LIS
Working Page Series

No. 903

Correcting for the Missing Rich: Imputing Missing Incomes into Household Data Proportional to Wealth Distribution

Halit Güzelsoy, Hasan Tekgüç

September 2025

CROSS-NATIONAL DATA CENTER
in Luxembourg

Luxembourg Income Study (LIS), asbl

Correcting for the Missing Rich: Imputing Missing Incomes into Household Data Proportional to Wealth Distribution

Halit Güzelsoy[1] & Hasan Tekgüç[2]

31 August 2025 Version

Abstract

We compare alternative imputation methods to correct for missing incomes of the rich in inequality estimates for countries where official tax statistics are not available or reliable. First, we demonstrate that correcting household survey data by imputing missing incomes proportional to wealth distribution is both (i) economically intuitive and macroeconomically consistent, and (ii) yields results close to benchmark studies for Turkey for 2019. We follow Alvaredo et al. (2019) and use the average of the US, China, and France's wealth distribution (from WID) for countries where wealth distribution data is missing. Second, we validate our method with summary statistics data for Brazil and Chile from LIS. We compare our estimates for corrected Gini for Brazil and Chile with De Rosa et al. (2024), which employs official tax statistics. Our estimates approximate De Rosa et al. (2024) findings both in magnitude of correction and in trends over time. Third, we estimate corrected Gini coefficients for developing countries where no reliable tax data is available and show that official Gini estimates under-estimate inequality by more than 10 percentage points in most places and times. To conclude, we propose a method to correct inequality estimates for missing incomes of the rich that requires only publicly available summary data from LIS (for income distribution), WID (for wealth distribution), and World Bank WDI for GDP and its expenditure components, and no specialized software.

Keywords: Income distribution, wealth distribution, macroeconomic consistency.

JEL Codes: D31, E01.

Acknowledgments: This research is funded by TUBITAK Grant No 122G241. We thank Ayşe Göç, Emre Almaz, and Çağlar Karabulut for their assistance and helpful comments. All the remaining errors are ours.

[1] University of Missouri – Kansas City Economics Department PhD student.
[2] Kadir Has University Department of Economics, Istanbul. Corresponding author: hasan.tekguc@khas.edu.tr

Introduction

Income inequality estimates mostly rely on household surveys. Estimates from these surveys are downward-biased because of the 'missing rich problem' (Lustig, 2020). The main reason causing this problem is that wealthy households are less willing to participate in surveys. This is referred to as *unit non-response*. Additionally, another significant issue is that rich households may not accurately report their income, and the surveys are not detailed enough to capture various types of income (*item non-response)*. Three alternative methods to correct downward bias in income inequality estimates are: (i) complement survey data with tax data for the upper tail, (ii) collect specialized data over-sampling the rich, and (iii) use existing data to impute incomes to the upper tail of the income distribution.

In the last two decades, researchers have used official tax statistics to correct for downward bias in inequality estimates. However, official tax statistics are not available for researchers in most developing countries outside of Latin America. Moreover, official tax statistics are probably not useful in correcting inequality estimates in countries where tax evasion is rampant (van der Weide et al., 2018). Furthermore, in certain countries, a significant portion of national income, often exceeding half, is generated in the informal sector and remains untaxed. Similarly, certain components of capital income (such as corporate retained earnings or imputed rental income) could be missing from fiscal income data, even when there is no tax evasion (Alvaredo et al., 2016, and Piketty et al., 2017). Therefore, the tax data in developing countries also might be biased in representing the income and wealth distribution (Ravallion, 2022). Additionally, not all sources of income are taxed, so even in cases of perfect compliance, income from untaxed sources will not be reported in tax data. Zucman (2023) notes that the current tax system was created in the 1950s to reflect the postwar socio-economic state of countries. However, today's world is much different than the postwar conditions, as private wealth has increased rapidly and the share of capital income in GDP is now much higher. Despite its potential shortcomings and biases, tax data has demonstrated its ability to correct the underestimation of income distribution in developed economies (Yonzan et al., 2022). De Rosa et al. (2024) use tax data to correct the downward bias in income distribution for various Latin American countries where data is available. Ledic et al. (2024) apply a similar correction for Croatia using an improved version of such imputation.

The second alternative for addressing the missing rich problem in household surveys can be relying on specialized data sources. Methods requiring specialized data sources are often costly and more complicated. Additionally, they cannot be perpetuated back in time, hence they are not useful in analyzing change over time. The Household Finance and Consumption Survey (HFCS) is one of the specialized sources that is specifically designed for this purpose. The Central Bank of the Republic of Turkey conducted an HFCS to explore households' asset and liability structure. To address the missing rich problem, they oversample the rich neighborhoods using house prices as a criterion (for the methodology of identification of wealthy households, see Ceritoğlu and Sevinç (2020); also see the construction of the survey: Betti et al., (2022)). Based on this survey, Ceritoğlu et al. (2023) estimate the income and wealth inequality statistics for Turkey.

Alternatively, Van der Weide et al. (2018) use house prices to correct the missing rich problem in Egypt. They find more than a 10 Gini point increase after correction. Karabulut (2024) applies van der Weide et al. (2018) method to Turkey. Using more than 200,000 house advertisements, he constructs a distribution intended to correct especially for missing capital and entrepreneurial income. His findings are quite compatible with the results of Ceritoğlu et al. (2023).

The third alternative is to impute the discrepancy between National Accounts and household surveys, *missing incomes*, to household data based on a-priori assumptions or secondary data. These methods that use already existing data allow researchers to observe the change over time. Scholars, using survey data, developed methods to improve the precision of the surveys to approximate the actual distribution. Korinek et al. (2005) and Hlasny (2021) use adjusted sample weights to better estimate top incomes. They construct coefficients from the weights, which help them estimate the non-response rate. Alternatively, Lakner and Milanovic (2013) propose adding the discrepancy between incomes from surveys and national accounts to the income of the top 10%. Alvaredo et al. (2019) use the Lebanese tax data and the average of the United States, France, and China wealth distribution to correct income distribution for all Middle Eastern countries. In these studies, results are sensitive to assumptions made during the imputation stage. The discrepancy between National Accounts and surveys for each income type is not uniform. For example, the missing income in surveys for labor income is much less than capital income. To implement a detailed comparison, the detailed breakdown of aggregate income must be available both in surveys and national Accounts. Once the appropriate comparison is made, the remaining question is determining the method of imputation (Zwijnenburg, 2022).

In this paper, we follow the third alternative and impute the discrepancy between surveys and National Accounts. First, we compare alternative imputation methods for data from Turkey. Turkey is a very suitable case for comparing alternative imputation methods because (i) Ceritoğlu et al. (2023) and Karabulut (2024) provide alternative income inequality estimates where they also grapple with the missing rich problem. Ceritoğlu et al. (2023) document the findings of a first-of-its-kind survey for Turkey, where they deliberately over-sample neighborhoods with expensive homes in order to better capture higher net worth households.[3] Karabulut (2024) follows van der Weide et al. (2018) and corrects downward bias in inequality estimates using house price data. (ii) Tekgüç and Eryar (2025) follow Lustig (2018) and construct gross incomes from survey data. The discrepancies between gross incomes from survey data and National Accounts estimated by Tekgüç and Eryar (2025) allow us to experiment with alternative imputation scenarios. We consider alternative methods proposed by Assouad (2023), Alvaredo et al. (2019), Hlasny (2021), and Lakner & Milanovic (2013), as well as hybrid versions of these methods to estimate alternative corrections. In our preferred alternative, Scenario 6, we distribute the discrepancy between surveys and National Accounts with respect to the average of US, China, and France wealth distribution (see Table 1 for comparisons of estimates from alternative scenarios). This method incorporates the least number of assumptions compared to the alternatives and corroborates Ceritoğlu et al. (2023) and Karabulut (2024) findings. Scenario 6 also does not require microdata and can be performed with publicly available summary statistics such as average income and wealth shares by deciles/ventiles.

Second, we apply our preferred method to various selected countries. We started with Turkey and extended the data backward and forward in time. We show that our preferred method performs well in capturing the rapid worsening of income inequality in 2021 and 2022 in the high-inflation environment (see Figures 1 and 2). We then check the external validity of our preferred method with LIS data for Brazil and Chile. We obtain income and wealth shares of the ventiles from LIS and the World Inequality Database (WID), respectively. We chose Brazil and Chile for external

[3] This survey is sponsored by Central Bank of Turkey and is compatible with European Central Bank – Household Finance and Consumption Survey (Betti et al. 2022).

validity because De Rosa et al. (2024) correct Gini coefficients for selected Latin American countries using administrative tax data. We show that our preferred method, Scenario 6, performs well for Brazil and Chile (Figure 3) in the sense that both the amount of correction (in terms of percentage points) and trends over time approximate De Rosa et al. (2024) findings. Finally, we employ LIS income data and WID wealth data and generate corrected Gini estimates for selected countries (Table 2).

The rest of the paper is organized as follows. Section 2 reviews alternative imputation methods under *alternative scenarios;* application for Turkey for an 18-year period; and criteria for international comparisons. Section 3 presents findings. Section 4 discusses the findings and concludes the paper.

2. Methods

2.1 Alternative Imputation Scenarios

In this section, we employ the Household Budget Survey (HBS) for the years 2003, 2007, 2011, 2015, and 2019. We chose HBS over Survey of Income and Living Conditions (SILC) because (i) Alvaredo et al. (2019) also use HBS, and Tekgüç and Eryar (2025) employ HBS to construct pretax incomes for different types of incomes. Tekgüç and Eryar (2025: Table A1) also provide comparisons of different kinds of incomes in HBS versus the National Accounts, which is necessary for scenarios 1, 2, 6, and 7.[4]

Lower Bound: Hlasny (2021) proposes a method that addresses only the issue of unit non-response bias found in surveys. Hlasny (2021) obtains non-response rates for different income groups (he finds that non-response increases with income) and adjusts survey weights taking into account the non-response. We adapt the coefficients the author obtained for Mexico to Turkey. We refer to these as lower-bound estimates because only one aspect of measurement issues is addressed.

