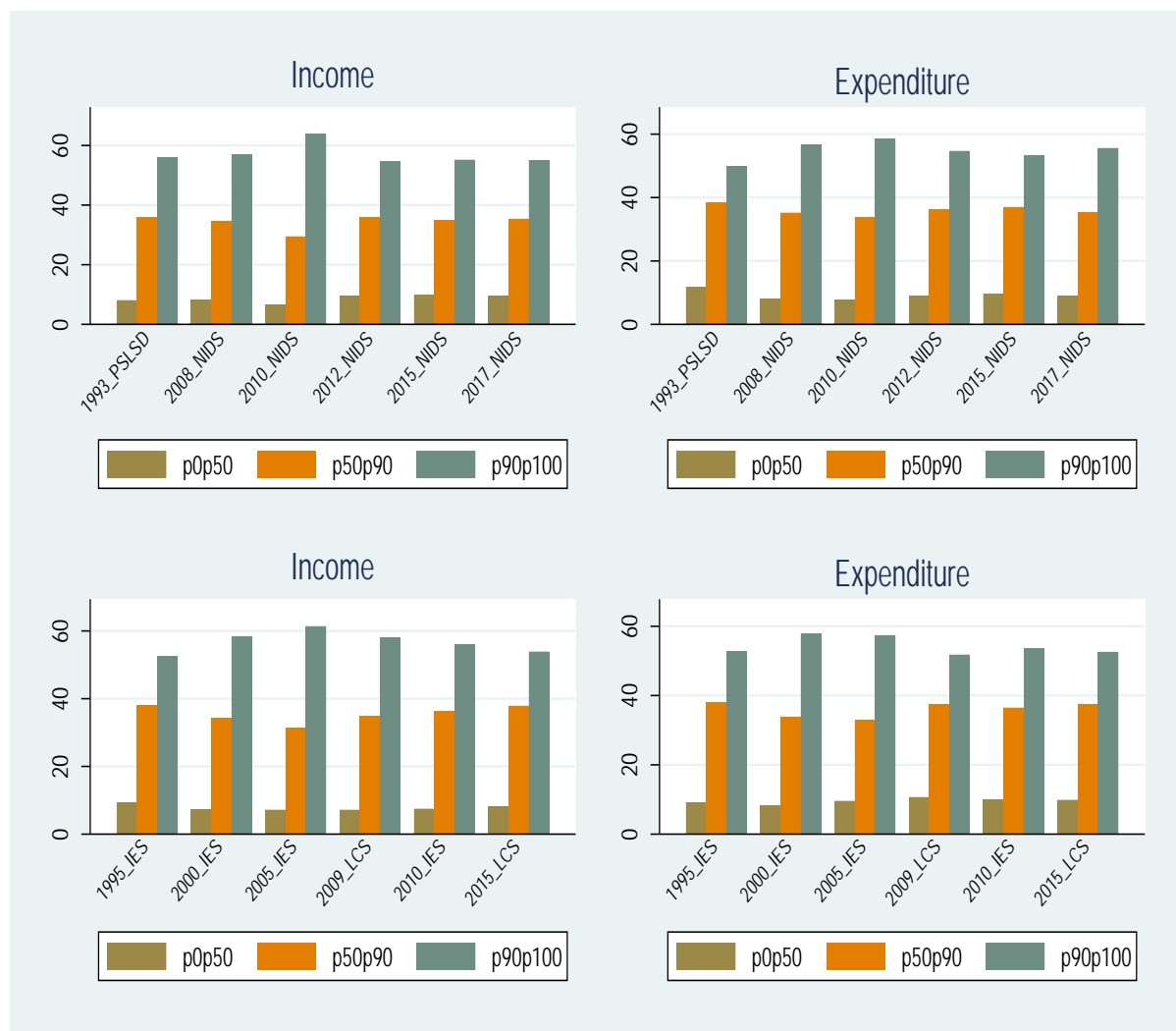
Source: authors' illustration based on survey data.

When we use data from PSLSD and NIDS, we see a similar pattern. The income shares of the top 10 per cent increased from 55.9 per cent in 1993 to 63.8 per cent in 2010, then decreased to 55.0 per cent in 2017, while the income share of the middle 40 per cent decreased from 36 per cent in 1993 to 29.5 per cent in 2010. It then increased to 35.9 per cent in 2012 and remained stable thereafter. The bottom 50 per cent's share of income decreased from 8.1 per cent in 1993 to 6.6 per cent in 2012 then increased to 9.5 per cent in 2012 and 2017.

Other studies that use household survey data to estimate top income shares also found a slight decrease in the income share of the top 10 per cent in recent years. Using data from the Luxembourg Income Study (LIS) database, Checchi et al. (2018) show that the income shares of the top 10 per cent and top 5 per cent declined between 2008 and 2015, while the figure for the bottom 50 per cent increased slightly. Similarly, according to Statistics South Africa's inequality report (Stats SA 2019), the top 10 per cent's share of consumption expenditure decreased from 57 per cent in 2005/06 to 52 per cent in 2014/15. However, as discussed earlier, studies that combine

tax and survey data find increasing top income shares and declining bottom 50 per cent income shares (Assouad et al. 2018; Bassier and Woolard 2021). Overall, regardless of the instruments employed to track the evolution of top income shares, the evidence is consistent in revealing that the income share of top 10 per cent remains above 50 per cent, while individuals in the bottom 50 per cent earn less than 10 per cent. The persistently higher top income shares indicate that economic growth has largely benefited top income earners.

4.2 Gini coefficient estimates

Figure 7 depicts Gini coefficient estimates for per capita income and consumption, while Figures B7 and B8 (Appendix B) present the corresponding Lorenz curves. The Gini coefficient estimates using the IES and LCS datasets show that inequality estimates based on income and consumption data were identical in 1995, but for the rest of the years, the Gini coefficient for income inequality became higher than the Gini coefficient for consumption inequality and had slightly different patterns.

Based on the IES and LCS datasets, the Gini coefficient for income increased from 65.5 in 1995 to 70.3 in 2000 and 71.9 in 2005, before declining to 69.0 in 2010 and 67.4 in 2015. Similarly, consumption inequality increased between 1995 and 2005, then decreased between 2005 and 2009, before increasing slightly in 2010 and 2015. Gini coefficient estimates based on the PSLSD and NIDS datasets show that inequality estimates based on the income and expenditure variables are very close, except in 1993 and 2010, when the Gini coefficient for income inequality was much higher than that of the consumption. The level of income inequality is roughly the same in 1993 and 2008; it increased significantly in 2010, before declining in subsequent years. Consumption inequality was lowest in 1993 but rose in 2008 and 2010, then fell in subsequent years.

Figure 7: Gini coefficient of per capita income and expenditure, IES, LCS, PSLSD, and NIDS
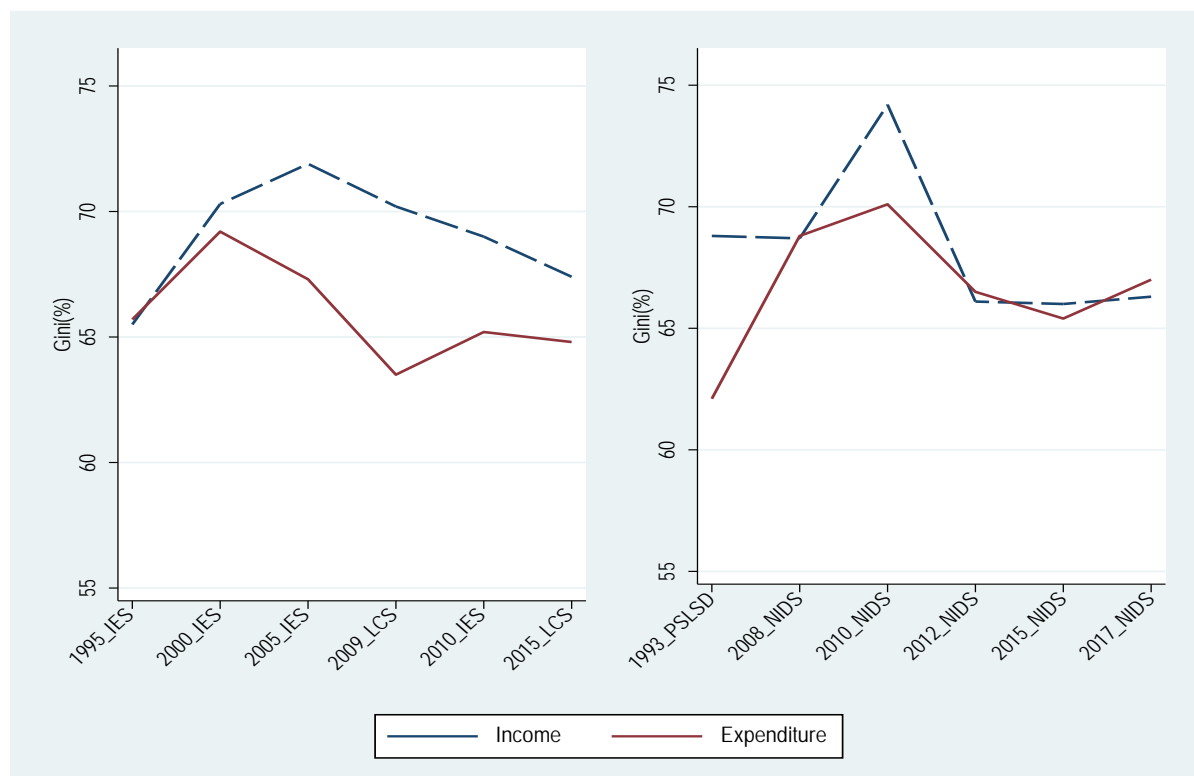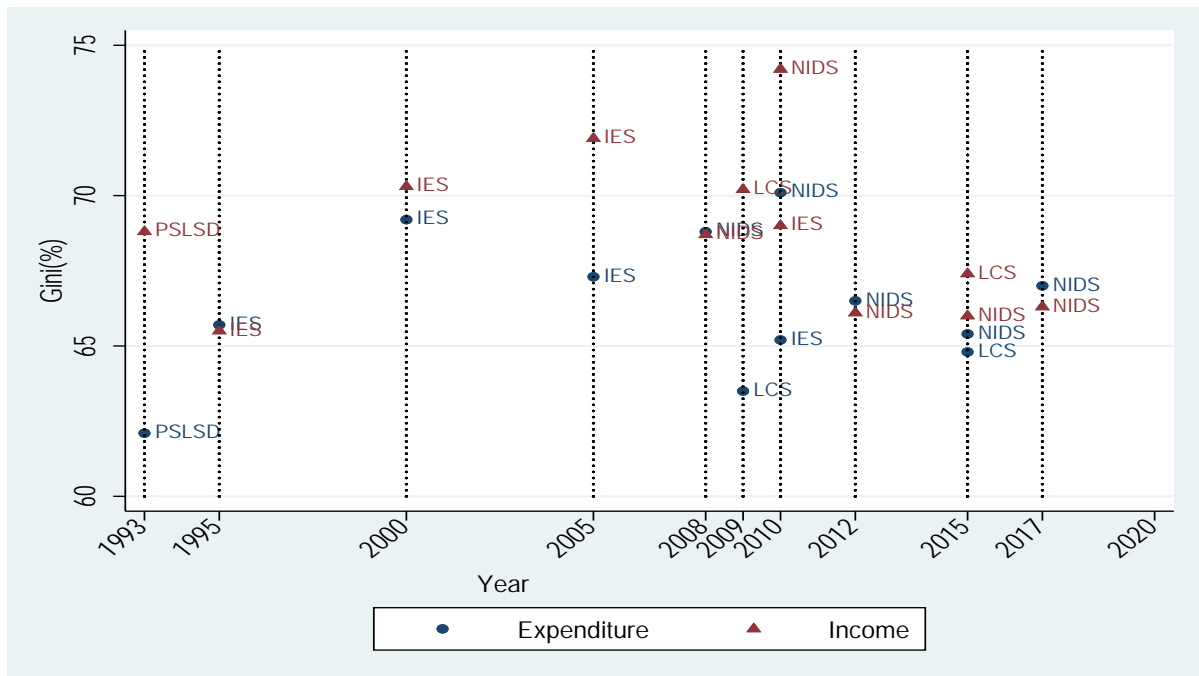

Source: authors' illustration based on survey data.

