CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 600

Drawing a Line

Comparing the Estimation of Top Incomes Between Tax Data and Household Survey Data

Nishant Yonzan, Branko Milanovic, Salvatore Morelli and Janet Gornick

January 2021

University of Naples Federico II   **University of Salerno**   **Bocconi University, Milan**

WORKING PAPER NO. 600

Drawing a Line

Comparing the Estimation of Top Incomes Between Tax Data and Household Survey Data

Nishant Yonzan[*]**, Branko Milanovic**[**]** , Salvatore Morelli**[***]** and Janet Gornick**[**]

Abstract
The paper uses the flexibility of household survey data to align their income categories and recipient units with the income categories and units found in data produced by tax authorities. Our analyses, based on a standardized definition of fiscal income, allow us to locate, for top-income groups, the sources of discrepancy. We find, using the cases of the United States, Germany, and France, that the results from survey-based and tax data correspond extremely well (in terms of total income, mean income, composition of income, and income shares) above the 90th percentile and up to the top 1 percent of the distribution. Information about income composition, available in the US, allows us to investigate the determinants of this gap in the US. About three-fourths of the tax/survey gap is due to differences in non-labor incomes, especially self-employment (business) income. The gap itself may be due to tax-induced re-classification of income from corporate to personal or/and to lower ability of surveys to capture top 1 percent incomes.

Keywords: Income, Inequality, Survey Data, Tax Data.

JEL Classification: D31.

[*] PhD Program in Economics, Graduate Center, CUNY. *Corresponding author*.
    Email: nyonzan@gradcenter.cuny.edu.

[**] Stone Center on Socio-Economic Inequality, Graduate Center, CUNY.

[***] Stone Center on Socio-Economic Inequality, CUNY Graduate Center (NY) and CSEF.

Table of contents

1. Introduction

Research on income distributions has intensified, since the beginning of the twenty first century, after a long hiatus following Kuznets' (1955) foundational contribution. This recent wave of research has focused heavily on the richest segments of populations, which have seen increasing income shares, especially in high-income countries (A. B. Atkinson and Piketty 2007; A. B. Atkinson and Piketty 2010; and Alvaredo et al. 2018). Hence, a complete understanding of the evolution of income inequality, within countries, requires sound data on income concentration at the top. The growing need to capture distributional dynamics at the top of income distributions has put standard data sources for the study of income distribution, such as data based on household surveys, under intense scrutiny, and has stimulated the use of previously unexplored data sources, such as those generated by tax authorities.

The use of the latter source, tax data, indicates that top income shares are generally larger than what has been estimated from household survey data. Careful work by Burkhauser et al. (2012) compared, for the case of the US, the evolution of top income shares between survey data from restricted (not publicly available) files of the Current Population Survey (CPS) and tax-based estimates from Piketty and Saez (2003). Their main results find a close match between the income share trends reported in the two sources "with the exception of the richest 1 percent" (page 372).

Similarly, Bartels and Metzing (2019) compared results based on income tax data with results based on data from a widely-used survey, the German Socio-Economic Panel (SOEP), and found that "the estimates of the income share of the top 10-5% and top 5-1% are of similar magnitude in both data sources. The income share of the [top] (10-5%) is around 12 percent in the SOEP data and between 11.2% and 11.8% in the income tax data. The upper 4% do not differ significantly until 2008 in both datasets and are between 13.4% and 15%" (page 129). Bartels and Metzing (2019) also found that "there are large quantitative differences for the top 1 percent between SOEP and tax data. Tax data measure 3 to 6%-points higher income shares for the top 1%" (page 129).

In line with the US research, recent work by Burkhauser et al. (2018), using UK data, finds that "household surveys are the main source of information about overall inequality levels and trends in most countries around the world but do not capture income at the extreme top of the distribution very well." However, somewhat differently from the US case, the UK tax and survey data appear to match relatively well, except within the top 2 percent of the income distribution. This indicates that the point in the distribution where the two types of data diverge varies somewhat across countries.