[4] We reproduced this Table as Table A.1 in the Appendix. We use the averages/trends of 2011-2019 for years 2003 and 2007 where detailed National Accounts are not available.

Upper Bound: Lakner & Milanovic (2013) propose adding the discrepancy between surveys and National Accounts directly to the income of the top 10%. We create an extreme version of Lakner & Milanovic (2013) to serve as an upper bound for other scenarios: we allocate 2/3 of the discrepancy to the income of the top 5% and the remaining 1/3 to the second-highest 5% income group. As De Rosa et al. (2024) also highlight, it is arbitrary to attribute the whole discrepancy solely to the top 10%.

Benchmark: Ceritoğlu et al. (2023) and Karabulut (2024) estimate the income inequality Gini coefficient for 2019 as 0.517 and 0.501, respectively.

Scenario 1: We compare the total labor income, transfers, and imputed rent incomes in surveys with their counterparts in the National Accounts for Turkey. Each year, there is approximately a 10% discrepancy for these categories with their counterparts in the National Accounts (Tekgüç and Eryar, 2025). Therefore, the relevant income types for the population comprising the lowest 90% of the income distribution have been adjusted upward by 10%. Next, the capital income discrepancy (the total of capital and mixed income) from National Accounts is added to the income of the top 10%, while the income of the bottom 90% is increased by 10%. After this adjustment, the matching rate between the survey and the total disposable income of households in the National Accounts rises to 98%. This adjustment is expected to reduce income inequality with respect to upper-bound estimates.

Scenario 2: In this scenario, we use the distribution of entrepreneurial income in the survey as a benchmark for mixed income. Specifically, while capital income is added to the income of the top 10% income group, mixed income is distributed according to the entrepreneurial income distribution in the survey. Entrepreneurial income refers to the income earned by individuals working as entrepreneurs within a year, including self-employed. In the Household Budget Surveys, the share of the top 10% income group in entrepreneurial income is around 35%[5], except for 2019, in which it is 27%. The share of the bottom 50% in the entrepreneurial income is around 25-30%. Therefore, in this scenario, as we impute entrepreneurial income to the households at the bottom of the distribution, the result of the correction is less unequal than in the previous scenario.

[5] Since this is an underestimate of actual entrepreneurial income, the households that should be in the top of the distribution appear more towards the middle.

Scenario 3: Assouad (2023) applies various steps to correct the income inequality estimates for Lebanon. The most noteworthy adjustment she makes is based on tax data. Using Lebanese tax data, the author employs the "Generalized Pareto Interpolation" method. She assumes that the income of the bottom 80% of the population is well-measured by surveys, while for the top 20%, fiscal data is presumed to be more accurate. The inverted Pareto coefficients are calculated for each percentile from the fiscal data, which are then used to adjust the income data derived from surveys. A similar adjustment for Lebanon is also applied to all Middle Eastern countries in the study Alvaredo et al. (2019). Following the methodologies of Assouad (2023) and Alvaredo et al. (2019), we use Lebanese tax data to calculate adjusted Pareto coefficients for Turkey and re-estimate the income distribution.

Scenario 4: Alvaredo et al. (2019) also use wealth distribution as another correction step. In the study, they estimate that missing capital income for the Lebanese economy is approximately 20% of national income. As a modest assumption, they assume this discrepancy, on average, to be 10% for other Middle Eastern countries. Then, they distribute 10% of the national income according to the wealth distribution. In this scenario, we replicate this step.

In order to apply this step, we first needed the wealth data. The World Inequality Database (WID) provides wealth estimations for various countries and years. The data for the US, China, and France, among these estimations, are reported to be the most reliable data (Alvaredo et al., 2019). Following Alvaredo et al. (2019), we use the arithmetic average across the three countries for each ventile. Then, we distribute 10% of the national income according to wealth distribution.

Scenario 5: Applying both steps together, Alvaredo et al. (2019) and Assouad (2023) re-estimate the income inequality in all Middle Eastern countries, especially focusing on the top 1%.[6] Following them, we combine scenarios 3 and 4 in this scenario. That is to say, we use wealth data to redistribute 10% of national income and inverted Pareto coefficients to correct for the top 20% of income groups.

Even though wealth is much more unequally distributed, the result of applying this step is more equal than in previous scenarios. The reason is that the actual missing capital income, which is

[6] At the third step they also use billionaires list to correct the very top of the distribution.

correlated with wealth distribution, is much greater than 10% of the national income. Based on our calculations, the missing income is 27.9% of national income in 2011, 24.8% in 2015, and 24.4% in 2019 for Turkey. Therefore, for a more accurate measure of inequality, the total discrepancy between national accounts and surveys must be considered.

Scenario 6 – Preferred Method

In addition to employing Lebanese tax data for each Middle Eastern country, Alvaredo et al. (2019) distribute an additional 10 % of GDP National Income to households using the average of wealth distribution of China, France, and the USA (obtained from the WID database). Wealth distribution is considered to be a good proxy for the distribution of capital income.[7] However, 10% of GDP or National Income for each year is an arbitrary figure, and the discrepancy we estimate for Turkey is much higher than 10 % (see also Figure 1 for the discrepancy between survey totals and National Accounts). Therefore, instead of an arbitrary constant percentage, we distribute the discrepancy for disposable household income between National Accounts and surveys (similar to Lakner and Milanovic, 2013) according to the wealth distribution for each year.[8]

Scenario 7

The adjustment of income distribution calculations can be assessed across a broad spectrum of criteria. We have developed scenarios in which we allocate the total discrepancy according to specific rules. Additionally, we can separate this discrepancy into mixed income and capital income. It is reasonable to distribute capital income according to wealth distribution. However, the rule by which mixed income should be distributed remains open to discussion. In this scenario, we distribute the capital income discrepancy based on wealth distribution, while mixed income is allocated according to the distribution of entrepreneurial income.

Wealth is distributed very unequally, favoring high-income groups. Allocating capital income in line with this distribution increases inequality measures. On the other hand, in Turkey, the average income of self-employed individuals among entrepreneurs in surveys is relatively low. Therefore,

[7] For a comparison between the wealth distribution obtained from the average of the three countries and the wealth distribution reported by Ceritoğlu et al. (2023) for Turkey, see Appendix A, Table A.2.

[8] Detailed National Accounts data containing household disposable income totals are available for Turkey starting in 2009.

distributing mixed income according to the entrepreneurial income distribution improves the income inequality measurement. As a result, in this scenario, the income distribution estimate remains closer to the initial point compared to other scenarios. The issue mentioned in Scenario 2 regarding the 2019 household survey is also present here: between 2015 and 2019, there is an unexpectedly rapid improvement in income distribution that cannot be fully explained by economic conditions. This underscores the importance of accurately estimating the distribution of mixed income.

2.2 Application for Turkey: Employing only publicly available summary data

We extend Scenario 6 to forward and backward in time using publicly available data. TurkStat implements the Survey of Income and Living Conditions (SILC) to estimate official inequality and poverty statistics. TurkStat publicly releases summary statistics for SILC roughly one year after the survey's completion. These summary statistics include Gini coefficients, total and equalized per capita household incomes by deciles and ventiles. Microdata from SILC and other surveys are available for outside researchers with further lags. Hence, employing publicly available summary statistics allows us to generate up-to-date corrected Gini estimates very rapidly. As we see in Section 3.2, the very high inflation environment of 2021 to present worsened income inequality in Turkey to unprecedented levels. Generating corrected inequality estimates in a timely manner has obvious value for the public in such cases.

We also use the final consumption expenditure of resident households, C, as a proxy for household disposable income from National Accounts because total disposable household incomes are not available for years before 2009 for Turkey, whereas GDP and its expenditure components are available for years before 2009. Additionally, GDP and its expenditure components are publicly released before the detailed National Accounts tables. Proxying total disposable incomes with the final consumption of households will underestimate income inequality because some of the disposable incomes are saved. Furthermore, as we discuss in further detail in Section 2.4, some of the profits are not distributed to households to avoid taxes. Hence, we estimate a second correction with respect to the Net Private Sector instead of total incomes. First, as the saying goes, *neither capitalists nor workers can eat depreciation,* so we deduct consumption of fixed capital (CFC) from GDP. This is equivalent to pretax National Income in Alvaredo et al. (2019) or De Rosa et al. (2024). Second, we deduct Government Spending and spending by non-profit organizations from total National Income. For most countries, surveys report disposable incomes, not pre-tax incomes; hence, their corollary in aggregate figures should not include Government Spending. Furthermore, determining and assigning final beneficiaries of government spending requires some microdata with demographic variables like age and school attendance to assign public in-kind health and education to households. Such corrections are not possible with summary data. We name the remaining total as *Net Private Sector,* which is the total of final consumption of households plus net investment (I – CFC) plus net exports (X – M) and change in stocks.

Our preferred income definition is Net Private Sector, which (i) takes into account the strategic allocation of profits to firms or households depending on the tax code, (ii) does not include any implicit assumption about the sources of the taxes, unlike pretax National Income. For example, correcting inequality estimates to obtain pretax National Income inequality by wealth distribution or tax records redistributes government spending back to wealthy households because government spending is part of the discrepancy between survey totals and pretax National Income. We believe that in most developing countries assigning government spending to the wealthy in pretax income is unwarranted. In many of these countries, the main source of government revenue is consumption taxes (including tariffs) or revenues from natural resource extraction, not taxes on income or wealth. Since these countries do not tax income and wealth, they do not have reliable income tax records to use in correcting inequality estimates in the first place.

2.3 External Validity

We apply our preferred method, Scenario 6, to LIS data from Brazil and Chile. We choose Brazil and Chile to test the external validity of Scenario 6 because De Rosa et al. (2024) provide corrected Gini estimates based on tax data and scaling data to national income in order to eliminate discrepancies between survey totals and National Accounts. They perform the following three steps: (i) First, they reweight the survey using tax data. (ii) Second, they scale the income totals from step 1 "…to their matched national accounts aggregates (these are wages, property incomes, mixed income, pensions and imputed rents)" (De Rosa et al., 2024: 2). (iii) Finally, they impute missing incomes (such as corporate retained earnings) to reach the national income.