When comparing inequality trends across the two datasets, some disparities emerge. While income inequality remained roughly constant based on the 1993 PSLSD and 2008 NIDS estimates, the IES and LCS datasets show a rise in income inequality between 1995 and 2009. Consumption inequality, on the other hand, increased between 1993 and 2008 using the 1993 PSLSD and 2008 NIDS datasets, whereas it first increased between 1995 and 2000 then declined between 2000 and 2009 using the IES and LCS datasets.

Overall, when we examine inequality trends based on all of the surveys, we find no clear pattern for either income or consumption inequality estimates (Figure 8). Income inequality appears to have dropped between 1993 (PLSDS) and 1995 (IES), then increased in 2000 (IES) and 2005 (IES), then declined again in 2008 (NIDS), only marginally increasing in 2009 (LCS). The estimate in 2010 (NIDS) shows a large spike compared with both the 2008 (NIDS) and the 2009 and 2010 (LCSs) estimates, indicating that the estimate in 2010 (NIDS) may not be credible. Except for the 2015 (LCS) estimates, which show an increasing trend compared with the 2012 (NIDS) values, income inequality declined in 2012 and has remained close to the 2012 level since then.

Figure 8: Gini coefficient estimates, PSLSD (1993), IES (1995–2010), LCS (2009–15), and NIDS (2008–17)

Source: authors' illustration based on survey data.

The initial rise in income and consumption inequality based on the two datasets is consistent with previous research that documented rising trends in income inequality during the 1990s and 2000s. On the other hand, the stable Gini coefficient between 1993 and 2008 based on PSLSD and NIDS data differs from previous research that used the same data source and found that the Gini coefficient for income inequality increased between 1993 and 2008 (Hundenborn et al. 2018; Leibbrandt et al. 2012). These studies accounted for methodological changes in the collection of income data between the 1993 PSLSD and the 2008 NIDS data (see Appendix A). This indicates that our estimates are affected by the data comparability problem.

The inequality trends from NIDS datasets contradict the pattern documented by Checchi et al. (2018), who used data from the LIS database. According to Checchi et al. (2018), income inequality decreased between 2008 and 2010, increased in 2012, and then declined in 2015. However, they

used a different equalization scale, and the income variable was top-coded at ten times the median of non-equivalized income.

Although there is widespread agreement that income inequality increased in the early 1990s, the ten-point increase in the consumption Gini coefficient estimates (based on the PSLSD and NIDS datasets between 1993 and 2008) may reflect a measurement issue rather than an actual change in the level of inequality. For example, the PSLSD total expenditure was calculated as current expenditure net of taxes, excluding spending on durable goods and all forms of income tax liabilities. As a result, total consumption measures may be underestimated, resulting in lower estimates of consumption inequality. Similarly, the large spike in the Gini coefficient between 2008 and 2010 could be due partly to a measurement error given that there is no change in the survey methodology. The attrition rate was higher in 2010 compared with other NIDS waves. However, the attrition rate tends to be higher among those with higher incomes. Thus, it is not clear whether attrition is part of the explanation for why we observed such a large gap between the 2008 and 2010 NIDS datasets.

The introduction of imputed rental values for owner-occupied dwellings in the IES, which began in 2005, can be expected to raise income levels for individuals at the top of the distribution, increasing income inequality. For example, Yu (2017) shows that the Gini estimate for income decreased from 71.7 when imputed rental income was included to 71.0 when imputed rental values were excluded. Thus, when excluding the imputed rental income, income inequality seems the same between 2000 and 2005. However, given that the survey techniques for the 1995 and 2000 IESs were very similar (Yu 2017), it is unclear what caused an increase in both income and expenditure inequality between 1995 and 2000. Previous studies have also documented a decline in average household income levels during this period, suggesting that the decline was a real change (Leibbrandt et al. 2005). This is possible given that economic growth had only begun to recover during the period after a decade of negative GDP per capita growth rates and rising unemployment.

4.3   A Comparison of the ACEIR and WIID narratives

Figure 9 illustrates the Gini estimates from the WIID Companion as well as our (ACEIR's) estimations starting from 1993. Our calculations do not include inequality estimates from the 1996 and 2001 censuses. We use the 1995 IES for 1996, whereas the WIID Companion uses the 1996 census.

Overall, income inequality estimates based on the original and standardized WIID data follow a similar pattern. Likewise, with the exception of the 2000 estimates, the ACEIR and WIID inequality estimates follow very similar patterns. Except for 1993, 2000, and 2010, ACEIR's income inequality estimates are lower than those published by the WIID Companion. In 1996, we find a relatively large disparity between ACEIR and WIID Gini estimates, with the Gini coefficient from the WIID Companion being 69 and the estimate from ACEIR being 65. This discrepancy may be due to the use of different survey sources, as the ACEIR estimate is based on the 1995 IES while the WIID estimate is based on the 1996 census.

Figure 9: Gini estimates based on WIID Companion and ACEIR estimates
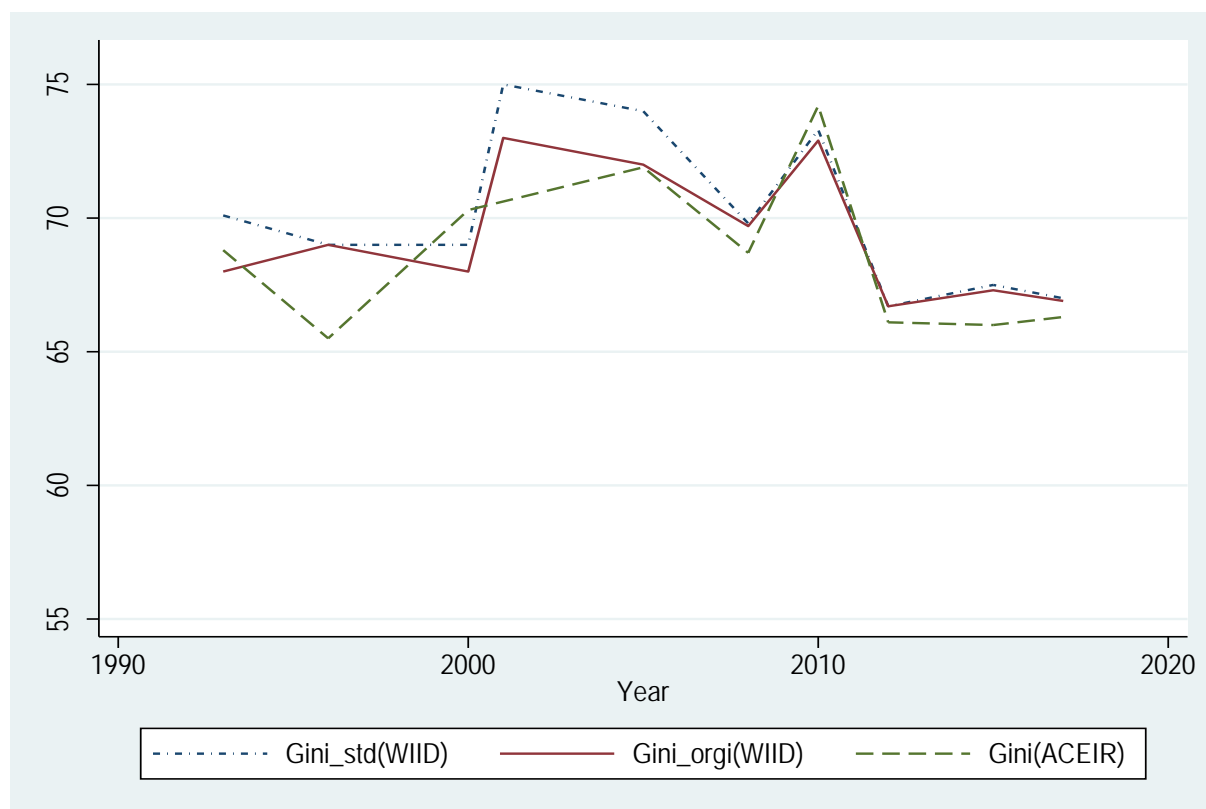
Note: Gini (ACEIR) represents our income Gini coefficient estimates based on data from the 1993 PSLSD, IES (1995, 2000, 2005), and NIDS (2008–17); Gini_orgi (WIID) represents income Gini estimates reported in the WIID database; Gini_std (WIID) represents the standardized Gini coefficient estimates reported in the WIID Companion.

Source: authors' own illustration based on data from UNU-WIDER (2021).

Data comparability also helps to explain the large increases in income inequality that occurred in a single year, from 68.0 in 2000 to 73.0 in 2001. The 2000 WIID estimate is based on the 2000 IES, whereas the 2001 estimate is based on the 2001 census, which collected income data using only a single question about income bands. These findings suggest that income inequality estimates based on census data are higher than those based on IES surveys for comparable survey years.

Overall, ignoring some of the survey years with large spikes (2001, 2005, and 2010), we can conclude that inequality has decreased only slightly in recent years compared with values in the early 1990s and early 2000s. Furthermore, WIID's standardization procedure changed the original inequality estimates mainly for survey years prior to 2008. Thus, despite the potential limitations of earlier surveys for estimating income inequality, the standardization process is useful because it allows one to evaluate income inequality trends beginning in the early 1960s, when representative household surveys aimed at analysing inequality and poverty were not available.

Figure 10 depicts standardized income Gini coefficient estimates from WIID as well as top income share estimates from WID. Between 1960 and 1987, the standardized income Gini estimates were obtained by raising the income Gini estimates published by Lachman and Bercuson (1992) using the Gini coefficient values reported by Whiteford and van Seventer (2000) in 1975 and 1985. However, the Gini estimates from Whiteford and van Seventer (2000) are based on interpersonal inequality estimates, whereas the estimates from Lachman and Bercuson (1992) are Gini coefficient estimates of racial inequalities.

The two estimates show different trends in inequality between 1975 and 1987. According to Whiteford and van Seventer (2000), the national Gini coefficient for income remained constant between 1975 and 1991 (i.e. 0.68). Lachman and Bercuson (1992), on the other hand, found a decreasing trend in income inequality, with the exception of rising inequality in the periods 1960–65 and 1975–1980. Consistent with Lachman and Bercuson's (1992) trend, the standardized Gini coefficient from WIID was very high in the 1960s and early 1970s, ranging from 0.72 to 0.76, before falling to 0.68 in 1975. Following an increase in 1980, the Gini estimate showed a declining trend until the end of the 1980s. Given that racial differences were the primary contributors to overall income inequality during this period, the overall downward trend in inequality between 1965 and 1987 is consistent with the narrowing of racial income inequalities documented during this time period (see Leibbrandt et al. 2009).