No systematic cross-country analyses have been conducted that assess where exactly the two types of data begin to differ and what drives the differences.[3] The aim of this paper is to shed light on this question, using cross-national harmonized microdata from the Luxembourg Income Study (LIS) database. The core research question in this paper is: Where do the two sources begin to diverge, and why?

There are several possible reasons underlying divergent estimates of top income shares based on data from administrative tax authorities versus data from surveys.

First, the definition of income used by the two may be different (i.e., tax data do not use a standard economic definition of income, but only include sources of income whose reporting is required for tax purposes and such sources differ from country to country, and over time within individual countries).

Second, income flows may refer to different time periods, and may accrue to different units of analysis (i.e., tax units vs households).

Third, the two types of data are plagued by different under-reporting problems which are source-specific. Income from surveys is self-reported whereas tax data reflect income reported for tax purposes, some of which may be also self-reported. The general presumption is that individuals have few incentives or direct benefits from misreporting their incomes to household surveys. On the contrary, individuals have clear incentives to misreport taxable income, i.e., to minimize their tax liabilities. Furthermore, whereas weaknesses in tax data may stem from tax exemptions, evasion, and avoidance, parallel problems with household survey data include under-reporting, survey non-response (refusal to participate in surveys), and item non-response (refusal to provide specific information). The two sources may be more or less affected by these problems. For instance, richer households or tax units typically have higher rates of survey non-response and, among those who respond, higher rates of income under-reporting, especially of particularly sensitive source of income such as capital or business income.

Fourth, even in the absence of non-response problems, household survey data may also return biased estimates of top income shares if their sampling frame does not allow for over-sampling of rich households. In other words, the presence of a thick tail and highly skewed distribution of income will

---

[3] Burkhauser et al. (2017) for the UK, and Bartels and Metzing (2019) for Germany, provide examples of country-specific survey versus tax data comparative exercises. The former study uses the Family Resource Survey (FRS), specific to the UK, and the latter uses the German Socio-Economic Panel (SOEP) survey data. Both compare survey-based results to results from tax data, derived from the World Inequality Database (WID).

3

render standard random samples less suitable to correctly estimate the top tail. This is particularly problematic as all results are highly dependent on income ranks. Who is considered as being in the top 1 percent does not depend only on their own reporting, but on everyone else's as well. Similar problem may apply to tax-based data too if researchers are provided only a sample of all fiscal units.

The first two differences between income reported on tax and survey data described above are due to differential definitions of both income and units of analysis. These differences are mechanical in nature and are adjustable ex post. The latter two sources of difference, instead, relate to households' behavior and sampling design, both of which are core differences linked to the purposes for which the two types of data are constructed. In this paper, we address the first two sources of difference: income definitions and units. We use data provided by tax authorities (see the data section below for details) and compare them with survey-based data, from the LIS Database, for Germany (DE), France (FR), and United States (US). After adjusting for these two key differences (income and unit of analysis), residual discrepancies that still exist between the two data sources are due to fundamental differences in the data sources, including differences due to households' behavior and the population of tax units respectively.[4]

In this paper, we assess where in the income distribution the differences between these two types of data start and what might be driving these differences. This is important because recent attempts (Bourguignon 2018; Blanchet, Flores, and Morgan 2018; Lustig 2020) to combine the two data sources in order to get a better grasp on income distributions, require a cutpoint, i.e., a point in the distribution above which the comparison is focused. Should comparisons (and possible replacement of survey data by tax data) address just the top 1 percent or should they reach "deeper", including, say, the top five percent? Moreover, the causes of discrepancies have not been clarified: Are they due to differences in recorded levels of capital or labor incomes? Thanks to the flexibility and granularity of the LIS microdata, we are able to define and construct income variables and units of analysis to match closely those from tax sources and thus to investigate the origins of the observed discrepancies.

---

[4] Attempts have been made to adjust for differences due to behavioral (non-response) and sampling factors. See, e.g., Piketty, Saez, and Zucman (2018), Hlasny and Verme (2018), and Bourguignon (2018).