We follow De Rosa et al. (2024) and estimate both raw survey and National income inequality for Brazil and Chile. However, our preferred income definition for international comparisons is the total of net private sector instead of total pretax national income (GDP minus CFC), as we discuss in the previous section.

2.4 International Comparisons: How to Account for Retained Earnings?

Alvaredo et al. (2019) point out that a significant portion of the wealthiest individuals' income is retained within the firms, and household survey-based estimations cannot account for this income. Behringer et al. (2020) go one step further and point out that the share of retained earnings in GDP is contingent upon national tax codes. In some countries, such as Germany, a much larger share of national income is retained as profits within firms to avoid highly progressive income taxes. Hence, inequality estimates based on private household incomes (even after being corrected with official tax statistics) will be biased downwards to a different degree for different countries and time periods. A natural implication of these arguments is that both international and long-run comparisons should take into account differences in retained profits. In other words, in international comparisons (or studies encompassing long periods with significant changes in income tax rates), countries will have different disposable household income inequality levels and trends even if their Net Private Sector income distributions are identical.

In Section 3.4, we provide corrections for Gini estimates, taking into account the discrepancy between disposable incomes in surveys and the Net Private Sector. Net Private Sector total includes retained profits within firms and deducts consumption of fixed capital. An additional benefit of this approach is that it accounts for different levels of industrialization. More industrialized countries, such as China, have much higher depreciation rates compared to other countries; hence, correcting for Net Private Sector (which takes into account differential rates of depreciation) makes a significant difference in such cases.

3. Findings

3.1 Comparing Scenarios

At first glance, Scenario 3 results appear closest to the benchmark in 2019. However, correcting Turkish survey data with inverted Pareto coefficients obtained from Lebanese fiscal data cause internal and macroeconomic consistency problems: (i) ranking of percentiles can change (i.e. 85th percentile move above the 87th percentile[9]) after the correction and (ii) imputed total amounts as a share of GDP increases from 2003 to 2019 despite the fact that surveys actually got better at capturing a larger share of income as time passed (see Table A.3). Furthermore, Scenario 3 is a computationally intensive method and requires both tax statistics and household income micro data. Alternatively, Scenario 6 is closer to the mid-point of the upper and lower bounds when all years are considered. Additionally, Scenario 6 can be applied with publicly available summary statistics: (i) household incomes at each decile/ventile, (ii) total disposable income (or household consumption as a proxy), (iii) average wealth distribution of China, France, and the USA (obtained from the WID database).

[Table 1: Comparing Scenarios]

3.2 Application for Turkish Data for a longer period

Official income inequality statistics in Turkey are based on SILC data. In this section, we switch from HBS to SILC summary data, which is publicly available for deciles and ventiles to produce results comparable to official statistics from TURKSTAT.[10] TURKSTAT provides Gini coefficients both for total disposable income and per capita disposable income adjusted for household size. We focus on total disposable income inequality because (i) total survey income in Figure 1 can only be calculated from total disposable income from publicly available summary statistics, and (ii) the benchmark study Ceritoglu et al. (2023) only reports total disposable income inequality. Figure 1 shows the ratio of total disposable income from SILC surveys to final household consumption (green bars) and the alternative ratio of total disposable income from SILC to Net Private Sector (grey bars). We call these ratios *survey coverage rates*. 100% minus survey coverage rates are our estimates for missing income shares. Figure 1 shows that survey coverage rates steadily increased from 2011 to 2020 and then suddenly dropped in 2021 and 2022 (and partially recovered in 2023). Turkey experienced very high inflation in 2021 and 2022. In this high inflationary environment, people on fixed incomes, such as wages and pensions, experienced erosion of real incomes, whereas people and firms owning assets experienced sudden enrichment. As Lustig (2020) would have predicted, the survey coverage ratio declined precipitously in those years. Surveys have a

[9] See Appendix A Table A.4 for the distortion in the ranking of household.
[10] WID wealth distribution data is available for each percentile, so if the official summary statistics were available at percentiles, Scenario 6 could be applied with percentiles to obtain the income share of top 1%.

hard time measuring the incomes of the richest households, and if these households experience sudden enrichment, survey coverage should decline.

[Figure 1: Survey Coverage Rate]

Figure 2 shows the Gini coefficients estimated with SILC data and the corresponding adjustments in Scenario 6, as well as results from the previous section employing the Household Budget Survey for selected years and the benchmark studies.[11] Corrected Gini estimates from HBS and SILC surveys are very close for 2011, 2015, and 2019 (green dots versus red line). As Figure 2 shows, Scenario 6 also performs in accordance with the expected direction in the high inflation environment of 2021 and 2022. Moreover, the increases in our corrected inequality estimates are much more than the increase in official inequality estimates. Comparing corrected Gini estimates to official data shows that not only the level but also the trends are different. Official inequality statistics show stable inequality between 2009 and 2013 and then a gradual increase between 2013 and 2017. Our corrected inequality estimates for disposable income (S6, SILC data - red line) show an increase between 2009 and 2011 and a gradual decline between 2011 and 2018. Trend in corrected Gini estimates aligns better with the timing of the Great Recession and decline in wage inequality in Turkey after 2013 (Tamkoç and Torul, 2020; Bakis and Polat, 2023). When we take into account retained earnings in our corrections (S6 Net Private Sector, SILC data - purple line), we do not observe the gradual decline between 2011 and 2018.

[Figure 2: Total Household Income Inequality, Official and Corrected Estimates]

Finally, we estimate S6-National Income (black line), where we impute the discrepancy between SILC survey total and National Income in accordance with Scenario 6. The results of this scenario are comparable with Alvaredo et al. (2019) since they also aim to estimate the inequality of national income. Appendix A Figure A.1 shows that correcting for National Income inequality with Scenario 6 produces estimates very similar to their findings, both in level and trend. These findings show that the inequality estimate correction is more sensitive to the amount of discrepancy (i.e., *how much* to correct) imputed into the household survey than the specific method (i.e., *how* to

[11] In Table 1 Gini coefficient has increased by 14 percentage points in 2019 from 0.351 in raw survey data to 0.491 in Scenario 6 with HBS data. In Figure 2, the magnitude of the correction is less, from 0.402 in raw survey data to 0.500 after correcting with Scenario 6. In other words, Gini estimates from raw survey data differ significantly between HBS and SILC. However, after the correction both surveys yielded similar results.

correct), and their complicated three-step procedure can be simplified and made macroeconomically consistent.

3.3 External Validity: Estimates for Brazil and Chile

We apply the method used in Scenario 6 to the summary survey data obtained from LIS for Brazil and Chile. We obtain the average total income by ventiles from the LIS website.[12] We use the average of China, France, and the USA wealth distribution for Brazil and Chile as well. We obtain the GDP and its expenditure components from the World Bank World Development Indicators. The level of our estimates differs from De Rosa et al. (2024) because they estimate inequality per capita, whereas we estimate total household income inequality to be consistent with previous sections.[13] As we discussed in Section 2.4, our corrected Gini estimates for National Income are comparable to their estimates in Figure 1.

For Brazil, they show that the discrepancy between raw survey Gini and corrected National Income Gini widens over time, from roughly 10 percentage points in 2001 to 14-15 percentage points in 2015 (De Rosa et al. 2024 Figure 1). Moreover, the trends of the raw survey and corrected Gini diverge after 2007. Raw survey Gini declines continuously between 2001 and 2015, whereas after 2007, their corrected Gini estimates fluctuate throughout the period. Similar to their findings, we find that the discrepancy between raw survey and corrected Gini estimates widens over time from roughly seven percentage points in 2001 to roughly 12 percentage points by 2015. Likewise, we also show that trends of raw survey Gini and corrected Gini estimates diverge after the mid-2000s. Our estimate for Chile has similar discrepancies and trajectories with De Rosa et al. (2024). The discrepancy between raw and corrected Gini estimates is more than 10 percentage points in both corrections, and in both cases, there is no declining trend in corrected Gini estimates in the 2000-10 period. Medeiros et al. (2018) use the 2010 Brazilian Census data and tax records to correct for

---

[12] Go to https://www.lisdatacenter.org/data-access/lissy/syntax/ web-site and click on "Login to Lissy". After login, run the following Stata code to obtain average household incomes by ventile for China in 2018:

```
use $cn18h, clear
xtile ventile_cn18h = hitotal, nq(20)
tab ventile_cn18h, summ(hitotal) nofreq nost noobs
ineqdeco hitotal
```

List of LIS data: https://www.lisdatacenter.org/our-data/lis-database/

Inequality estimates with household weights yield very similar results.

[13] De Rosa et al. (2024) estimates per capita income inequality without adjusting for household size whereas we estimate total household inequality. We estimate per capita income inequality with LIS data in Appendix B Figure B.1

the missing-rich problem. Using calibrated incomes above different threshold scenarios, in which they make the correction for the top 2.5% and 5%, they find a 6 to 8 Gini point increase in inequality estimates. An interesting feature of Brazil is that the corrected Gini estimates for Net Private is equal and lower than corrected Gini estimates for disposable income after 2012. In 2009 Consumption of Fixed Capital jumped from 14.8% 17.1% of GDP and continued to increase to 23.1% of GDP by 2016. Investments (and savings) did not increase to compensate the increase in Consumption of Fixed Capital. Hence, net Investment was negative after 2012. Consequently, Net Private Sector share of GDP dropped below Consumption share by 2012 and missing income correction for Net Private Sector was less than disposable income in those years.