Figure 10: Standardized income Gini coefficient estimates and top income shares
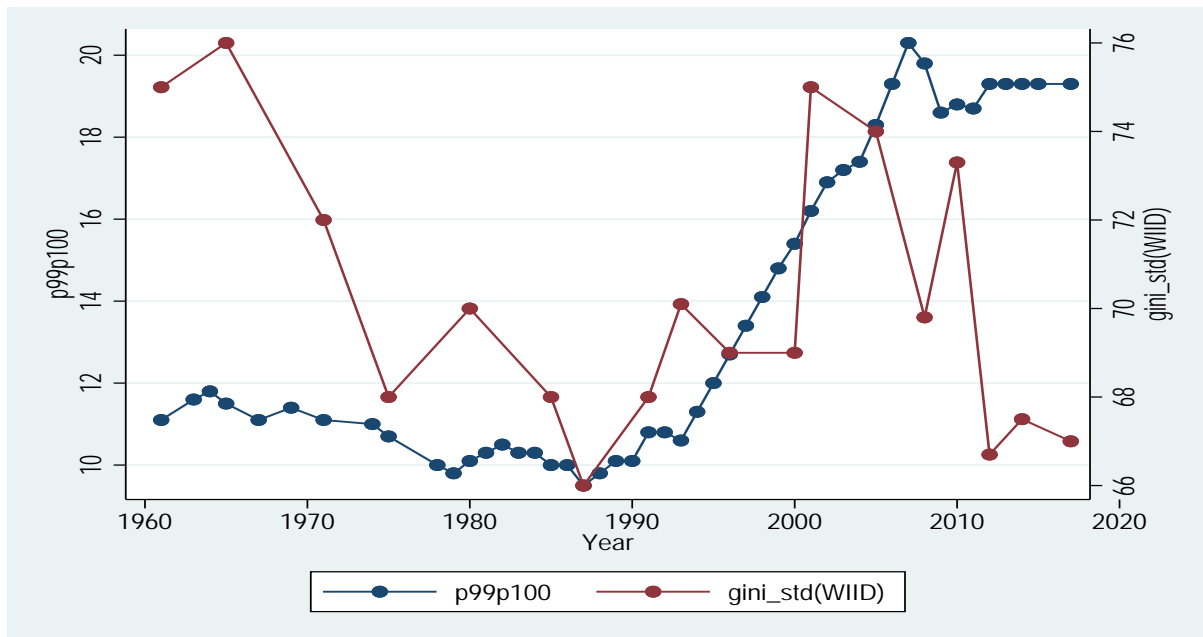
Source: author's illustration based on data from WID (2022) and UNU-WIDER (2021).

WIID estimates show an overall pattern of decreasing inequality during the 1970s and 1980s, which is consistent with evidence of declining top income share during this time period (Alvaredo and Atkinson 2022). The top 1 per cent income share increased slightly in the early 1960s, then fell between 1965 and 1987, before rising again afterwards. Beginning in the early 1990s, the top income share increased dramatically, rising from 10.1 per cent in 1990 to 20.3 per cent in 2007, before falling slightly during the final crises of 2008 (Figure 10). Thus, while the general downward trend in inequality prior to the 1990s is plausible, explaining the year-to-year volatility observed during this time period remains a challenge, regardless of whether we use the original or standardized inequality estimates. We do not have good survey data for this period to make more definitive (helpful) conclusions in this regard.

5   Conclusion

Although South Africa has census data and some survey data dating back to the early 1960s, nationally representative household surveys were not available until 1993. Databases such as the WIID have proved extremely valuable when analysing individual living standards and inequalities across time and across socioeconomic groups and, particularly, in locating South Africa's inequality evolution in an international perspective. However, the systematic study of long-term inequality and poverty in the South African context necessitates the use of many surveys with different methodologies for collecting data on income and expenditure. In this research, we document such methodology changes over time and across surveys using South African data.

We estimate income and consumption inequality using nationally representative survey data since 1993. Our inequality estimates across time are influenced in part by methodological modifications implemented across different surveys, making it difficult to draw definitive conclusions regarding inequality patterns over time. The results are not comparable within some of the same survey series or across surveys, with the latter being more problematic.

All of these datasets consistently measure levels of income (or expenditure) inequality that are extremely high in an international context. They thus strongly confirm that this has been and remains the reality of South African society. The same is true of the estimates in the WIID Companion that this paper has reviewed.

When comparing our income inequality estimates with those obtained from the WIID Companion, we can conclude that the pattern of inequality estimates over time are very similar, except for two survey years. However, the magnitude of the inequality estimates based on the WIID standardized values is higher than that of the ACEIR estimates, except in 2000 and 2010. Given that our comparison with the WIID standardized data is limited to a subset of data points collected since 1993, we cannot draw a definitive conclusion about how well the narrative from our analysis would match the narrative from the entire WIID combination series. However, triangulating data from previous studies that used survey data as well as tax and national account data to better measure income and wealth inequality appears to provide a consistent narrative about South Africa's income and wealth inequality trends. Despite fluctuations over time, the overall trend in post-apartheid South Africa is one of an increase in income and wealth inequality.

Overall, the levels of multidimensional inequality were very high in 1994 and remain so when appropriately benchmarked internationally. But they declined quicky from the late 1990s through to about 2010 and then flattened at the new lower (but still high) level. This decline in inequality was made possible due to large expansions in access to basic services and education. This tells a story that it is possible to intervene in the drivers of inequality but it takes a while for this impact on income inequality. Moreover, if this is not sustained through hard economic times, income inequality will start to rise again.

Thus, tracking trends in income and multidimensional inequality is crucial in supporting and monitoring progress against the policy imperative to reduce South African inequality from such stiflingly high levels.

References

Alvaredo, F., and A.B. Atkinson (2022). 'Top Incomes in South Africa in the Twentieth Century'. *Cliometrica*, 16: 477–546. https://doi.org/10.1007/s11698-021-00235-4

Ardington, C., D. Lam, M. Leibbrandt, and M. Welch (2006). 'The Sensitivity to Key Data Imputations of Recent Estimates of Income Poverty and Inequality in South Africa'. *Economic Modelling*, 23(5): 822–35. https://doi.org/10.1016/j.econmod.2005.10.009

Assouad, L., L. Chancel, and M. Morgan (2018). 'Extreme Inequality: Evidence from Brazil, India, the Middle East, and South Africa'. *AEA Papers and Proceedings*, 108: 119–23. https://doi.org/10.1257/pandp.20181076

Banerjee, A., S. Galiani, J. Levinsohn, Z. McLaren, and I. Woolard (2008). 'Why Has Unemployment Risen in the New South Africa?' *Economics of Transition*, 16(4): 715–40. https://doi.org/10.1111/j.1468-0351.2008.00340.x

Bassier, I., and I. Woolard (2021). 'Exclusive Growth? Rapidly Increasing Top Incomes amid Low National Growth in South Africa'. *South African Journal of Economics*, 89(2): 246–73. https://doi.org/10.1111/saje.12274

Bhorat, H., and C. Van der Westhuizen (2013). 'Non-Monetary Dimensions of Well-Being in South Africa, 1993–2004: A Post-Apartheid Dividend?' *Development Southern Africa*, 30(3): 295–314. https://doi.org/10.1080/0376835X.2013.817308

Bhorat, H., K. Lilenstein, M. Oosthuizen, and A. Thornton (2020a). 'Structural Transformation, Inequality, and Inclusive Growth in South Africa'. WIDER Working Paper 2020/50. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/807-8

Bhorat, H., K. Lilenstein, M. Oosthuizen, and A. Thornton (2020b). 'Wage Polarization in a High-Inequality Emerging Economy: The Case of South Africa'. WIDER Working Paper 2020/55. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/812-2

Brophy, T., N. Branson, R.C. Daniels, M. Leibbrandt, C. Mlatsheni, and I. Woolard (2018). 'NIDS Panel User Manual'. Release 2018. Version 1. Cape Town: Southern Africa Labour and Development Research Unit (SALDRU), University of Cape Town (UCT). Available at: www.nids.uct.ac.za/images/documents/20180831-NIDS-W5PanelUserManual-V1.0.pdf (accessed 15 March 2022).

Chatterjee, A., L. Czajka, and A. Gethin (2021). 'Can Redistribution Keep Up with Inequality? Evidence from South Africa, 1993–2019'. WID (World Inequality Database) Working Paper 2021/20. Paris: World Inequality Lab. Available at: https://wid.world/methodology/#library-key-reports (accessed 11 November 2022).

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

Checchi, D., A. Cupak, T. Munzi, and J. Gornick (2018). 'Empirical Challenges Comparing Inequality across Countries'. In C. Gradín, M. Leibbrandt, and F. Tarp (eds), *Inequality in the Developing World*. Oxford: Oxford University Press.

David, A., N. Guilbert, N. Hamaguchi, Y. Higashi, H. Hino, M. Leibbrandt, and M. Shifa (2018). 'Spatial Poverty and Inequality in South Africa: A Municipality Level Analysis'. SALDRU Working Paper 221. Cape Town: SALDRU, UCT.

Deininger, K., and L. Squire (1996). 'A New Data Set measuring Income Inequality'. *The World Bank Economic Review*, 10(3), 565–91. https://doi.org/10.1093/wber/10.3.565

Feinstein, C.H. (2005). *An Economic History of South Africa: Conquest, Discrimination, and Development*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139165457

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

Finn, A., M. Leibbrandt, and V. Ranchhod (2017). 'Patterns of Persistence: Intergenerational Mobility and Education in South Africa'. Updated, Version 3. SALDRU Working Paper 175/NIDS Discussion Paper 2016/2. Cape Town: SALDRU, UCT.

Goldman, M., I. Woolard, and J. Jellema (2020). The Impact of Taxes and Transfers on Poverty and Income Distribution in South Africa 2014/2015. AFD Research Papers 2020/198. Paris: AFD (Agence française de développement). Available at: www.afd.fr/sites/afd/files/2021-02-04-52-37/taxes-transfers-poverty-income-distribution-south-africa.pdf (accessed 11 November 2022).

Gradín, C. (2013). 'Race, Poverty and Deprivation in South Africa'. *Journal of African Economies*, 22(2): 187–238. https://doi.org/10.1093/jae/ejs019

Gradín, C. (2019). 'Occupational Segregation by Race in South Africa after Apartheid'. *Review of Development Economics*, 23(2): 553–76. https://doi.org/10.1111/rode.12551

Gradín, C. (2021a). 'Occupational Gender Segregation in Post-Apartheid South Africa'. *Feminist Economics*, 27(3): 102–33. https://doi.org/10.1080/13545701.2021.1906439

Gradín, C. (2021b). 'WIID Companion: March 2021: Data Selection'. WIDER Technical Note 2021/4. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/WTN/2021-4

Heintz, J., and K. Naidoo (2021). 'Unemployment in South Africa'. In A. Oqubay, F. Tregenna, and Imraan Valodia (eds), *The Oxford Handbook of the South African Economy*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780192894199.013.7

Hundenborn, J., M.V. 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

Kerr, A. (2017). 'Tax(i)ing the Poor? Commuting Costs in South African Cities'. *South African Journal of Economics*, 85(3): 321–40. https://doi.org/10.1111/saje.12161

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

Kingdon, G.G., and J. Knight (2004). 'Unemployment in South Africa: The Nature of the Beast'. *World development*, 32(3): 391–408. https://doi.org/10.1016/j.worlddev.2003.10.005

Lachman., D., and K. Bercuson (eds) (1992). 'Economic Policies for a New South Africa'. IMF Occasional Papers 91. Washington, DC: IMF. https://doi.org/10.5089/9781557751980.084

Leibbrandt, M., J.A. Levinsohn, and J. McCrary (2005). 'Incomes in South Africa since the Fall of Apartheid'. NBER Working Paper 11384. Cambridge, MA: National Bureau of Economic Research (NBER). Available at: www.nber.org/papers/w11384 (accessed 3 December 2022). https://doi.org/10.3386/w11384

Leibbrandt, M., L. Poswell, P. Naidoo, M. Welch, and I. Woolard (2006). 'Measuring Recent Changes in South African Inequality and Poverty using 1996 and 2001 Census Data'. In H. Bhorat and R. Kanbur (eds), *Poverty and Policy in Post-Apartheid South Africa*. Cape Town: HSRC Publishing.