2. Data and Variables

All of our analyses based on survey data utilize microdata from the LIS Database, accessed via LIS' remote-execution system. All of our analyses based on tax data utilize data available in the studies listed in Table A.1.[5] Although our study draws on specific sources of survey and tax data, in the remainder of this paper -- for clarity and convenience -- we refer to the data that we are analyzing using generic labels: *survey data* and *tax data*, respectively.

Although both survey and tax data capture income distributions, there are substantial differences between them. The share of the population that reports its income to tax authorities in advanced economies is often high: in the US in 2013, about 90 percent of the population filed taxes. Income definitions used in tax data are defined by tax authorities and are shaped by tax codes. As a consequence, they vary both across countries, and within countries over time. Income shares, based on tax data, may not be comparable over time if substantial changes in tax rules have intervened in the meantime. Changes in tax rules/policies are often driven by political decisions about what types of income are taxable (and hence reported) and which are not. Moreover, tax-based income definitions, in some countries and at time points, may be inadequate. Tax-data sources generally do not apply internationally-accepted definitions of income, drawing on, for example, the Canberra Group Handbook on Household Income Statistics (2011, updated from 2001), which is widely used to standardize survey-based income data.

Survey data, on the other hand, are based on samples of populations. Sampling creates the potential for differences between the two types of data, especially at the top of the distribution; weakness in survey data arises if samples are not large enough to "catch" the top of the distribution and/or if the rich systematically underreport their income or refuse to participate. Yet, one advantage of survey data is that the definition of income is more stable over time and is typically constructed according to international conventions. Income estimates from survey data are thus, in principle, more comparable across countries and over time. Survey data have a further advantage that they are often available in highly disaggregated form; they generally include more income categories than tax data. These features have allowed us to use survey data to construct income definitions that match those available in tax data.

---

[5] The LIS Database can be accessed here: https://www.lisdatacenter.org/. The studies listed in Table A.1 represent the building blocks of the income concentration estimates available in the WID database. The WID database can be accessed here: https://wid.world/data/.

In what follows, we outline the process of aligning the mechanical differences -- the units of analysis and the definition of income -- across the two data sources.

3. Aligning the units of analysis between survey and tax database

The first notable difference between the tax and survey data sources is the unit of analysis. Units in tax data are defined according to tax filing requirements within each country. The legal definition of a tax unit varies across countries, and also within them over time. For instance, Germany, France, and the US allow for either a joint-couple or a single-adult to be filing units. On the other hand, UK and Italy, for example, require all adult individuals to file taxes separately. In the former cases, a tax unit is defined as a couple or an adult single (with or without dependents), while in the latter cases, a tax unit is exclusively an adult individual (with or without dependents). Moreover, changes in country-specific tax codes can affect the definition of tax filing units. This happened, for example, in the UK in 1990 and in Spain in 1988 when tax filing units were changed from a mixture of joint-couple and single-individual filers to exclusively individual tax filers.

Another issue to consider is that, in all countries, portions of the population do not file taxes. Non-filers are generally individuals and households at the lower end of the income distribution. In order to obtain the total number of units in an entire distribution, these "missing" non-filers have to be affixed to the tax filers recorded in the administrative tax sources (see Piketty and Saez 2003 for the details on this imputation). Generally, richer countries have more advanced tax systems that cover most of their populations. Specifically, average coverage in the years included in this study are as follows: 92 percent in the US, 76 percent in France, and 54 percent in Germany; note that this still classifies almost half of the German population as non-filers. This is a greater problem for middle- and lower-income countries, where tax filers can amount to fewer than 5 percent of the actual number of tax units in the population. In the remainder of this paper, tax units (both tax filers and tax non-filers) derived from tax data are referred to as *Fiscal-TUs*.

While Fiscal-TUs do vary across space and time, households, the preferred unit for household surveys, are almost always defined according to uniform international standards. Survey households -- referred to throughout this paper as *Survey-HHs* -- are defined as a collection of individuals who live in the same dwelling and generally have a joint budget. Thus, it is often the case that multiple couples and/or multiple single tax filing individuals live within one Survey-HH. While this grouping of persons would constitute two

or more Fiscal-TUs in the tax data, they are categorized as a single Survey-HH in the survey data. This implies that, on average, in any given country (or year), the number of Survey-HHs will be smaller than the number of Fiscal-TUs.