[Figure 3: Correcting Gini Coefficients for Brazil and Chile, LIS data]

In Appendix C, we present the figures for the top 5%, top 10%, middle 40% and bottom 50% for the same income definitions in Figure 3 for Turkey, Brazil, and Chile. Our preferred income definition for international comparisons is S6, Net Private Sector. The increase in the share of the top 10% in Turkey is astonishing in 2021-2022 (Figure C.1, top right panel). The share of the top 10 % increased from 45% of national income to 50% in 2022, surpassing Brazil (Figure C.2, top right panel) and approaching Chilean levels. The trend in Chile is in the opposite direction (Figure C.3 top right panel), supporting the claims that the situation in Turkey is not a global phenomenon but the result of very high inflation.

3.4 Correcting for downward bias in selected countries

We apply the method used in Scenario 6 to the summary statistics of survey data obtained from LIS for China, Egypt, India, the Ivory Coast, Mali, and Vietnam. We obtain the average total income by ventiles from the LIS website. We use the average of China, France, and the USA wealth distribution for these countries as well (except China, for China, we use its own wealth distribution from WID). We obtain the GDP and its expenditure components from the World Bank World Development Indicators. In general, the discrepancies between Gini estimates from the raw survey and corrected estimates are more than 10 percentage points. Vietnam is the most curious case. In 2007 and 2009, the net private sector total is less than the total disposable income in surveys.

Hence, corrected Gini coefficients are lower than raw Gini estimates. This curious pattern reversed to a more normal pattern after 2011. In the case of the two most populous countries, our estimates point to worsening income inequality in the 21st Century.

[Table 2: Corrections for other countries]

4. Discussion and Conclusion

Measuring inequality is a major challenge for all countries. Household surveys often face various issues, and these challenges are even more pronounced in developing countries. Correcting income inequality estimates with tax data is the state of the art when tax data is available and reliable. However, in most developing countries, tax data is either unavailable or unreliable or both. Therefore, alternative methods have been developed to address the so-called "missing rich" problem. One such method is conducting a one-time, specialized survey, which can provide a more accurate estimate of inequality. Using house price data to estimate the right tail of the income distribution is another alternative. Nevertheless, these methods are costly in time and money to implement and do not allow for retrospective analysis of past inequality trends.

A more practical alternative is to use existing data, especially National Accounts, to improve income inequality estimates. Lakner and Milanovic (2013) propose to allocate the discrepancy between surveys and household consumption or disposable income in National Accounts to the top 10%. Alternatively, Alvaredo et al. (2019) propose a three-step method for Middle Eastern Countries involving Lebanese tax data, wealth distribution from the USA, France, and China, and rich lists. We show that Lakner and Milanovic's (2013) method ensures that corrected income totals are macroeconomically consistent. However, their allocation strategy is based on a heuristic. Alvaredo et al. (2019) allocation strategy is data-driven, but their three-step procedure does not ensure macroeconomic consistency, and their Generalized Pareto Interpolation undesirably re-orders the top twenty percentiles at least for Turkey, as we show in Section 3.1 and Appendix A. In our preferred Scenario 6, we combine the better parts of these alternatives. The total incomes after the correction should be consistent with the National Accounts, as in Lakner and Milanovic (2013). Allocation of missing income should be data-driven (in 2019 top 10 % share of wealth was 63% so they are allocated 63% of the discrepancy between survey and National Accounts, see

Table A2). As expected, surveys under-estimate capital and entrepreneurial incomes much more than others, like labor or pension incomes. Hence, Alvaredo et al. (2019)'s idea to allocate missing income with respect to wealth distribution is very intuitive. Furthermore, the wealth data adjustment they propose is highly practical and easy to implement, even with summary data, i.e., even when microdata is unavailable. Using this combined approach (Scenario 6) for Turkey, the estimates also align with the results of benchmark studies for 2019.

We apply Scenario 6 to 2005-2023 Turkish data to obtain corrected inequality estimates for household disposable income distribution, Net Private Sector income distribution, and pretax Net Income distribution. We show that after the correction, not only the level but also the trends in inequality are different. Trends of corrected inequality estimates for household disposable income align better with detailed studies on labor earnings.

We also estimate corrected inequality estimates for Chile and Brazil with Scenario 6. De Rosa et al. (2024) is a state-of-the-art paper, which corrects inequality estimates with tax data and allows us to cross-validate the findings of Scenario 6. We use summary statistics for income distribution from LIS, National Accounts data from World Development Indicators, and WID wealth distribution data for Brazil and Chile. We show that the magnitude of correction and trends are broadly similar to De Rosa et al. (2024) findings. Finally, we apply Scenario 6 to other selected developing countries with no tax data. We show that, on average, Gini coefficients are under-estimated by 10-12 points.

References

Alvaredo, F., Atkinson, A., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2016). Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in WID.world. *Working Papers*, Article 201602. https://ideas.repec.org//p/wel/wpaper/201602.html

Alvaredo, F., Assouad, L., & Piketty, T. (2019). Measuring lnequality in the Middle East 1990–2016: The World's Most Unequal Region? *Review of Income and Wealth*, *65*(4), 685–711. https://doi.org/10.1111/roiw.12385

Assouad, L. (2023). Rethinking the Lebanese economic miracle: The extreme concentration of income and wealth in Lebanon, 2005–2014. *Journal of Development Economics*, *161*, 103003. https://doi.org/10.1016/j.jdeveco.2022.103003

Bakis, O., & Polat, S. (2023). Wage inequality dynamics in Turkey. *Empirica*, 50(3), 657-694.

Behringer, J., Kowall, N., Theobald, T., & Van Treeck, T. (2020). Inequality in Germany: A macroeconomic perspective. *German Politics*, *29*(3), 479-497.

Betti, G., Ceritoğlu, E., Küçükbayrak, M., & Sevinç, Ö. (2022). Central Bank of the Republic of Turkey Household Finance and Consumption Survey Methodology. *Working Paper No:22/01*.

Ceritoğlu, E., Cılasun, S. M., Küçükbayrak, M., & Sevinç, Ö. (2023). Household portfolios in Türkiye: Results from the household finance and consumption survey. *Central Bank Review*, *23*(4), 100132. https://doi.org/10.1016/j.cbrev.2023.100132

Ceritoglu, E., & Sevinc, O. (2020). Identification of Wealthy Households from the Residential Property Price Index Database for Sample Selection for Household Surveys. *Working Papers*, Article 2010. https://ideas.repec.org//p/tcb/wpaper/2010.html

De Rosa, M., Flores, I., & Morgan, M. (2024). More unequal or not as rich? Revisiting the Latin American exception. *World Development*, *184*, 106737.

Hlasny, V. (2021). Redistributive effects of fiscal policies in Mexico: Corrections for top income measurement problems. *Latin American Policy*, *12*(1), 148–180. https://doi.org/10.1111/lamp.12206

Karabulut, F. Ç. (2024). Correcting Downward Bias in Inequality Estimates For Turkey with House Price Data. Kadir Has University Economics Department Master Thesis.

Korinek, A., Mistiaen, J. A., & Ravallion, M. (2005). *Survey Nonresponse and the Distribution of Income* (SSRN Scholarly Paper 695442). Social Science Research Network. https://papers.ssrn.com/abstract=695442

Lakner, C., & Milanovic, B. (2013). Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession. *World Bank Policy research Working Paper 6719.*

Ledić, M., Rubil, I., & Urban, I. (2024). Using tax records to correct for under-representation of top income sources in surveys. *Fiscal Studies*, 1475-5890.12363. https://doi.org/10.1111/1475-5890.12363

Luxembourg Income Study (LIS) Database (2025). http://www.lisdatacenter.org (multiple countries; June 2024 - January 2025). Luxembourg: LIS.

Lustig, N. (Ed.). (2018). Commitment to Equity Handbook: Estimating the Impact of Fiscal Policy on Inequality and Poverty. Brookings Institution Press, D.C.

Lustig, N. (2020). The 'missing Rich' in Household Surveys: Causes and Correction Approaches (Vol. 520). ECINEQ, Society for the Study of Economic Inequality.

Medeiros, M., de Castro Galvão, J., & de Azevedo Nazareno, L. (2018). Correcting the Underestimation of Top Incomes: Combining Data from Income Tax Reports and the Brazilian 2010 Census. *Social Indicators Research*, *135*(1), 233–244. https://doi.org/10.1007/s11205-016-1498-8

Piketty, T., Yang, L., & Zucman, G. (2019). Capital Accumulation, Private Property, and Rising Inequality in China, 1978–2015. *American Economic Review*, *109*(7), 2469–2496. https://doi.org/10.1257/aer.20170973

Ravallion, M. (2022). Missing Top Income Recipients. *The Journal of Economic Inequality*, 20(1), 205–222. https://doi.org/10.1007/s10888-022-09530-0

Tamkoç, M. N., & Torul, O. (2020). Cross-sectional facts for macroeconomists: wage, income and consumption inequality in Turkey. *The Journal of Economic Inequality*, *18*(2), 239-259.

Tekgüç, H., & Eryar, D. (2025). Redistribution Trends in Turkey: Unintended Consequences vs. Deliberate Policies. *Review of Development Economics*, https://doi.org/10.1111/rode.13215

TURKSTAT. 2021. "Income and Living Conditions Survey." Accessed March 2024. https://data.tuik.gov.tr/Bulten/Index?p=Gelir-ve-Yasam-Kosullari-Arastirmasi-2021-45581

TURKSTAT (2019). Household Budget Surveys. Accessed August 2023. https://evam.tuik.gov.tr/dataset/detail/25

TURKSTAT (2023)." National Accounts - Institutional Sector Accounts." Accessed June 2023. https://data.tuik.gov.tr/Kategori/GetKategori?p=ulusal-hesaplar-113&dil=2

Van Der Weide, R., Lakner C., & Ianchovichina, E. (2018). Is Inequality Underestimated in Egypt? Evidence from House Prices. *Review of Income and Wealth* 64: S55-S79.

World Bank (2024). "World Bank Databank World Development Indicators." Accessed April 2024. https://databank.worldbank.org/source/world-development-indicators

WID (2020). World Inequality Database. Accessed March 2024. http://wid.world.