Leibbrandt, M., I. Woolard, and C. Woolard (2009). 'A Long-Run Perspective of Contemporary Poverty and Inequality Dynamics'. In J. Aron, B. Kahn, and G. Kingdon (eds), *South African Economic Policy under Democracy*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199551460.003.0010

Leibbrandt, M., A. Finn, J. Argent, and I. Woolard (2010). 'Changes in Income Poverty over the Post-Apartheid Period: An Analysis based on Data from the 1993 Project for Statistics on Living Standards and Development and the 2008 Base Wave of the National Income Dynamics Study'. *Studies in Economics and Econometrics*, 34(3): 25–43. https://hdl.handle.net/10520/EJC21508

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. https://doi.org/10.1787/5kmms0t7p1ms-en

Leibbrandt, M., A. Finn, and I. Woolard (2012). 'Describing and Decomposing Post-Apartheid Income Inequality in South Africa'. *Development Southern Africa*, 29(1): 19–34. https://doi.org/10.1080/0376835X.2012.645639

Leibbrandt, M.V., V. Ranchhod, and P. Green (2018). 'Taking Stock of South African Income Inequality'. WIDER Working Paper 2018/184. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2018/626-5

Maboshe, M., and I. Woolard (2018). 'Revisiting the Impact of Direct Taxes and Transfers on Poverty and Inequality in South Africa'. WIDER Working Paper 2018/79. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2018/521-3

McGrath, M. (1984). 'Inequality in the Size Distribution of Incomes in South Africa'. Staff Paper 2. Durban: Development Studies Unit, Centre for Applied Social Sciences. Available at: https://core.ac.uk/download/pdf/19917711.pdf (accessed 3 December 2022).

Mosomi, J. (2019). 'Distributional Changes in the Gender Wage Gap in the Post-Apartheid South African Labour Market'. WIDER Working Paper 2019/17. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2019/651-7

Piraino, P. (2015). 'Intergenerational Earnings Mobility and Equality of Opportunity in South Africa'. *World Development*, 67: 396–405. https://doi.org/10.1016/j.worlddev.2014.10.027

PSLSD (Project for Statistics on Living Standards and Development) (1994). *South Africans Rich and Poor: Baseline Household Statistics*. Cape Town: SALDRU, UCT.

Ranchhod, V., and R.C. Daniels (2021). 'Labour Market Dynamics in South Africa at the Onset of the COVID-19 Pandemic'. *South African Journal of Economics*, 89(1): 44–62. https://doi.org/10.1111/saje.12283

Rogan, M., and C. Skinner (2018). 'The Size and Structure of the South African Informal Sector 2008–2014: A Labour-Force Analysis'. In F. Fourie (ed.), *The South African Informal Sector: Providing Jobs, Reducing Poverty*. Pretoria: HSRC Publishing.

Seekings, J., M. Leibbrandt, and N. Nattrass (2004). 'Income Inequality after Apartheid'. CSSR (Centre for Social Science Research) Working Paper 75. Cape Town: SALDRU, UCT. http://hdl.handle.net/11090/643

Shifa, M., A. David, and M. Leibbrandt (2021). 'Spatial Inequality through the Prism of a Pandemic: Covid-19 in South Africa'. *Scientific African*, 13: e00949. https://doi.org/10.1016/j.sciaf.2021.e00949

Shifa, M., and M. Leibbrandt (2022). 'Spatial Inequality in Sub-Saharan Africa'. *African Geographical Review*, online: 15 July. https://doi.org/10.1080/19376812.2022.2099916

Shilpi, F., L. Xu, R. Behal, and B. Blankespoor (2018). 'People on the Move: Spatial Mismatch and Migration in Post-Apartheid South Africa'. World Bank Urbanisation Review Paper 1. Pretoria: National Treasury.

Simkins, C. (2004). 'What Happened to the Distribution of Income in South Africa between 1995 and 2001?' Unpublished paper. Available at: https://sarpn.org/documents/d0001062/P1175-simkins_Nov2004.pdf (accessed 22 June 2022).

Stats SA (Statistics South Africa) (1995) *Income and Expenditure Survey: Questionnaire*. Pretoria: Stats SA.

Stats SA (1996). 'Population Census 1996: Questionnaire'. Pretoria : Stats SA. Available at: https://microdata.worldbank.org/index.php/catalog/915/related-materials (accessed 12 August 2022).

Stats SA (1997). *Earning and Spending in South Africa: Selected Findings of the 1995 Income and Expenditure Survey*. Pretoria: Stats SA.

Stats SA (2001). *Population Census: Metadata Report*. Pretoria: Stats SA.

Stats SA (2008a*). Income and Expenditure of Households 2005/6: Analysis of Results*. Pretoria: Stats SA.

Stats SA (2008b). *Income and Expenditure of households 2005/6: Statistical Release P0100*. Pretoria: Stats SA.

Stats SA (2011b). *Population Census: Metadata Report*. Pretoria: Stats SA.

Stats SA (2012a). *Income and Expenditure of Households 2010/2011: Statistical Release P0100*. Pretoria: Stats SA.

Stats SA (2012b). *Poverty Profile of South Africa: Application of the Poverty Lines on the LCS of 2008/9*. Pretoria: Stats SA.

Stats SA (2017a). *Living Conditions of Households in South Africa: An Analysis of Household Expenditure and Income Data using the LCS 2014/15*. Pretoria: Stats SA.

Stats SA (2017b). *Poverty Trends in South Africa: An Examination of Absolute Poverty between 2006 and 2015*. Pretoria: Stats SA.

Stats SA (2019). *Inequality Trends in South Africa: A Multidimensional Diagnostic of Inequality*. Pretoria: Stats SA.

Treiman, D.J., M. McKeever, and E. Fodor (1996). 'Racial Differences in Occupational Status and Income in South Africa, 1880 and 1881'. *Demography*, 33(1): 111–32. https://doi.org/10.2307/2061717

UNU-WIDER (2021). 'World Income Inequality Database (WIID) Companion'. Version 31 May 2021. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/WIIDcomp-310521

Van der Berg, S., M. Louw, and D. Yu (2008). 'Post-Transition Poverty Trends based on an Alternative Data Source 1'. *South African Journal of Economics*, 76(1): 58–76. https://doi.org/10.1111/j.1813-6982.2008.00161.x

Van der Berg, S. (2011). 'Current Poverty and Income Distribution in the Context of South African History'. *Economic History of Developing Regions*, 26(1): 120–40. https://doi.org/10.2307/2061717

Whiteford. A., and D.E. van Seventer (2000). 'South Africa's Changing Income Distribution in the 1990s'. *Studies in Economics and Econometrics*, 24(3): 7–30. https://doi.org/10.1080/03796205.2000.12129274

WID (World Inequality Database) (2022). 'Data Tables'. Available at: https://wid.world (accessed 24 November 2022).

Wilson, F. (2011). 'Historical Roots of Inequality in South Africa'. *Economic History of Developing Regions*, 26(1): 1–15. https://doi.org/10.1080/20780389.2011.583026

Wittenberg, M. (2017). '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

Wittenberg, M., and M. Leibbrandt (2017). 'Measuring Inequality by Asset Indices: A General Approach with Application to South Africa'. *Review of Income and Wealth*, 63(4): 706–30. https://doi.org/10.1111/roiw.12286

World Bank (2022). *Inequality in Southern Africa: An Assessment of the Southern African Customs Union*. Washington DC: The World Bank. https://doi.org/10.1596/37283

Yu, D. (2016). 'Factors influencing the Comparability of Poverty Estimates across Household Surveys'. *Development Southern Africa*, 33(2): 145–65. https://doi.org/10.1080/0376835X.2015.1120646

Yu, D. (2017). 'The Comparability of Income and Expenditure Surveys 1995, 2000 and 2005/2006'. Stellenbosch Economic Working Papers 11/08. Stellenbosch: Stellenbosch University. http://hdl.handle.net/10566/5968.

Appendix A

A.1   List of datasets used in the WIID and what is available

Most inequality estimates used in WIID before the early 1990s were based on Lachman and Bercuson (1992), which cited Nel and Van Wyk (1984). Nel and Van Wyk used South African Bureau of Market Research (BMR) data for 1975, 1980, and 1985 and IMF interpolations. Population censuses were also used to estimate income inequality in earlier years for South Africa. However, African households were not asked the income question in censuses conducted before 1980. In the censuses of 1980, 1985, and 1991, the former homeland areas were not represented, excluding large sections of the black population (Treiman et al. 1996), rendering any analyses unrepresentative of the entire South African population.

Table A1: Description of WIID data sources used for South Africa

Year	No. of estimates	WIID Companion	Source on income inequality Gini	Survey
1960	1	yes	Lachman and Bercuson 1992	Not specified
1965	4	yes (1 obs)	Lachman and Bercuson 1992	Not specified
1970	1	yes	Lachman and Bercuson 1992	Not specified
1975	2	yes (1 obs)	Whiteford and van Seventer 2000	Census 1975
1980	1	yes	Lachman and Bercuson 1992	BMR
1985	1	yes	Lachman and Bercuson 1992	BMR
1987	1	yes	Lachman and Bercuson 1992	Not specified
1990	1	No	Web 14 June 1999	SHE (Survey of Household Expenditures)
1991	1	yes	Whiteford and van Seventer 2000	Census 1991
1993	2	yes (both)	Hundenborn et al. 2016	PSLSD
1995	1	No	Web 14 June 1999	IES
1996	2	yes (both)	Whiteford and van Seventer 2000	Census 1996
1997	1	No	Deininger and Squire 1996	October Household Survey (OHS)
2000	1	yes	Leibbrandt, Woolard et al. 2010	IES
2001	2	yes (both)	Leibbrandt et al. 2009	Census 2001
2005	2	yes (both)	Leibbrandt et al., Woolard et al. 2010	IES
2008	13	yes (1 obs)	Own construction based on LIS[b] database	NIDS
2009	1	yes	PovcalNet	Not specified
2010	13	yes (1 obs)	Own construction based on LIS database	NIDS
2011	2	No	Commitment to Equity Project	IES
2012	12	yes (1 obs)	Own construction based on LIS database	NIDS
2015	16	yes (1 obs)	Own construction based on LIS database	NIDS
2017	12	yes (1 obs)	Own construction based on LIS database	NIDS

Source: authors' construction based on UNU-WIDER (2021).

Table A2 lists nationally representative household surveys available since 1993 for assessing income inequality and other living standard measures. The 1993 PSLSD, 1994 OHS, and 1996 census were South Africa's first nationally representative surveys. The OHS, done between 1994 and 1999, was split into two surveys: the General Household Survey (GHS), conducted annually since 2002, and the Labour Force Survey (LFS), conducted between 2000 and 2007. Since 2008, the LFS has been conducted quarterly as the Quarterly Labour Force Survey (QLFS).