3.1 Total tax units from tax data

A tax unit is defined as a cohabiting married couple with or without dependents or as a single adult with or without dependents. The number of units that are recorded in tax statistics (i.e., the number of tax filers) are, as discussed above, fewer than the actual number of potential tax units in the population. In countries that allow for joint tax-filers, as in Germany, France and the US, the number of total tax units can be calculated from national population statistics as the sum of married couples and non-married adults (aged 20 or more). This is exactly the definition of tax units used by Bartels (2018) for the estimation of top income shares series in Germany. In the United States, the number of married women who file tax returns separately is "fairly small (about 1 percent of all returns in 1998)" (Piketty and Saez 2003). Hence, the control total of tax units used in Piketty and Saez (2003) to derive the top income shares series for the US is defined as the sum of married males and all nonmarried individuals aged 20 and over.[6]

The sources for the actual number of Fiscal-TUs are given in Table A.1. This affects the total income and income shares held by each fractile in the distribution.

3.2 Total tax units from survey data

Survey data sample representative households from populations. These representative households are then assigned weights relative to the total population in national population statistics. For instance, there are 16,703 sampled households in Germany in 2010 in the LIS database. Applying the sample household weights, we find that these sampled households represent 39.9 million households in the population. As noted, an average household is larger than an average tax unit. Hence, although both data sources use national population statistics, the number of households (39.9 million in Germany in 2010) in survey data is less than the number of tax units in the same year (49.2 million). As reported in Bricker et al. (2016), in the US, "Families in the bottom 99 percent are often split into multiple tax units, but a tax unit in the top 1 percent is almost always a family. Counting the top 1 percent (1.61 million) of tax units, then, effectively

---

[6] In this paper, we use term "control total" to refer to the total number of units or total income used in the denominator in order to calculate the number of units in each percentile or the share of income held by each income group respectively.

includes more families than counting the top 1 percent (1.22 million) of families in a survey" (page 266). The number of the x percent richest households would be smaller than the number of the x percent richest tax units. This mechanical difference affects the calculation of total income held by top income groups.

We, therefore, working with the survey data, we construct tax units from households using the following procedure:

A.   We separate the head-of-household (and his or her spouse if married) from each household. We use the weight of the head-of-household to calculate the total number of such units in the population.

B.   We treat other married couples within the household as additional tax units. To derive the exact number of such tax units, we use the weight of the closest kin of the household head. For instance, if the added couple within the household is the daughter of the head-of-household and her spouse, we use the weight of the daughter to calculate the total number of such tax units in the population. In most cases, the closest kin is either the son or daughter of the household head.

C.   All other identified single adults (age > 20) living within each household and having positive income are treated as separate tax units. For all such singles, we apply their own sampling weights to calculate the total number of such units in the population.

D.   Finally, the total number of tax units in a survey-based dataset is calculated by adding the total units from the above three steps, that is A + B + C. We refer to the tax units that we constructed from the survey data as *Survey-TUs*.