Yonzan, N., Milanovic, B., Morelli, S., & Gornick, J. (2022). Drawing a Line: Comparing the Estimation of Top Incomes between Tax Data and Household Survey Data. *The Journal of Economic Inequality*, *20*(1), 67–95. https://doi.org/10.1007/s10888-021-09515-5

Zucman, G. (2023). Globalisation, taxation and inequality. *Fiscal Studies*, *44*(3), 229–235. https://doi.org/10.1111/1475-5890.12341

Zwijnenburg, J. (2022). The Use of Distributional National Accounts in Better Capturing the Top Tail of the Distribution. *The Journal of Economic Inequality,* 20(1), 245–254. https://doi.org/10.1007/s10888-022-09534-w

TABLES AND FIGURES

Table 1: Comparing Scenarios

Gini Coefficient	2003	2007	2011	2015	2019
Raw Data: Household Budget Survey	0.408	0.357	0.372	0.380	0.351
Lower Bound: reweighting to account for non-response	0.477	0.382	0.405	0.450	0.379
Upper Bound: imputing all the discrepancy to top 10 %	0.605	0.576	0.586	0.594	0.547
Scenario 1	0.587	0.558	0.568	0.576	0.529
Scenario 2	0.499	0.462	0.459	0.486	0.415
Scenario 3	0.529	0.514	0.517	0.520	0.514
Scenario 4	0.468	0.436	0.449	0.457	0.433
Scenario 5	0.556	0.547	0.551	0.556	0.548
Scenario 6	0.533	0.516	0.522	0.531	**0.491**
Scenario 7	0.469	0.436	0.442	0.465	0.403
Alvaredo et al. (2019) (per adult)	0.610	0.560	0.580	0.590	-
Benchmark 1: Ceritoğlu et al. (2023)					0.517
Benchmark 2: Karabulut (2024)					0.501

Notes: Estimations are based on Household Budget Survey data. We present Alvaredo et al. (2019) estimates for Turkey for comparison purposes. They also employ Household Budget Survey data for Turkey.

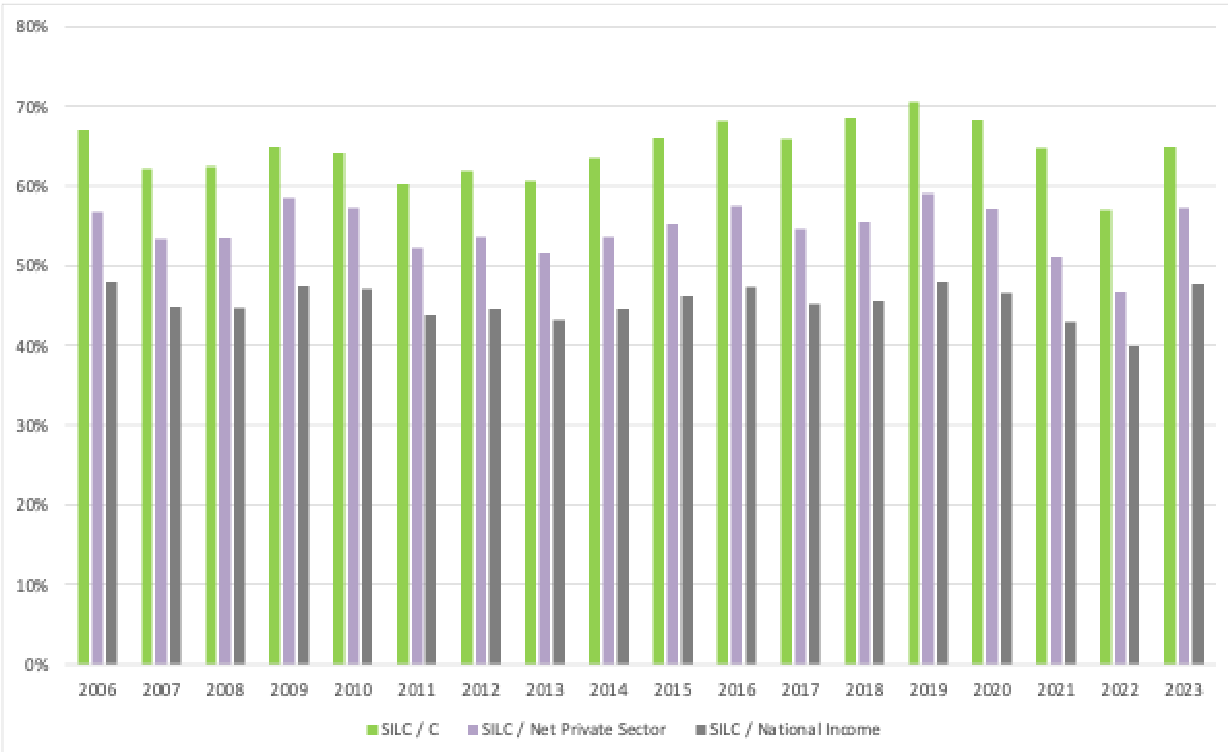
Table 2: Corrections for other countries

	Mali		**Egypt**		**India**		**Vietnam**		**Ivory Coast**		**China**	
	LIS Gini	Corrected Gini	LIS Gini	Corrected Gini	LIS Gini	Corrected Gini	LIS Gini	Corrected Gini	LIS Gini	Corrected Gini	LIS Gini	Corrected Gini
2000												
2002											0.39	0.44
2003												
2004					0.51	0.60						
2005							0.38	0.39				
2006												
2007							0.41	0.39				
2008												
2009							0.41	0.37				
2011	0.52	0.64			0.52	0.63	0.39	0.43				
2012			0.53	0.66								
2013	0.51	0.59					0.38	0.44			0.42	0.53
2014	0.40	0.54										
2015	0.40	0.52							0.58	0.68		
2016	0.40	0.53										
2017	0.38	0.51										
2018	0.40	0.50									0.42	0.54
2019	0.41	0.52										
2020	0.40	0.51										

Sources: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.

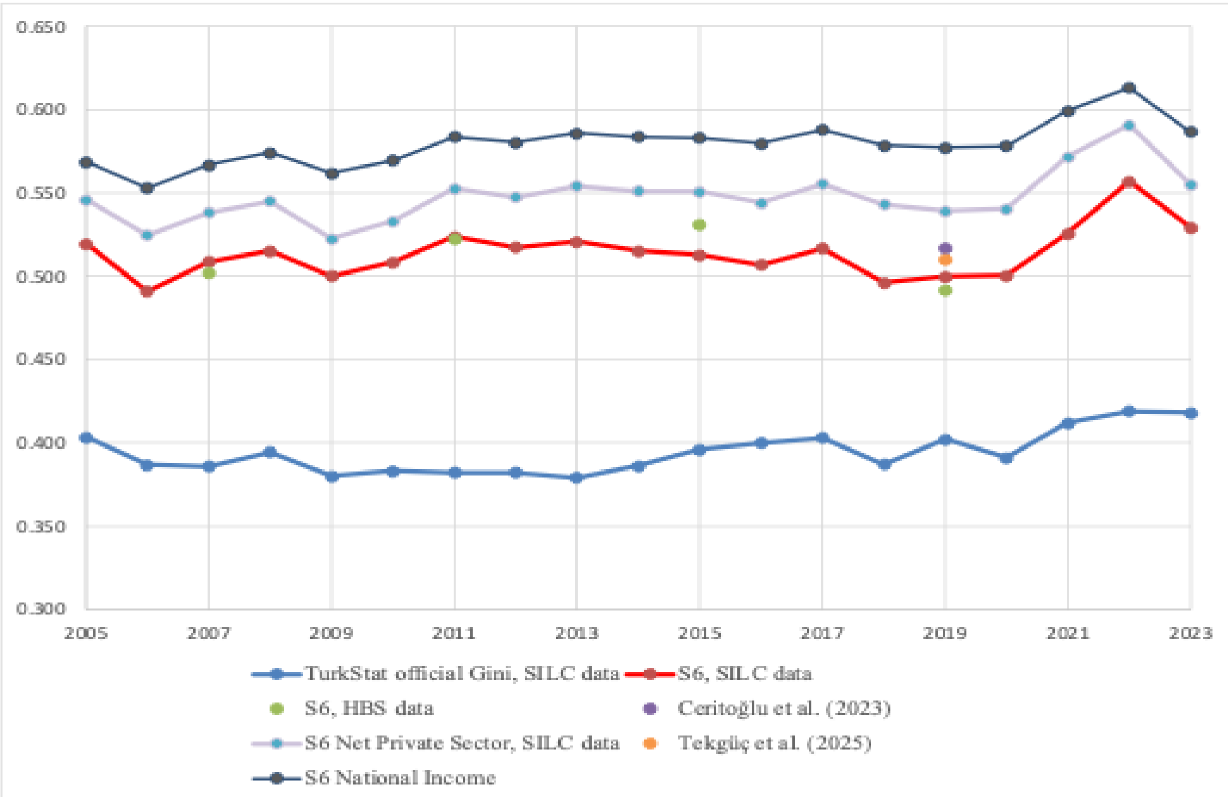
Notes: We correct Gini estimates using the average of China, France, and the USA wealth distributions for all countries except China. For China, we use China's wealth distribution estimates from WID. LIS Gini: Gini coefficient obtained from LIS data. Corrected Gini: S6 including retained earnings, missing income between survey total and net private sector (GDP – G – CFC) is distributed with respect to wealth distribution (scenario 6).

Figure 1: **Survey Coverage Rate**



Notes: SILC: Survey of Income and Living Conditions, official survey for income inequality in Turkey. C: Final consumption expenditure of resident households. Net Private Sector = GDP – Government Spending (G) – Consumption of Fixed Capital (CFC) = C + (I – CFC) + X – M + change in stocks. National Income = GDP – CFC.