The GHS collected information on total household income and expenditure as a point value or in bands. On the other hand, the population censuses of 1996, 2001 and 2011, and the 2007 and 2016 Community Surveys (CS) collected details of personal income in bands. The LFS and QLFS primarily collected earnings data and were used to analyse earnings inequality. Among these datasets, only the OHS from 1997, the PSLSD, and the censuses from 1996 and 2001 were utilized to assess income inequality in the WIID.

Beginning in 2000, the WIID income inequality estimates were based on data from the IES (2000, 2005, and 2010), LCS (2009), and the NIDS collected between 2008 and 2017. Unlike most prior surveys, which collected income data in bands and expenditure data as a single aggregate household-level estimate, these datasets collected detailed information on income from multiple sources and expenditure based on a detailed list of consumption items.

Table A2: Potential survey datasets for the analysis of inequality in South Africa

Survey	Years available	Used in WIID	Income	Consumption
PSLSD	1993	Yes	Exact amounts	Exact amounts
OHS	1994–99 (annual)	Only 1997	Earnings income 1994–98; income bands in 1999	Bands
GHS	Since 2002 (annual)	No	Earnings income	Bands
LFS	Since 2000–07 (biannual)	No	Earnings income	Bands (2001, 2003, 2004)
QLFS	Since 2008 (quarterly)	No	Earnings income	No consumption
Census	1996, 2001, 2011	Yes	Bands	No consumption
IES	1995, 2000, 2005, 2010	Yes	Exact amounts and bands	Exact amounts and bands (2010)
LCS	2008/09 and 2014/15	Yes	Exact amounts	Exact amounts
NIDS	2008–17	Yes	Exact amounts	Exact amounts and bands
CS	2007 and 2016	No	Bands	No consumption

Source: authors' construction based on Yu (2017).

A.2 Within-country comparability of the measurements over time

In this section, we look at methodological issues affecting the comparability of income and consumption measures across surveys. Although all of the datasets listed in Table A2 can be used to estimate income or consumption inequality, our discussion in this section focuses primarily on surveys that have been used to estimate income and consumption inequality in the WIID Companion since 1993.

Comparability of IES and LCS

The IESs and the LCSs are the main datasets that collect detailed information on household expenditure and income in South Africa. Statistics South Africa conducted the IES in 1995, 2000, 2005, and 2010. The main purpose was to update and reweight the Consumer Price Index (CPI) (Stats SA 2012a, b) and the surveys were not necessarily to measure poverty and inequality. On the other hand, the LCSs introduced in 2008/09 and collected again in 2014/15 were designed to measure multidimensional poverty (Stats SA 2017b) and collected detailed information on income, expenditure, and other wellbeing indicators.

The questionnaire structure used to collect information on income for the IESs (1995, 2000, 2005, and 2010) and the LCSs (2009 and 2015) is very similar. In general, the annual gross income of individuals in the household was collecting using the following income categories:

1. salaries and wages;
2. net profit from business/farming;
3. income from letting fixed property;
4. royalties;
5. interest received and accrued on deposits, loans, and savings certificates;
6. dividends on shares (listed and unlisted companies);
7. regular receipts from pension;
8. other social welfare grants;
9. alimony, maintenance, and similar allowances received from divorced spouse, family members, etc., living elsewhere, etc.;
10. other income.

There are some differences in the methodologies used across the IESs that should be considered when comparing poverty and inequality estimates over time (See Table A3). While a lot of methodological changes can be observed for South African surveys conducted in the pre-2000 relative to those in the post-2005 period, the LCSs benefited from the questionnaire design of the IES of 2005. The LCS of 2008/09 used the income and expenditure modules of the IES of 2005 (Stats SA 2017b). Therefore, except for a few changes in 2014/15, the income and expenditure modules remained largely the same between the 2009 and 2015 LCS.

Table A3: Methodological comparisons between the IES surveys

		IES 1995	IES 2000	IES 2005	IES 2010
Visits per household		1	1	6	4
Survey period		Oct 1995 (one month)	Oct 2000 (one month)	Sept 2005 – Aug 2006 (one year)	Sept 2010 – August 2011 (one year)
Classification of expenditure items		Standard Trade Classification	Standard Trade Classification	COICOP	COICOP (Classification of Individual Consumption according to Purpose)
Methodology		Recall	Recall	Diary and recall (four weekly diaries)	Diary and recall (two weekly diaries)
Expenditure collection method	Goods		Payment approach	Acquisition approach	Acquisition approach
	Services		Payment approach	Payment approach	Payment approach
	Own production		Consumption approach	Consumption approach	Consumption approach
Housing expenditure		Mortgage payments	Mortgage payments	Imputed rent	Imputed rent
Income collection questions		Regular income for past 12 months 'Other' income for past 12 months	Regular income for past 12 months 'Other' income for past 12 months	Regular income for previous month Regular income for 11 months prior to survey months Other income for past 12 months	Total income from salaries/wages and/or income from business Income bands for salaries/wages and/or income from business Total income from other sources 12 months prior to survey Income bands (total income from other sources) Income of persons not currently household members in 12 months prior to survey Annual 'other' income for past 12 months
Number of respondents used to capture income		Maximum of five respondents allowed to answer income question	Income collected from only ten household members	No restrictions	No restrictions
Actual amounts or Bands		Actual amounts	Actual amounts	Actual amounts	Actual amounts and income bands
Sample size		29,582 households (1991 census)	26,263 households (1996 census)	21,144 households (2001 census)	25,328 households (2001 census)

Sources: authors' construction based on Stats SA (2008, 2012); Yu (2017).

In 1995, only five household members responded to the IES income questions, reporting their annual regular income and 'other' income (Stats SA 1995). Individual incomes were then aggregated to get total household income. After 1995, the restriction on only using five household members was removed. Income bands were added to the 2010 IES as a follow-up to one-shot

income questions that had not been used in previous IESs. The introduction of the back-up question on income bands improved the collection of income data in the 2010 IES. However, these income band questions were not included in subsequent LCSs.

The IES captures housing expenditure differently, potentially affecting the comparability of expenditure estimates. In the 1995 and 2000 IESs, mortgage payments were used as part of housing costs for owner-occupied households (Stats SA 2008b). Homeowners reported their capital payments and interest separately; however, the reporting on these two was not always reliable. This was changed in 2005 when international standards were applied and rent imputations for owner-occupied properties were used instead (Stats SA 2008b). In 2005, Stats SA used 7 per cent of the value of the house to estimate annual rent for owner-occupied dwelling units.

Similar imputation approaches were used for the 2010 IES, and the 2009 and 2015 LCSs. However, while in the LCSs a single rental yield was applied across the board, in the case of the 2010 IES a segmented approach was used in the application of rental yields. The segmentation was done by type of housing and province, except for Limpopo and Northwest, where an average national rental yield was used (Stats 2012a). In 2005, using the new approach for calculating housing costs for owner-occupied dwellings, housing expenditure increased by 18 per cent and 26 per cent for those in the first and fifth expenditure quintiles, respectively (Stats SA 2008a). While this is expected to reduce poverty estimates over time, the impact on inequality may be ambiguous, as expenditure increased for both the rich and the poor.

There are also significant differences in how consumption data was collected between IESs. The IES of 1995 and 2000 collected expenditure data using recall methods. In 2005 and 2010, the diary method was introduced, and both the recall and diary methods were used. There are further differences in the time period used to keep diaries in the two years. While four weekly diaries were used in the IES of 2005, this was reduced to two weeks' diaries in 2010 to reduce respondent fatigue (Stats SA 2012a). The LCSs used the same approach as the 2010 IES for collecting data on expenditure. While the 2008/09 LCS used four weeks' diaries, the 2014/15 LCS used two weeks' diaries.

The adoption of the diary technique, which can better track consumption expenditure, may have had an ambiguous effect on estimates of income and poverty levels. When compared with 1995 and 2000, per capita expenditure increased in 2005 and 2010, particularly for those at the bottom of the distribution (Figure B3). However, which component of expenditure changed because of the use of the diary method is debatable. For example, using the diary method in the IESs resulted in an underestimation of food expenses but an increase in transportation expenditure (see Yu 2016).

In addition, there are some differences in how undeclared income/expenditure was dealt with across the surveys. We find no missing household-level income and expenditure values in any of the publicly available datasets. However, there are households with zero incomes for some of the surveys (Table A4). In most situations, it is unclear from the metadata file or the reports how many families reported missing or zero incomes, and what steps were taken to impute for zero individual and household incomes, if any.

Table A4: Households with zero income and expenditure values

Year	Income	Expenditure	No. of households
IES_1995	-	-	29,582
IES_2000	-	34	26,263
IES_2005	7	-	21,144
LCS-2009	378	-	23,708
IES_2010	73	-	23,655
LCS_2015	47	-	21,706
Total	505	34	146,058

Source: authors' construction based on survey data.

In the IES of 1995, if a household did not declare its total annual expenditure but provided its total annual income, total annual income was used as a proxy for annual household expenditure (Stats SA 1997). The same logic was applied if a household did not declare its annual income but did provide annual expenditure. In the 2000 IES, the final income variable is a derived variable: 'If the total recorded expenditure exceeded the total recorded income, or where no income was recorded, the total income was estimated as being equal to the total expenditure.' This suggests that if total income was missing, total expenditure was used to impute for the missing income. Since 2005 (in the 2005 and 2010 IES and LCSs), the imputation approach has been used to impute for missing income where an individual reported having a source of income but did not declare the amount. Missing income was generally approximated using average values from households with similar characteristics. To match households, variables such as province, settlement type, type of dwelling unit, number of rooms, and house value were employed.

Regarding expenditure, imputations were done for missing diaries—a household had to complete at least two weeks' diaries and the main questionnaire to be included in the IES. Thus, households that completed fewer than two weeks of diaries were excluded and treated as non-respondents. Missing diaries for households with at least two or three weeks' diaries were imputed. This was done as follows: if a household had diary information for two weeks, a donor household was randomly selected from a group of households with similar characteristics to donate information for the two missing diaries. The characteristics used to match households for imputations were province, type of area, type of dwelling, household size, expenditure patterns of the available diary information, and access to facilities and services. A similar approach was used in the case of the LCS, except that households were required to complete only at least one week's diary to be induced in the 2015 LCS, and imputation was done for the missing weeks.

Another change across the IESs and LCSs was a significant decrease in response rates following the introduction of the diary technique. Response rates in the LCSs were aggravated further by the length of the questionnaire, with lower response rates recorded among high-income areas (Stats SA 2017a). In the 2014/15 LCS, for example, the response rate for Gauteng was 65.3 per cent while the corresponding figure for the 2008/09 LCS was 79.7 per cent.

Comparability of PSLSD and NIDS

The NIDS is a South African nationally representative panel of individuals that provides detailed information on living conditions and labour market status of individuals and households over time (Brophy et al. 2018). It was conducted in 2008, 2010, 2012, 2014, and 2017 and in the WIID, all waves of NIDS are used. NIDS collects information on income and expenditure using various income and expenditure categories. Both actual amounts and bands are used to collect income data. In addition, respondents are asked to provide total household income and expenditure as a 'one-shot' amount.

In general, NIDS collects income from the following categories and creates an aggregated household total income variable:

1. labour market income (main job, causal jobs, self-employment etc.);
2. government grant (disability, child support, foster care, care dependent);
3. other income from government (unemployment insurance fund, worker's compensation);
4. investment income (interest/dividend income, rental income, private pension and annuities);
5. remittance income;
6. subsistence agriculture income (including value of own production);
7. imputed rent for owner-occupied housing.