Figure 1 compares the total number of Survey-HHs, Survey-TUs, and Fiscal-TUs for the US, in 2013. There were a total of 163 million Fiscal-TUs compared to 123 million Survey-HHs. This is very similar to what was recorded in 2012 by Bricker et al. (2016) using different household survey data (i.e., the Survey of Consumer Finances); they reported that the "number of tax units (about 161 million in 2012) is approximately 30 percent higher than the number of families" (page 278). Once we adjusted our survey data to match the tax unit definition, almost all of the survey-versus-tax data discrepancy disappears. Indeed, our work produces 161 million Survey-TUs. The discrepancy, vis-a-vis Fiscal-TUs, is less than one percent.
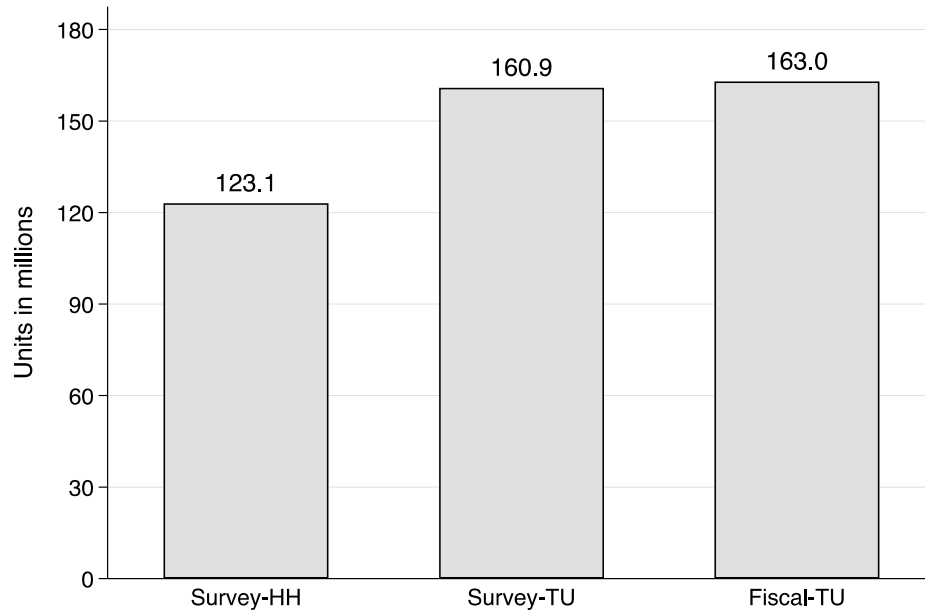
Figure 1: Comparison of Units for US in 2013

Source: See Table A.1; LIS Database.

Notes: This figure shows the total number of units in the US in 2013. The total number of households (Survey-HHs) is the aggregate household weights in the LIS US 2013 dataset; the total Survey-TU is the total number of tax units constructed using the household weights in LIS as described in the text; and the Fiscal-TU is the total units in tax data. The units are presented in millions.

Table 1 lists units for all available country-year combinations. For the US, the number of Survey-HHs was on average four-fifths of the total number of Fiscal-TUs. This "gap" almost disappears once we transform the Survey-HHs into Survey-TUs. After 2000, the difference in units between the two data sources narrows to, on average, less than one percent.

Table 1: Comparison of Units of Analysis in DE, FR, & US

| | Survey Data | | Tax Data | Ratio of units |
	Households (1)	Tax Units (2)	Tax Units (3)	Survey/Tax (4)
DE				
1989	28,623,306	33,531,863	34,376,745	0.98
2001	39,197,676	43,245,932	46,801,900	0.92
2004	39,672,688	43,636,233	46,338,000	0.94
2006	40,182,524	44,323,200	47,941,500	0.92
2007	40,554,568	45,055,381	48,296,900	0.93
2010	39,863,596	44,656,222	49,192,250	0.91
FR				
1984	20,325,672	25,566,811	24,572,248	1.04
1989	21,201,890	25,592,103	27,360,033	0.94
1994	23,155,000	27,530,216	30,038,236	0.92
2000	24,522,572	28,376,820	32,923,000	0.86
2005	24,918,384	28,864,379	35,572,000	0.81
2010	28,524,272	31,922,020	36,962,517	0.86
US				
1979	79,367,808	101,975,681	97,457,046	1.05
1986	89,835,680	116,832,994	110,683,650	1.06
1991	95,979,024	126,592,945	120,453,262	1.05
1994	99,087,648	128,096,618	124,715,805	1.03
1997	102,584,216	132,822,947	129,301,257	1.03
2000	108,289,768	139,337,395	134,473,000	1.04
2004	113,475,720	146,656,550	143,982,000	1.02
2007	116,880,816	150,464,925	149,875,000	1.00
2010	118,748,720	154,546,955	156,167,000	0.99
2013	123,052,856	160,926,621	162,998,000	0.99

Source: See Table A.1 for tax data; survey data units are calculated using LIS Database.