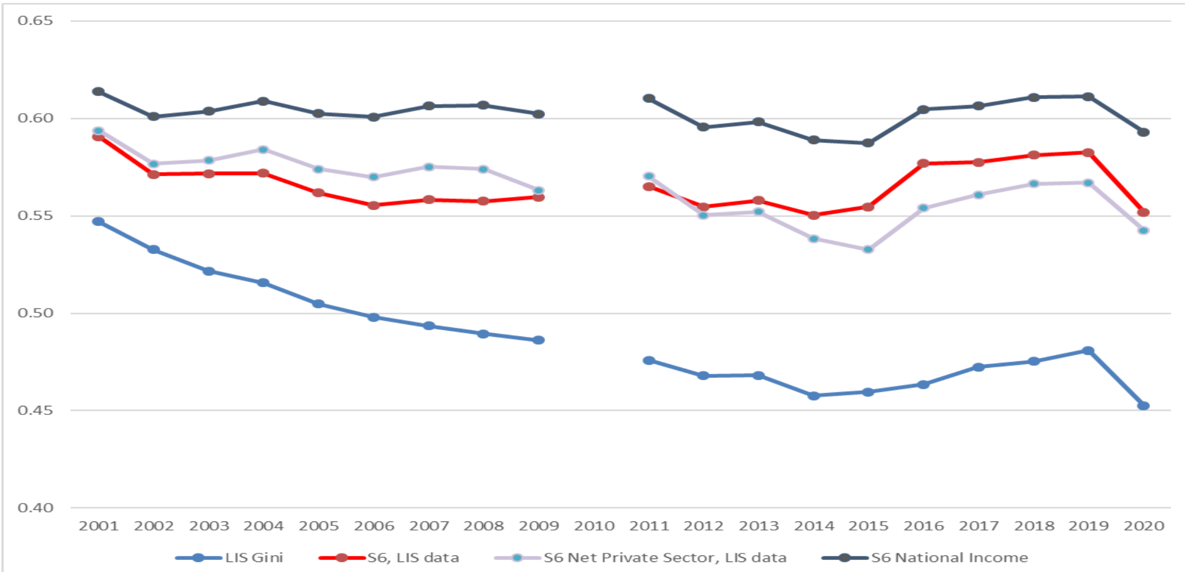
Figure 2: Total Household Income Inequality, Official and Corrected Estimates



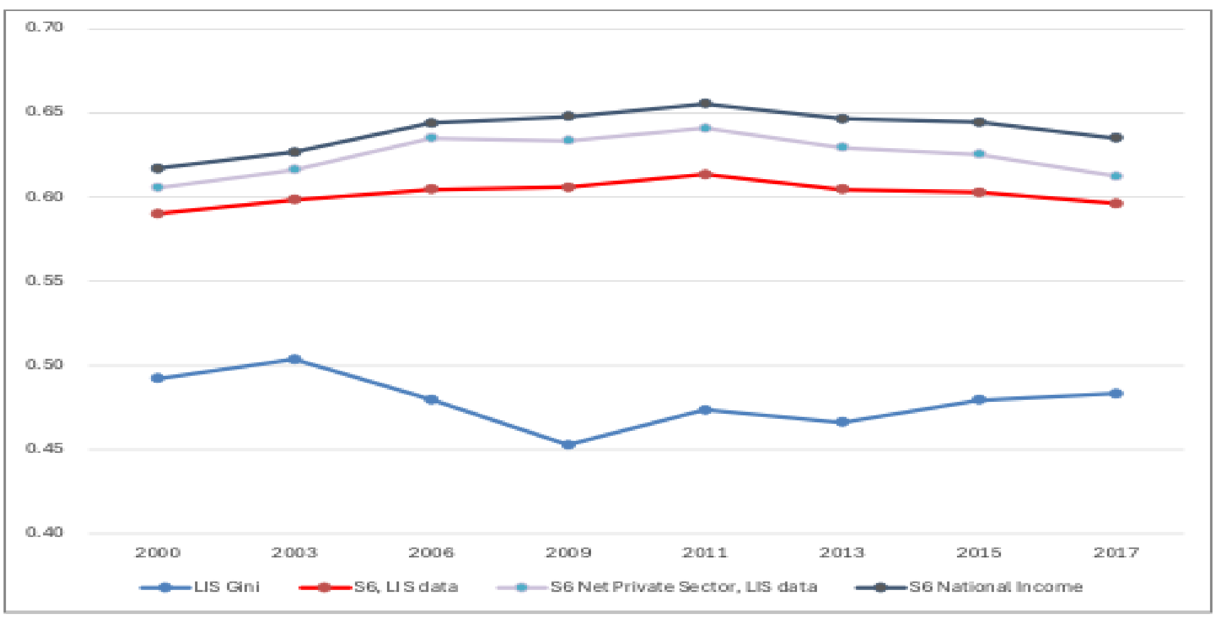
Notes: SILC: Survey of Income and Living Conditions, official survey for income inequality in Turkey. HBS: Household Budget Survey. S6: Scenario 6 correction. *TurkStat official Gini, SILC data*, shows official household disposable income inequality per TURKSTAT. *S6, SILC data*, represents our estimates for household disposable income inequality corrected for missing rich households. *S6, HBS data*, shows our estimates based on HBS data for household disposable income inequality corrected for missing rich households. *S6 Net Private Sector* represents our estimates for inequality of Net Private Sector including household disposable income plus retained profits. Retained earnings are profits of firms not distributed households but are controlled by richest households. *S6 National Income* represents our estimates for inequality of national income corrected for missing rich. In practice, national income is calculated as GDP minus consumption of fixed capital (GDP – CFC).

Figure 3: **Correcting Gini Coefficients for Brazil and Chile, LIS data, HH Total Income**

Panel A: Brazil



Panel B: Chile



Sources: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.

Notes: S6: missing income between survey total and final consumption expenditure of households is distributed with respect to wealth distribution (Scenario 6). *S6 Net Private Sector*: missing income between survey total and Net Private Sector (=GDP – G – CFC) is distributed with respect to wealth distribution (scenario 6). *S6 - National income*: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.

APPENDIX A

Table A.1. Comparison of HBS with National Accounts

		2011	
(million TL, nominal values)	**HBS**	**NA**	**HBS/NA**
Mixed income, except own final use	107,231	296,286	36%
Own final use	66,970	87,690	76%
Gross wages	349,781	372,235	94%
Capital income (gross)	25,836	241,530	11%
Public & private transfers	115,467	130,835	88%
Direct taxes	(35,693)	(55,752)	64%
Social Security contr.	(77,960)	(103,937)	75%
Disposable income	551,633	968,887	57%
Total Spending	507,464	876,892	58%
		2015	
(million TL, nominal values)	**HBS**	**NA**	**HBS/NA**
Mixed income, except own final use	180,171	450,611	40%
Own final use	109,431	127,769	86%
Gross wages	617,943	682,611	91%
Capital income (gross)	41,623	443,891	9%
Public & private transfers	196,852	207,892	95%
Direct taxes	(70,507)	(94,513)	75%
Social Security contr.	(156,423)	(195,345)	80%
Disposable income	919,089	1,622,915	57%
Total Spending	824,253	1,403,965	59%
		2019	
(million TL, nominal values)	**HBS**	**NA**	**HBS/NA**
Mixed income, except own final use	284,259	798,022	36%
Own final use	184,573	199,158	93%
Gross wages	1,244,931	1,346,679	92%
Capital income (gross)	73,389	557,249	13%
Public & private transfers	367,110	427,172	86%
Direct taxes	(138,865)	(175,278)	79%
Social Security contr.	(321,223)	(372,404)	86%
Disposable income	1,694,175	2,780,598	61%
Total Spending	1,525,118	2,441,247	62%

Notes: Reproduced from Tekgüç and Eryar (2025: Table A.1). HBS: Household Budget Survey. NA: National Accounts. In HBS, we classified all incomes from entrepreneurial activities under mixed income. Undoubtedly, some of these entrepreneurial incomes are from profits from incorporated firms. Unfortunately, there is no information in HBS to identify which entrepreneurial activities are incorporated firms. Social security contributions include unemployment contributions.

Table A.2: Wealth Distribution

2019		
Ventiles	**WID (USA, China, France)**	**Ceritoğlu et al. (2023)**
1	-0.1%	1.9%
2	0.1%	
3	0.1%	
4	0.2%	
5	0.3%	
6	0.4%	
7	0.5%	
8	0.8%	
9	1.0%	
10	1.3%	
11	1.6%	42.8%
12	2.0%	
13	2.4%	
14	3.0%	
15	3.7%	
16	4.7%	
17	6.1%	
18	8.4%	
19	14.8%	13.3%
20	48.6%	42.0%

Sources: Wealth distribution of China, USA, and France from https://wid.world/, Turkey's wealth distribution: Ceritoğlu et al. (2023).