The total household income variable in NIDS collects regular monthly income and is derived from the above income categories from three questionnaires; Adult, Proxy, and household (Brophy et al. 2018). NIDS in all the five waves used three ways of aggregating individual-level income data to get household total income:

1. Where all adult household resident members were successfully interviewed, household income was derived by aggregating all income sources for all individuals in the household.
2. Where an adult respondent refused to be interviewed or was not available, the 'one-shot' income question was used as a measure of household income.
3. For households with partial unit non-response and where one-shot income was missing, any income data collected from responding household resident members were used.
4. NIDS added imputed rental income from owner-occupied housing to all households, irrespective of the method of aggregation, where appropriate.

Where respondents were unable to declare their 'one shot' income, bracket responses were given, the mid-point of the bracket was assigned if a respondent's income fell withing the income bracket, and individuals with income above the value of the highest bracket were assigned twice the value of the upper bound of the top bracket, in line with the Pareto Index method of estimating income for the upper tail (Brophy et al. 2018).

There are variations in the way NIDS captured income from subsistence agriculture. In 2008, NIDS collected income from subsistence agriculture using the household questionnaire; however, from 2010, income from this category was collected using the adult questionnaire and an additional new question was added in 2010. Brophy et al. (2018: 48) report that NIDS did not use agricultural income in calculating aggregate household income in 2010. Further changes on income from subsistence agriculture were made in 2012: apart from asking the actual rand value that accrued to the household from the sale of agricultural produce and livestock, the income question on subsistence agriculture was moved back to the household questionnaire (Brophy et al. 2018: 54).

Differences also exist in the way income from bonus payments was captured across NIDS waves. In 2008, the questionnaire asked an annual question on 13th cheques,[5] profit shares, and bonus payments in the previous 12 months, and the monthly average derived from this annual figure was used. In 2010, the question captured the same sources of income only for a period of one month. In 2012, 2015, and 2017, NIDS used both a monthly and annual question. Thus, with few exceptions, income was consistently measured and imputation for missing values was also consistent. Table A5 shows the number of households with missing and zero income and expenditure values.

Table A5: Households with zero income and expenditure values

| | Zero values | | | Missing values | |
year	Income	Expenditure	Income	Expenditure	No. of households
1993	34	-	178	65	8,809
2008	17	-	-	-	7,058
2010	-	-	-	-	6,580
2012	-	-	-	-	7,824
2015	-	-	-	-	9,615
2017	-	-	-	-	10,842
Total	51	-	178	65	50,728

Source: authors' construction based on survey data.

The household questionnaire in NIDS asks about food and non-food spending in the previous 30 days. Households reported their expenditure on various food items, which was aggregated into a household food expenditure value. To calculate total household expenditure, rental expenditure and imputed income from owner-occupied dwellings were added to total food and non-food household expenditure. Imputed rental income from owner-occupied housing was added to both income and expenditure (Brophy et al. 2018).

There are also some comparability issues between the 1993 PSLSD survey and the NIDS datasets. The questions used to measure income in the 1993 PSLSD survey and the 2008 NIDS differ only slightly (see Leibbrandt et al. 2010). Individual-level income was obtained from all adult household members in the 2008 and later NIDS waves, whereas in the 1993 survey, one respondent per household answered the income and other questions in the survey. Furthermore, there were changes in how agricultural income and imputed rent were captured in the 1993 PSLSD and NIDS datasets. As a result, research that used these datasets to examine changes in inequality exclude farm income and imputed rent values over all survey years (Hundenborn et al. 2018). In addition, total expenditure in PSLSD was calculated as current expenditure net of taxes, excluding investment in durables and all forms of income tax liabilities (PSLSD 1994).

---

[5] A bonus paid to employees on top of their salary, usually in December.

Comparability of other surveys

We briefly discuss the comparability of other datasets utilized in South Africa to analyse income poverty and inequality. The OHS and GHS, among other data sources, were utilized to estimate poverty levels in South Africa. In the WIID data only the 1997 OHS was used, while the WIID Companion used neither. Both types of survey collected information on income and expenditure. Because the OHS and GHS were gathered annually, these datasets are more frequent than the IES or LCS. In both the OHS and the GHS, respondents were asked to provide point estimates for income; however, expenditure data were collected in bands. When respondents were unable to provided point estimates, income bands were used to collect income.

South African population censuses and CS have also been used to analyse income poverty and inequality in South Africa. However, these datasets collect data on income but not on expenditure/consumption. Only the 1996 and 2001 censuses were used in the WIID Companion for income inequality estimation. One of the key limitations of the income data in the censuses is that income was collected only using income bands, not actual income values, and income brackets are not consistent over time. While the 1996 census used 14 income bands (including zero income) to collect monthly personal income and the upper limit was set at ZAR30,001 or more, the 2001 census used 12 bands, and the upper limit was set at ZAR204,801 or more. Thus, the derived household incomes have a different number of bands and band width (see Table A6). Both annual and monthly income categories were used to collect individual-level income in the 2001 census. In 2011, both personal and household annual incomes were collected using the same number of bands and band width as those in the 2001 census. Individual-level income was used to obtain household total income, which can be used to calculate per capita income for the purpose of welfare analysis. Thus, the derived household total income variable is an aggregation of all individual incomes of household members and other additional income, including remittances (Stats SA 1996, 2001, 2011b). Table A6 shows the income brackets for derived household income available in the publicly released data in the three censuses.

The population censuses asked for personal income in different ways over the years. The 1996 income question seems more comprehensive than the questions in 2001 and 2011. Two follow-up questions were also added in 1996 that asked about any 'additional money' in the household as well as remittances or alimony. These were removed in 2001 and 2011. The 2007 and 2016 CS followed a similar approach to that used in the 2001 and 2011 censuses.

Another issue with the use of census data is that large number of households report zero income (Ardington et al. 2006; Yu 2016). For example, Yu (2016) shows that 13 per cent, 21 per cent, and 15.1 per cent of households reported zero income in the 1996, 2001, and 2011 censuses, respectively. Stat SA used different imputation methods for zero and missing income values across the censuses. In both 1996 and 2001, if households reported zero income, no adjustment was made. In 1996, in deriving household income from individual-level income, if income was missing for at least one household member aged 15 and above, derived household income was set to 'missing'.

Table A6: Derived household annual income categories

1996	2001	2011
Annual household income	Annual household income	Annual household income
None	No income	No income
ZAR1–2,400	ZAR1–4,800	ZAR1–4,800
ZAR2,401–6,000	ZAR4,801–9,600	ZAR4,801–9,600
ZAR6,001–12,000	ZAR9,601–19,200	ZAR9,601–19,200
ZAR12,001–18,000	ZAR19,201–38,400	ZAR19,201–38,400
ZAR18,001–30,000	ZAR38,401–76,800	ZAR38,401–76,800
ZAR30,001–42,000	ZAR76,801–153,600	ZAR76,801–153,600
ZAR42,001–54,000	ZAR153,601–307,200	ZAR153,601–307,200
ZAR54,001–72,000	ZAR307,201–614,400	ZAR307,201–614,400
ZAR72,001–96,000	ZAR614,401–1,228,800	ZAR614,401–1,228,800
ZAR96,001–132,000	ZAR1,228,801–2,457,600	ZAR1,228,801–2,457,600
ZAR132,001–192,000	ZAR2,457,601+	ZAR2,457,601+
ZAR192,001–360,000	Not applicable (institutions)	Unspecified
ZAR360,001+		
Unspecified		

Sources: authors' construction based on Stats SA (1996, 2001, 2011b).

Appendix B

Figure B1: Kernel density of per capita income and expenditure, IES (1995–2010) and LCS (2009 and 2015)
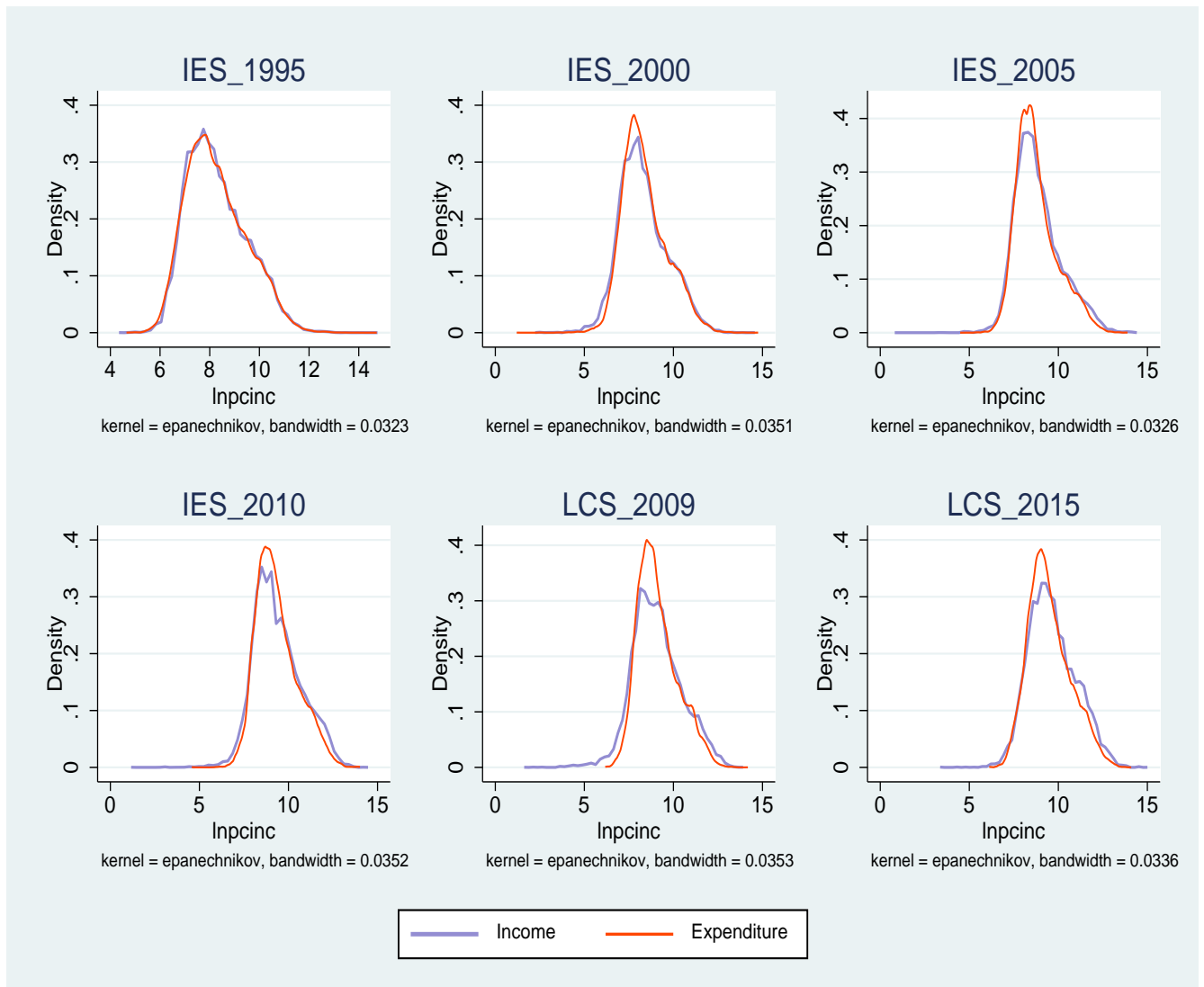
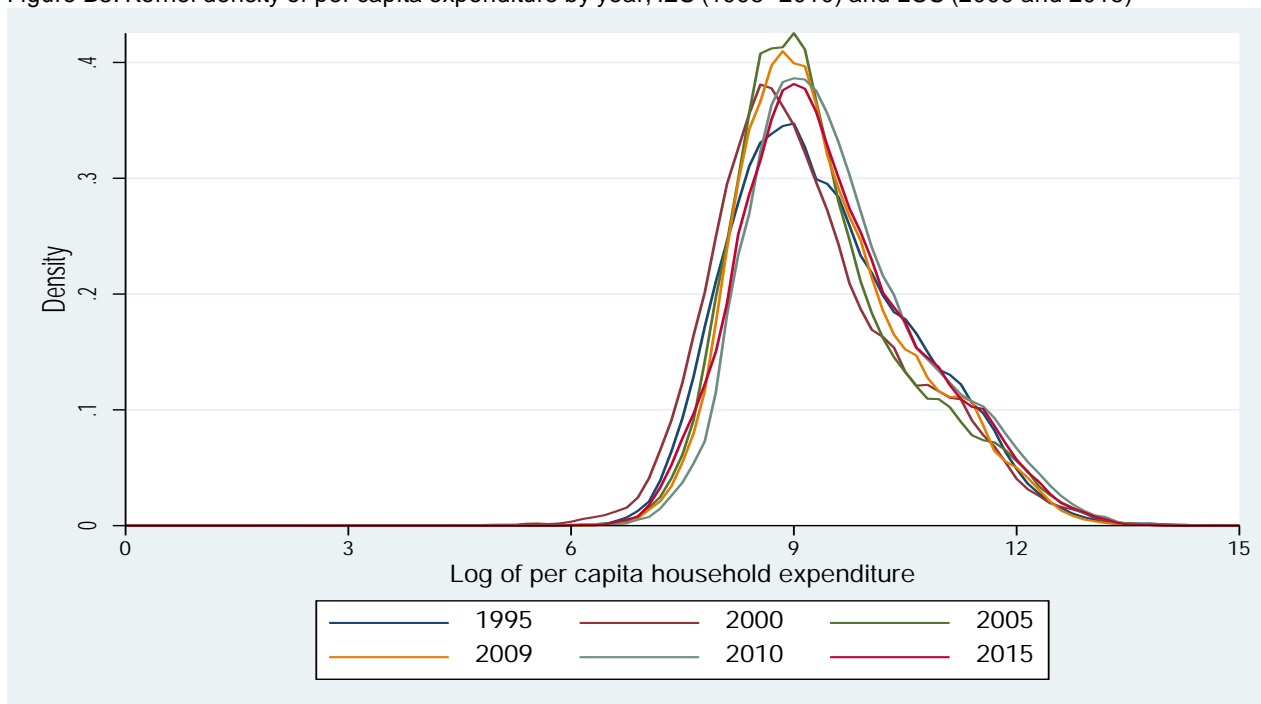Source: authors' illustration based on survey data.