Notes: This table reports total households (column 1) and tax units (column 2) from survey data, and total tax units from tax data (column 3). Column 4 reports the ratio of tax units from the survey data (column 2) to tax units from the tax data (column 3).

The discrepancy in the numbers of units, between the two types of data, is slightly greater for Germany and France. For the entire period, on average, the difference between Fiscal-TUs and Survey-HHs is 16 percent for Germany and 24 percent for France (in favor of Fiscal-TUs). This difference is reduced to 6.6 percent for Germany and 9.5 percent for France once we convert Survey-HHs into Survey-TUs.

The different magnitude of the discrepancies between the Fiscal-TUs and Survey-TUs in Germany and France compared to the US may arise for a number of reasons. First, the choice of the weights used to

expand the number of identified couples in survey data may turn out to be more or less representative of the actual number of couples in the population. This source of variance is not often acknowledged in this type of exercises and it is hard to assess its specific relevance for each country.

Second, part of the discrepancy may arise from the slightly different procedures used to derive the number of non-filers or the total number of tax units from population statistics.

Third, there might be differing incentives across countries for some couples to file taxes as separate individual tax units despite being married. For instance, in France, married persons must file joint returns, except under strictly limited circumstances. In Germany and the US, tax authorities do not force married couples to file joint returns consistently; they can opt to file separate returns some years and joint returns in other years. Typically, if one spouse commands a higher level of income than the other spouse, it is beneficial to file a joint tax return. This is because the partner with the higher income may be able to use tax credits which the other partner is unable to use, as his/her income is not high enough. Hence, filing jointly typically provides married couples with more tax breaks. In 2019, the standard deduction for a married couple filing jointly in the US was $24,400. Conversely, for those filing separately, the standard deduction was halved, i.e., $12,200, the same as for single people. Moreover, joint filers are often eligible to receive additional tax credits, such as the Earned Income Tax Credit (EITC) for working people with low to moderate income; the American Opportunity and Lifetime Learning Education Tax Credit, reducing the amount of taxes allowing for college or graduate school tuition costs; the reimbursement or refund for adoption expenses when legally adopting a child; and the Child and Dependent Care Tax Credit, which can help offset the costs of a caregiver to look after children below age 13 or after a disabled spouse.

In what follows, we will use the Survey-TUs and Fiscal-TUs for the comparative exercise. Reference to tax units or fiscal units will imply only these concepts in the respective data sources.

3.3 Aligning the definition of income

The second important source of difference between the two data sources is the income concept used to construct total income. For the tax data, as explained, the definition of income is derived from fiscal authorities in each country. In the US tax data, the main income indicator includes "salaries and wages, small business and farm income, partnerships and fiduciary income, dividends, interest, rents, royalties, and other small income reported as income" (Piketty and Saez 2003). Consistent with income tax-based reporting, this definition of income includes taxable withdrawals and payments from retirement plans (for example, Old-Age, Survivors, and Disability Insurance) and other forms of taxable cash transfers (for

example, unemployment benefits). This definition excludes all untaxed (cash or in-kind) transfers (for example, Supplemental Security Income).

As noted above, the definition of income, based on tax sources, varies across countries, as well as within countries over time. Incomes based on survey data, on the other hand, are generally defined by standard international conventions and are more comparable across countries and over time. A key point to stress is that, in an alignment exercise such as the one we report in this paper, there is a fundamental asymmetry: We can use survey data to construct "fiscal income" (the income definition used in estimations based on tax data), but we cannot use tax data to construct "survey income" (the standard income definitions derived from survey data). So, while using income definitions based on survey data may be preferable in general, because they are based on concepts grounded in economic theory, the nature of the two types of data -- especially the greater flexibility and higher level of disaggregation in survey data -- requires us to follow the strategy of creating "fiscal income" from survey data, rather than the reverse.

Table 2 lists the income variables in the LIS Database that we use to construct fiscal income for each country. In the remainder of this paper, when we use the term "fiscal income", we refer to both income reported in the tax-based data and to the income indicators that we have created in our survey data.