Table A.3: Disposable Income, Coverage Rate, and Amount of Correction

	2003	2007	2011	2015	2019
GDP (2019 Prices)	1,997,617	2,682,584	3,081,287	3,842,143	4,317,810
NA Household Income (2019 Prices)	1,345,987	1,719,966	2,124,963	2,652,330	2,766,126
NA Household Consumption (2019 Prices)	1,296,524	1,656,760	1,923,199	2,294,500	2,441,758
HBS Disposable Income (2019 Prices)	776,426	982,950	1,209,841	1,502,068	1,694,175
(NA-HBA)/NA Household Disposable Income	42%	43%	43%	43%	39%
(C -HBS) / C	40%	41%	37%	35%	31%
Amount of Correction as a % of NA Household Disposable Income					
Upper Bound	36%	37%	37%	38%	33%
Scenario 1	40%	41%	41%	42%	37%
Scenario 2	40%	41%	41%	42%	37%
Scenario 3	**24%**	**29%**	**28%**	**27%**	**33%**
Scenario 4	15%	16%	15%	15%	16%
Scenario 5	39%	45%	42%	41%	**48%**
Scenario 6	36%	37%	37%	38%	33%
Scenario 7	36%	37%	37%	38%	33%

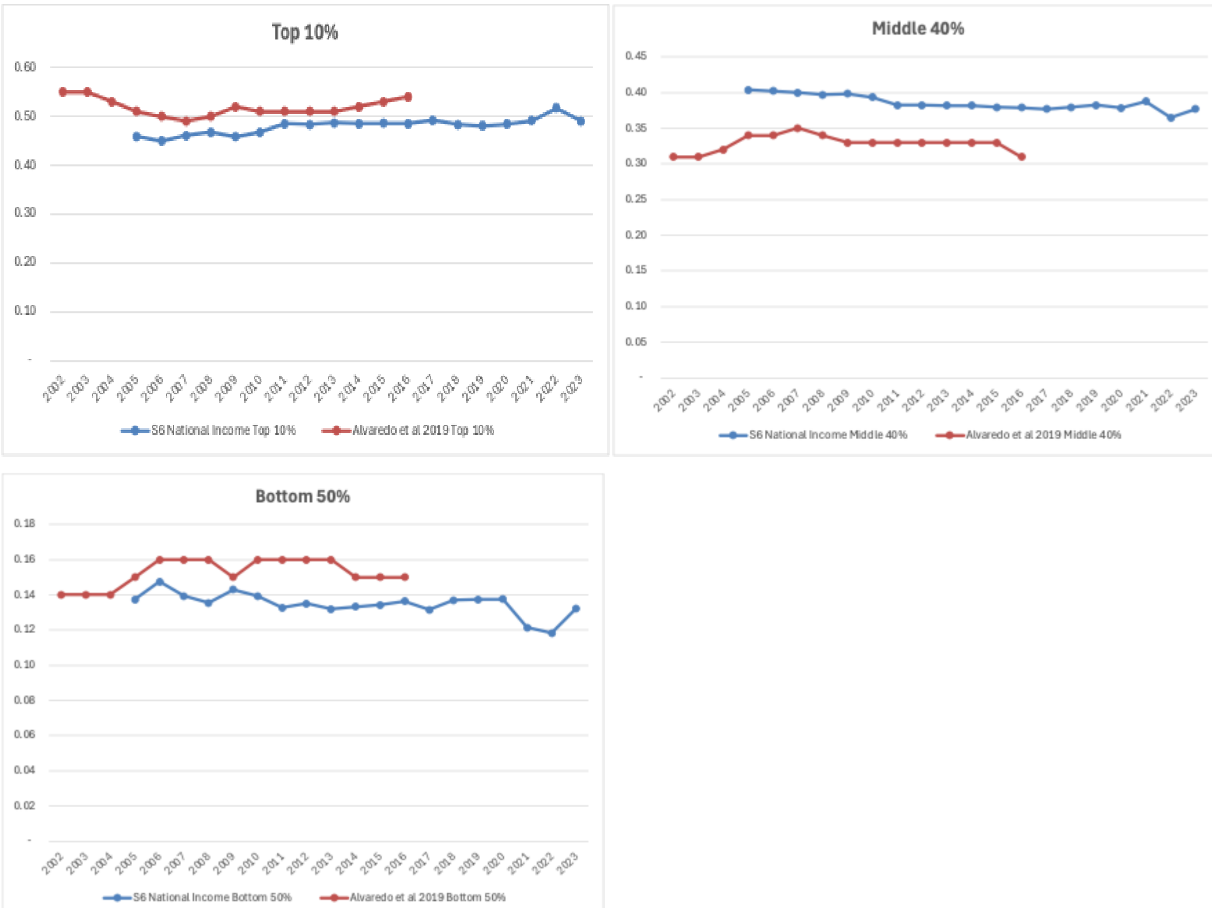
Source: Authors' own calculations based on TURKSTAT Data.
Notes: In scenario 5, the total amount of imputation exceeds the discrepancy between HBS and NA. This is because in the first step, we impute an amount based on Pareto coefficients, which come from Lebanese tax data, and then we impute 10% of GDP.

Table A.4: Correcting with Lebanese Tax Data

>%Y	Threshold	HBS Pareto Coefficients	Lebanon Income Tax 2015-2017 Coefficients	HBS Above-Threshold Average Annual Household Income	Above-Threshold Average Annual Household Income After Correction
80	29,017	1.73	3.48	50,205	100,978
81	29,944	1.71	3.43	51,295	102,822
82	30,846	1.70	3.39	52,457	104,500
83	31,609	1.70	3.28	53,708	103,828
84	32,650	1.69	3.19	55,059	104,020
85	33,788	1.67	3.04	56,524	102,865
86	34,961	1.66	2.92	58,095	101,925
87	36,217	1.65	2.76	59,858	99,977
88	37,495	1.65	2.63	61,755	98,473
89	39,124	1.63	2.48	63,877	97,129
90	41,227	1.61	2.37	66,261	97,559
91	43,166	1.60	2.28	68,927	98,358
92	45,433	1.59	2.22	72,021	100,944
93	48,140	1.57	2.19	75,652	105,535
94	51,249	1.56	2.18	79,931	111,810
95	54,857	1.56	2.18	85,386	119,453
96	60,080	1.54	2.16	92,341	129,835
97	67,160	1.52	2.12	101,971	142,183
98	77,111	1.52	2.08	117,172	160,763
99	97,930	1.51	2.29	148,107	224,604

Source: Alvaredo et al. (2019) and authors' own calculations.
Notes: Last column shows the above-threshold average annual household income after applying the Pareto coefficients from the Lebanese tax data. As seen in the last column, the ranking of each decile is distorted.

Figure A.1 Comparing Alvaredo et al. (2019) versus S6 National Income corrected estimates for Turkey



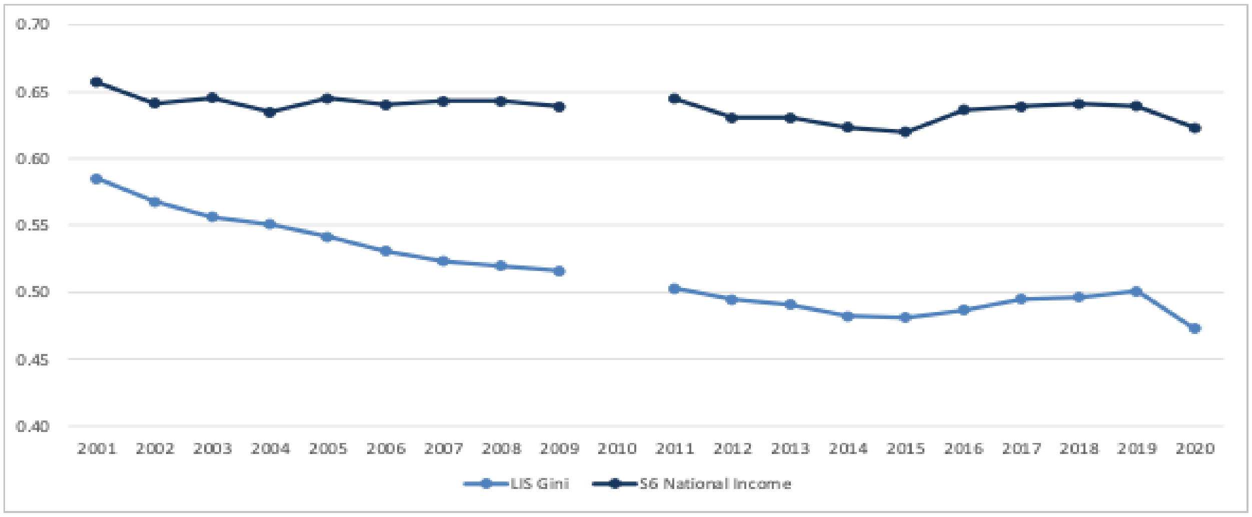
Sources: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.
Notes: S6 - National income: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.

APPENDIX B

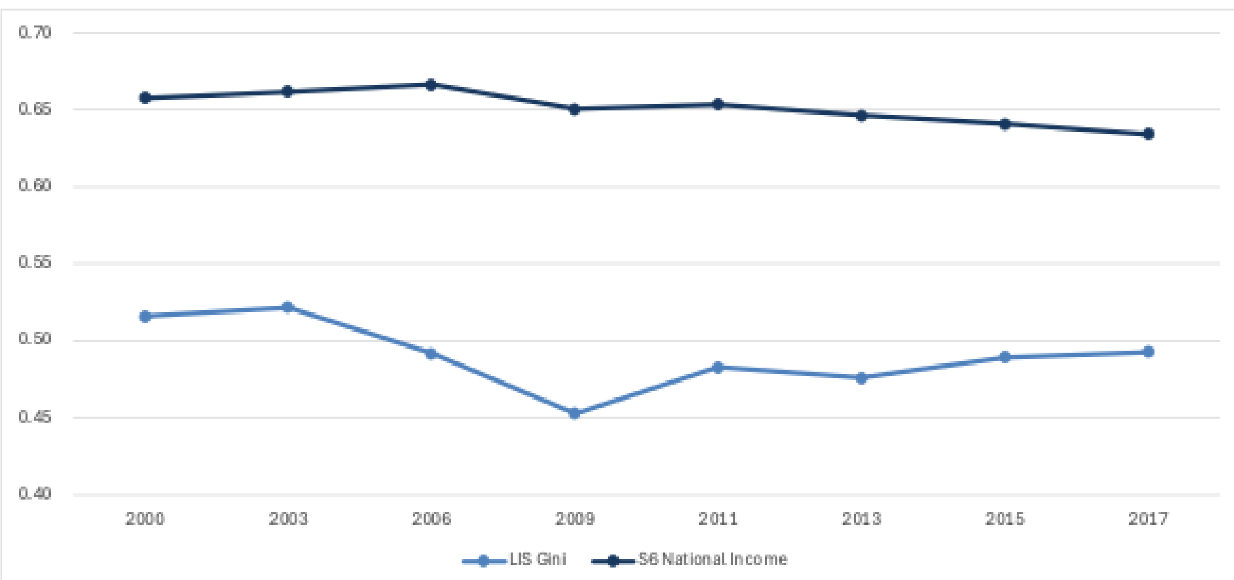
In our study, we calculate Gini coefficients using household total income. In contrast, De Rosa et al. (2024) compute the Gini coefficient based on household per capita income. To enable a meaningful comparison, we need to reconcile this difference. Therefore, we recalculated the Gini coefficients for the relevant countries using household per capita income. The main difference between our raw estimates and those of De Rosa et al. (2024) lies in the uncorrected (raw) Gini estimates. One of our aim is to develop an easily replicable procedure. Hence we use total household income in LIS database (*hitotal*) without any modification. Our uncorrected LIS Gini estimates (blue line) show the Gini coefficients derived from the raw data without any correction, and for Brazil, they are lower than De Rosa et al. (2024: Figure 1) raw Gini estimates. However, the size of the correction between raw data and National Income Gini is roughly the same in both studies: roughly 7-9 percentage points in 2001 and roughly 14-15 percentage points in 2015. Moreover, there is no sustained decline in Gini estimates between 2001 and 2015 after the correction. In De Rosa et al. (2024), the National Income Gini estimate hovers around 0.69 throughout the period, whereas in this study, the National Income Gini estimate hovers around 0.64. The uncorrected Gini estimates from LIS for Chile are lower than De Rosa et al. (2024) estimates. Unlike the case of Brazil, our corrections for Chile are larger than De Rosa et al. (2024). In both studies there is sustained but incremental decline in Gini estimates for National Income.