Figure B2: Kernel density of per capita income by year, IES (1995–2010) and LCS (2009 and 2015)
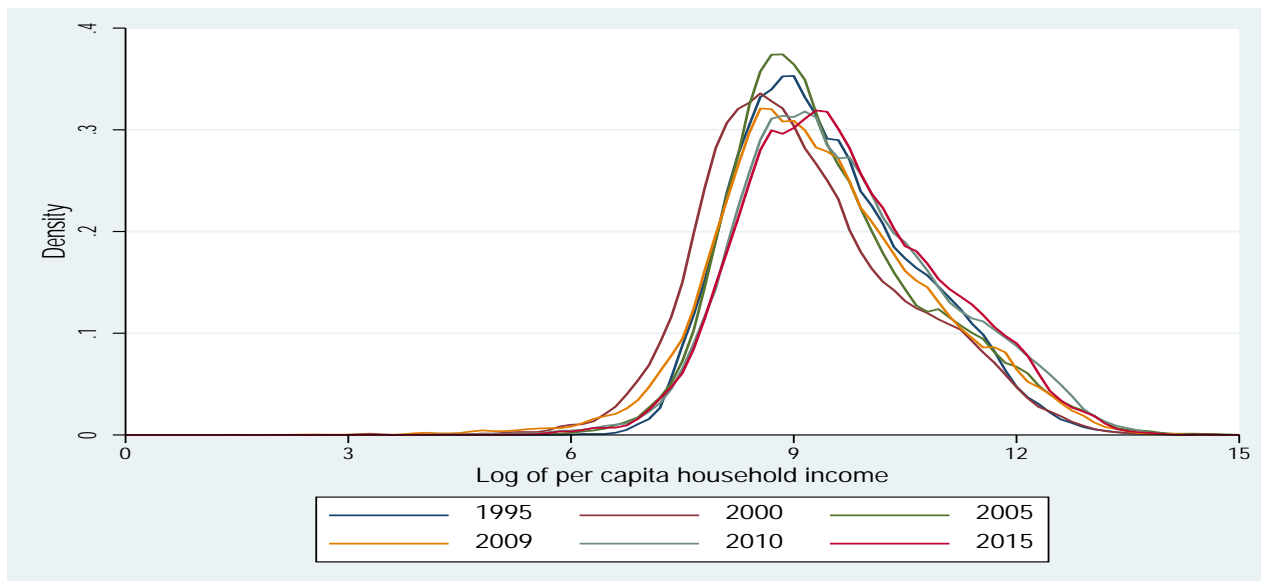
Note: real values are calculated using the 2015 prices.

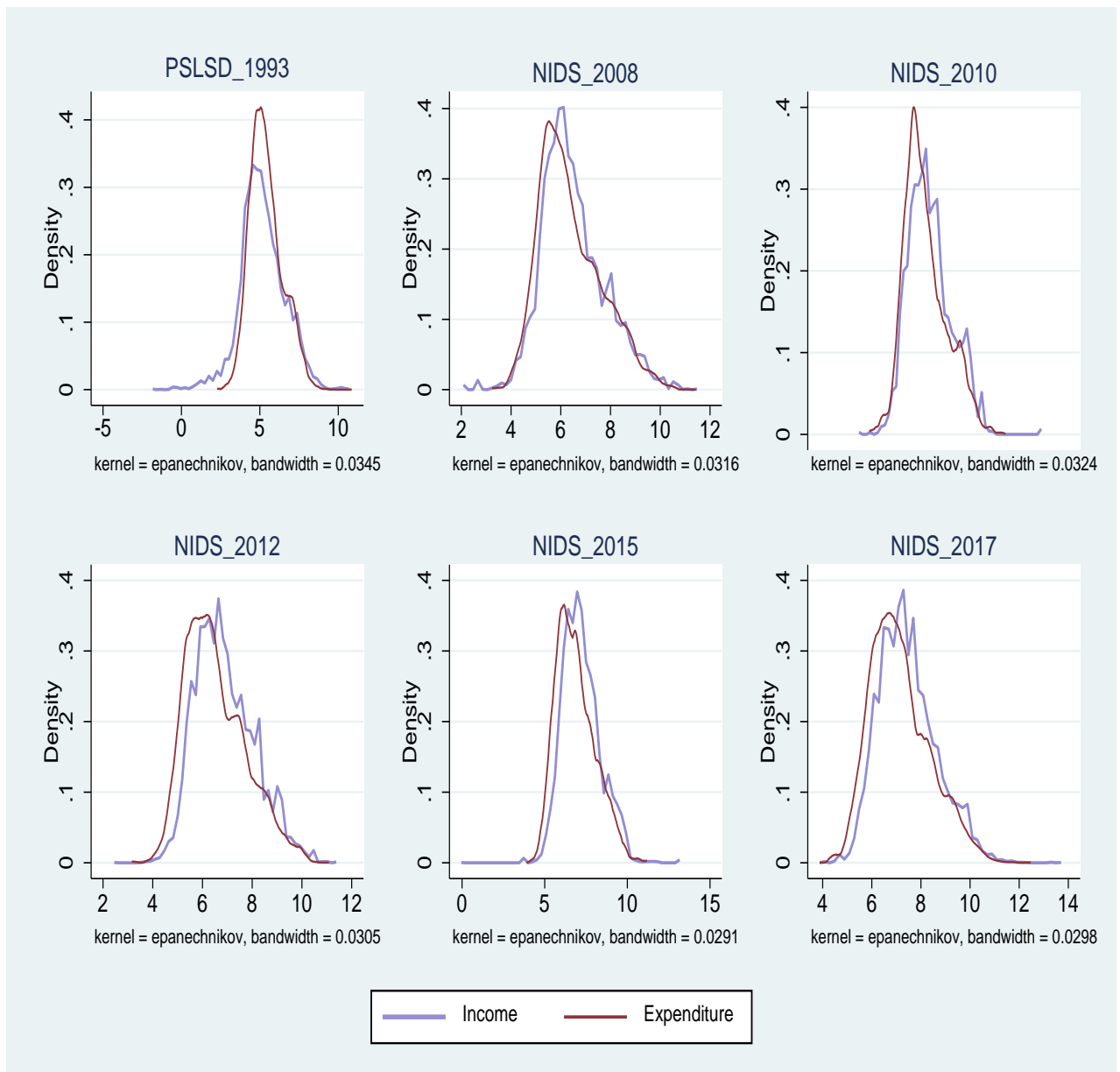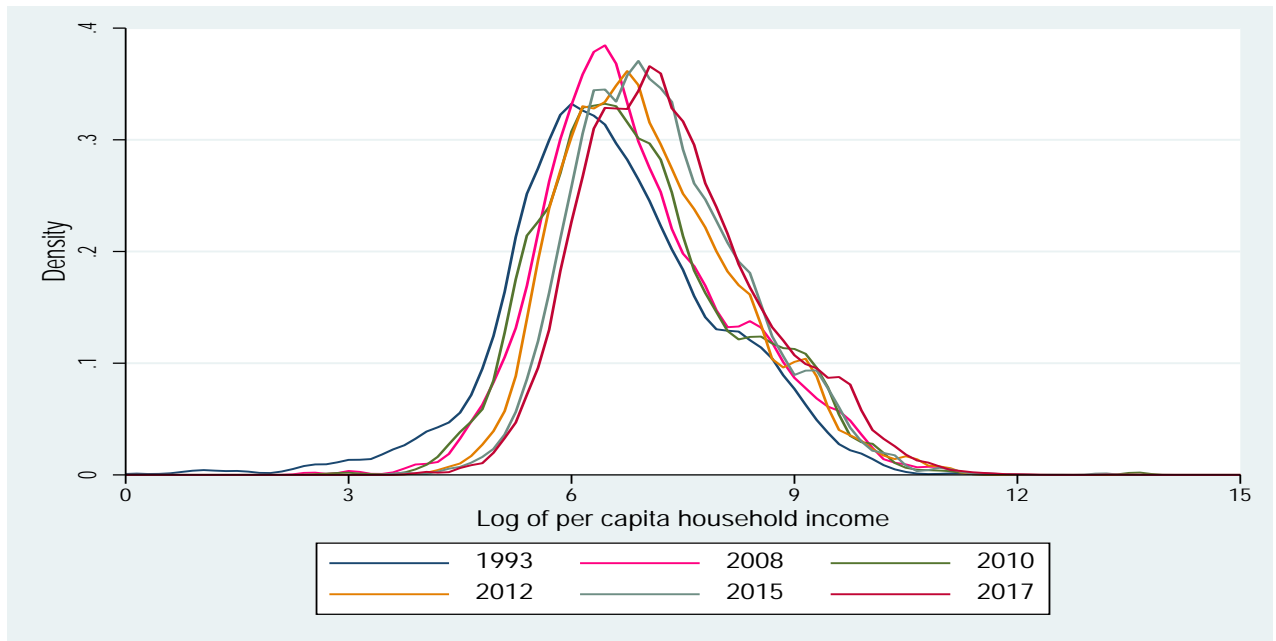Source: authors' illustration based on survey data.