For the US, we further disaggregate fiscal income into three income components. (The available tax data do not allow us to do this in the other two countries). Our interest is not only limited to matching the definition of income and units, but also in identifying the sources of discrepancies between the two data sources. These discrepancies may arise due to one or more income components. To find out which sources drive the mismatches, we disaggregate fiscal income into income from labor, income from business (or self-employment), and income from capital.

To construct the components of fiscal income in the survey data, we relied on the definitions provided by Piketty and Saez (2003) for the US. Fiscal income, in their work, can be summarized as the sum of labor, business, and capital income.

Piketty and Saez define *labor income* as income from wages, salaries, and pensions. (In this context, "pensions" refers to taxable income transfers). To match this definition of labor income, we summed these variables, available from LIS: paid employment income (variable *i11* in the LIS Database) and a selection of pension/transfer variables, specifically public contributory pensions (*i32*), private pensions (*i33*),

allowances for maternal and parental leave (*i411*), and unemployment benefits (*i42*).[7] (We selected these specific pension/transfer variables because they are taxable).

Piketty and Saez define *business income*, also referred as entrepreneurial income, as non-corporate business profits. To match this, using our survey data, we use self-employment income from farm and non-farm activities (*i12*). Finally, Piketty and Saez define *capital income* as income derived from ownership of capital, specifically interest, dividend, and rental income. To match this, using our survey data, we constructed capital income as the sum of interest and dividend income (*i21*) and rental income (*i22*). Finally, we construct total fiscal income from the survey data by aggregating these labor, business, and capital income components as defined here.

Similarly, the income items used to construct fiscal income for Germany and France are listed in Table 2. In addition to the definition of taxable income outlined in the sources from Table A.1, we have used tables available from Mutual Information System on Social Protection (MISSOC) to identify which transfers are taxable in each country.[8] For example, with regards to unemployment benefits in Germany, the MISSOC tables indicate that "[b]enefits are not subject to taxation." Hence, we exclude unemployment benefit from the fiscal income for Germany.[9] The MISSOC table describes the taxable portion of pensions in Germany as follows: "The 'taxable share' of a pension for pensions granted before the end of 2005 amounts to 50%. On this basis the tax-free amount of the pension granted to the pension recipient each year is calculated. The taxable share – on which the tax-free amount is determined – for new pensions will be increased each year by 2 percentage points until 2020, and from 2021 to 2040 by 1 percentage point, respectively." Our fiscal income concept constructed using survey data adjusts pension income in Germany, as described above.

---

[7] The income item codes in parenthesis, such as *i11*, refer to specific variables in the LIS Database. See Table 2 for definitions of these income items. For more detail on the availability of a specific variable in each year, see Table A.2.

[8] MISSOC tables can be accessed here: https://www.missoc.org.

[9] MISSOC also states that "[a] part of the benefits (e.g., unemployment benefit, partial unemployment benefit (Teilarbeitslosengeld), short-time working allowance (Kurzarbeitergeld)) is subject to 'progression'". "Progression", here, means that these benefits are themselves tax-free, but their receipt can affect a taxpayer's overall tax rate.

Table 2: Definition of Fiscal Income

Fiscal Income Category	Tax data variable	LIS variable	LIS variable Definition
DE			
Labor Income	Wages and Salaries	pi11	paid employment income
	Pensions	pi32	public contributory pensions
		pi33	private pensions (occupational + individual)
Business Income	Entrepreneurial Income	pi12	self employment income (farm + non-farm)
Capital Income	Interest, Dividend	hi21	interest and dividends income
	Rent	hi22	rental income
FR			
Labor Income	Wages and Salaries	pi11	paid employment income
	Pensions	hi32	public contributory pensions
		hi33	private pensions (occupational + individual)
		hi42	unemployment benefit
		hi43	sickness and work injury pay
Business Income	Entrepreneurial Income	hi12	self employment income (farm + non-farm)
Capital Income	Interest, Dividend	hi21	interest and dividends income
	Rent	hi22	rental income
US			
Labor Income	Wages and Salaries	pi11	paid employment income
	Pensions	pi32	public contributory pension
		pi33	private pension
		pi42	unemployment benefits
Business Income	Entrepreneurial Income	pi12	self employment income (farm + non-farm)
Capital Income	Interest, Dividends	hi21	interest and dividend income
	Rents	hi22	rental income

Source: See Table A.1; LIS Database.