Figure B.1: Correcting Gini Coefficient for Brazil and Chile, LIS data, HH Per Capita Income

Panel A



Panel B



Sources: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.
Notes: S6: missing income between survey total and final consumption expenditure of households is distributed with respect to wealth distribution (Scenario 6). *S6 National income*: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.

APPENDIX C

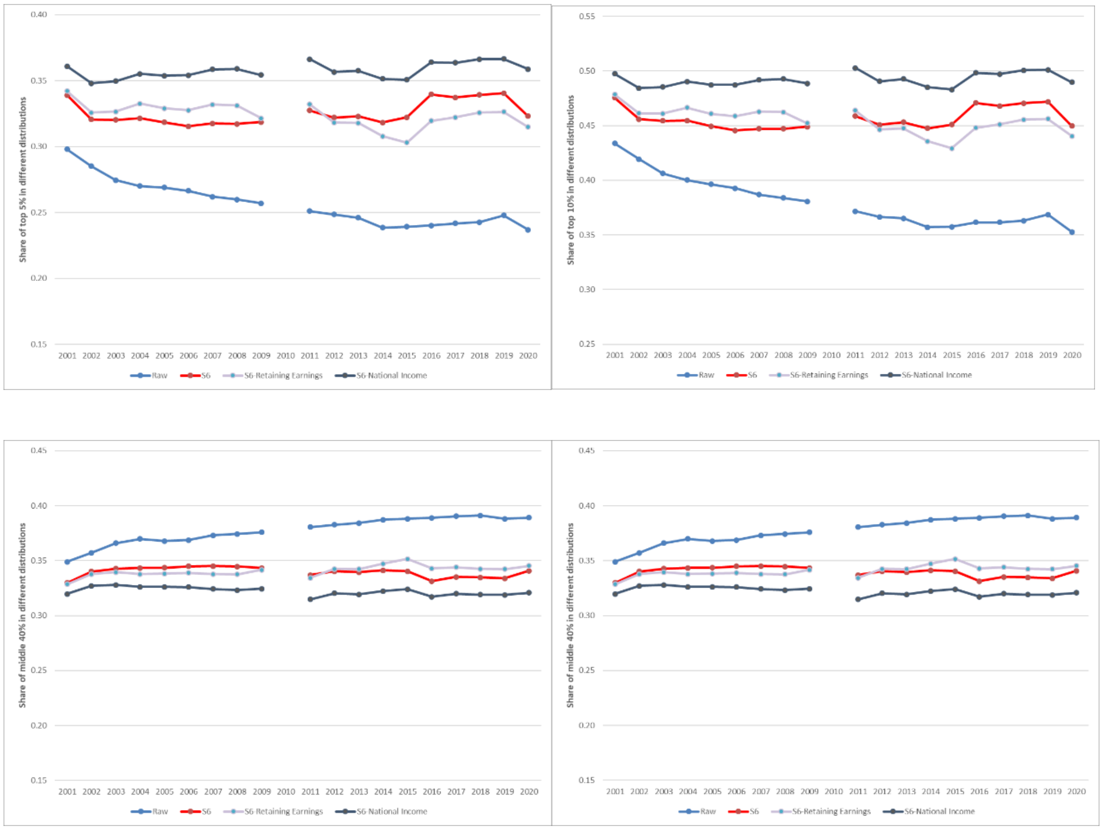
Figure C.1. Turkey

Income shares of the top 5%, top 10%, middle 40%, and bottom 50% in four different distributions: raw data, scenario 6, scenario 6 net private sector, and scenario 6 national income.

Source: Household income (SILC) and National Accounts data from TURKSTAT, wealth data from World Inequality Database (WID).
Notes: S6: missing income between survey total and final consumption expenditure of households is distributed with respect to wealth distribution (Scenario 6). *S6 Net Private Sector*: missing income between survey total and Net Private Sector (=GDP – G – CFC) is distributed with respect to wealth distribution (scenario 6). *S6 - National income*: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.

Figure C.2. Brazil

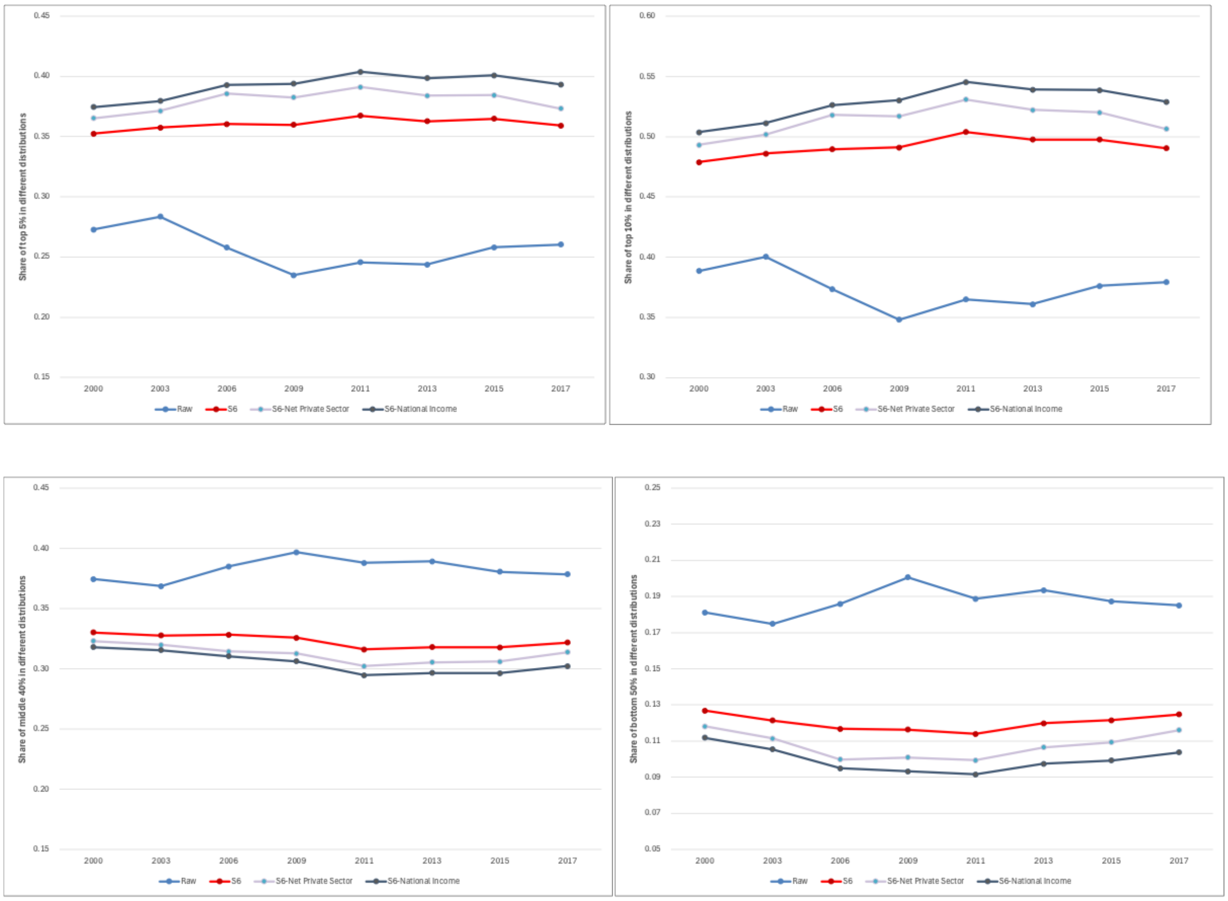
Income shares of the top 5%, top 10%, middle 40%, and bottom 50% in four different distributions: raw data, scenario 6, scenario 6 net private sector, and scenario 6 national income.



Source: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.
Notes: S6: missing income between survey total and final consumption expenditure of households is distributed with respect to wealth distribution (Scenario 6). *S6 Net Private Sector*: missing income between survey total and Net Private Sector (=GDP – G – CFC) is distributed with respect to wealth distribution (scenario 6). *S6 - National income*: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.

Figure C.3. Chile

Income shares of the top 5%, top 10%, middle 40%, and bottom 50% in four different distributions: raw data, scenario 6, scenario 6 net private sector, and scenario 6 national income.



Source: Income survey data from LIS; Wealth data from WID; and GDP and its expenditure components from the World Bank Databank WDI.
Notes: S6: missing income between survey total and final consumption expenditure of households is distributed with respect to wealth distribution (Scenario 6). *S6 Net Private Sector*: missing income between survey total and Net Private Sector (=GDP – G – CFC) is distributed with respect to wealth distribution (scenario 6). *S6 - National income*: missing income between survey total and national income (GDP-CFC) is distributed with respect to wealth distribution.