Figure B3: Kernel density of per capita expenditure by year, IES (1995–2010) and LCS (2009 and 2015)

Note: real values are calculated using the 2015 prices.

Source: authors' illustration based on survey data.

Figure B4: Kernel density of per capita income and expenditure, PSLSD (1993) and NIDS (2008–17)

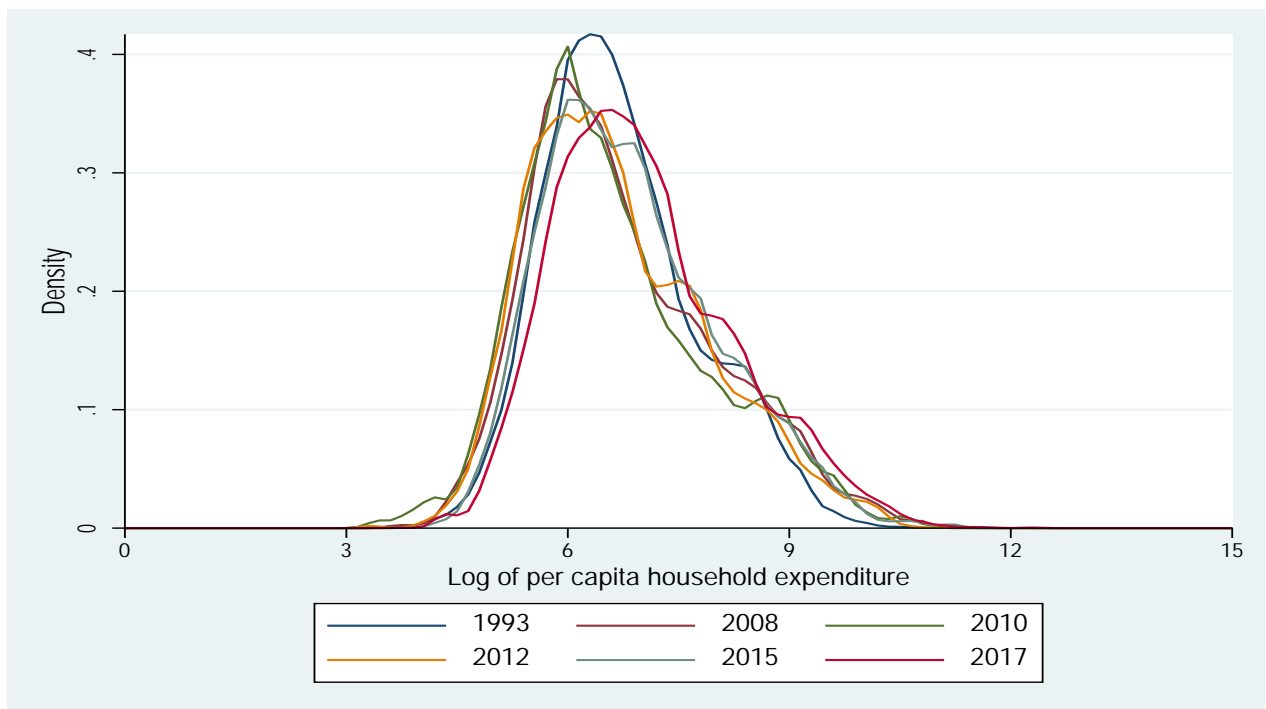Source: authors' illustration based on survey data.

Figure B5: Kernel density of per capita income by year, PSLSD (1993) and NIDS (2008–17)

Note: real values are calculated using the 2015 prices.

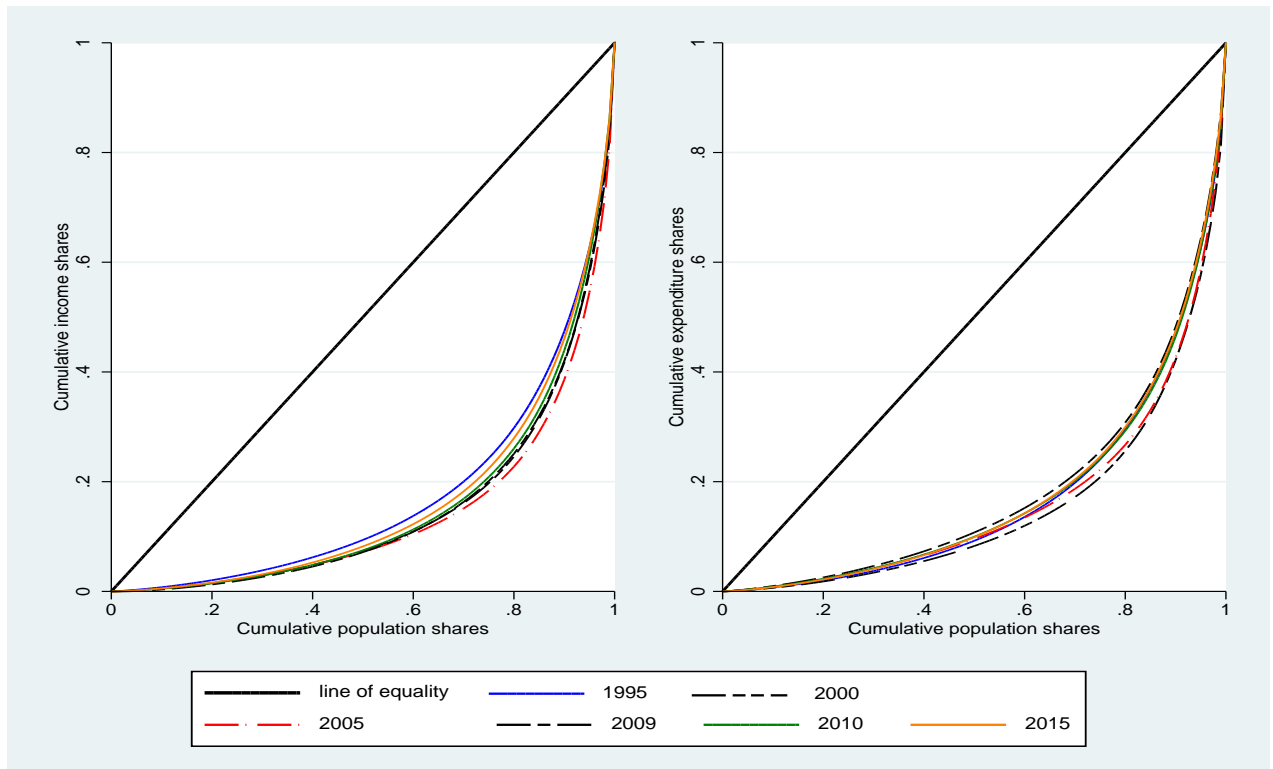Source: authors' illustration based on survey data.

Figure B6: Kernel density of per capita expenditure by year, PSLSD (1993) and NIDS (2008–17)

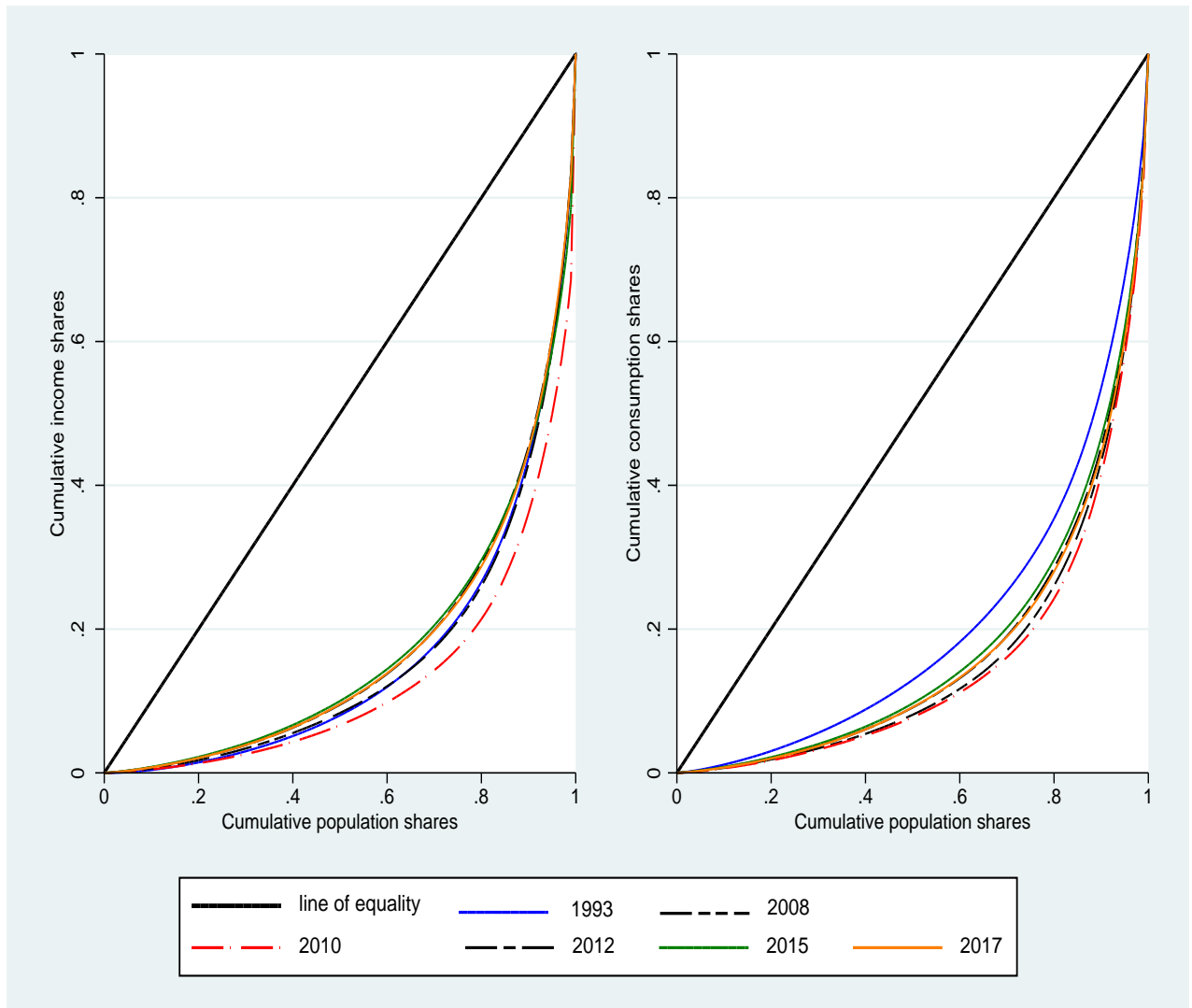Note: real values are calculated using the 2015 prices.

Source: authors' illustration based on survey data.

Figure B7: Lorenz curves of per capita income and expenditure, IES (1995–2010) and LCS (2009 and 2015)

Source: authors' illustration based on survey data.

Figure B8: Lorenz curves of per capita income and expenditure, PSLSDS (1993) and NIDS (2008–17)

Source: authors' illustration based on survey data.