Notes: Fiscal income is defined as pre-tax income, including taxable transfers, and excluding capital gains. Income components are classified depending on how income is earned (i.e., labor, business, or capital). We determined whether transfers in the US are taxable, using US-based sources. We determined whether transfers in Germany and France are taxable, using tables available from MISSOC (https://www.missoc.org/). The letters "p" and "h" in front of income items refer to person-level or household-level variables, respectively; we used person-level data where available. See Table A.2 for a detailed list of variables available, and used, in the LIS Database.

3.4 Adjustments in the matching process

There are several additional challenges to note in the matching process. First, directors' wages are part of the paid employment income variable in the survey data. It is likely, however, that directors' wages are reported as entrepreneurial income in the tax data. Unfortunately, we could not separate out directors' wages from the employment income variable in the LIS Database. So, we include directors' wages as part of labor income.

Second, in our survey data, capital income is not allocated to individuals but to households. Given that capital income is highly concentrated compared with wage income, it is likely that only one unit within a household derived most of the income from capital. In the fiscal income constructed from our survey data, we allocated household-reported capital income only to the Survey-TU that includes the household head. The discrepancy caused by this data challenge, we believe, is minimal.

Third, public contributory pensions, private pensions, unemployment benefits, sickness and work injury pay, and self-employment income are also only available at the household level for France. As above, we assign income from these variables to Survey-TU heads-of-household. Unlike in the previous case, the discrepancy could be significant; it is more likely that a person who is not a household head received unemployment benefits than capital income.

Fourth, realized capital gains are not included in our definition of fiscal income. This income source is only sporadically available in our (or any) survey data, so we omit capital gains from fiscal income in both data sources (in income from tax data, and fiscal income constructed from survey data).

To sum up, the income definition that we use in our comparative exercise can be described as pre-tax income, including taxable transfers, and excluding realized capital gains. In the reminder of the paper, we refer to this as *fiscal income*.

3.5 Calculation of total control income

The total control income is the aggregate annual fiscal income of the entire population based on each data source. In the tax data, it is total fiscal income, including income reported by tax filers and income imputed for non-filers; we take the total directly from the studies listed in Table A.1. In the survey data, we aggregate the constructed fiscal income of the Survey-TUs for the entire distribution. For instance, Germany in 2010 had total fiscal income in the tax data equal to $1.96 trillion, and total fiscal income

constructed from our survey data equal to $1.91 trillion (both in 2013 USD). We use these control totals to calculate the income shares of each fractile.

As a further check, we also use total income obtained from national accounts (NA) as a common income control. This has two benefits over using income totals from the respective databases: first, it accounts for the missing income not accounted for by the fiscal definition of income specified by the tax authorities; second, it removes the differences in the calculated income shares of various income groups between survey and tax data which may be due to differences in the control totals of their respective databases.

Column 1 of Table 3 reports the aggregate income as given in national accounts (NA). Column 2 reports total fiscal income constructed from survey data, and column 4 reports total fiscal income from tax data. In columns 3 and 5, we present, the respective fiscal income shares of total income in NA.[10]

For the United States in 2013, total fiscal income from tax data amounted to $9.1 trillion, which is 59.4 percent of NA income, while total fiscal income constructed from survey data amounted to $9.3 trillion, which is 61.0 percent of the NA amount. Fiscal incomes for the US from both sources are approximately equal in all years. This is also the case for Germany. This suggests that the LIS Database does well to "mimic" very closely the overall fiscal income based on tax sources.

For France, however, there are larger differences in total fiscal income between the survey and tax data. Moreover, it is interesting to note that, on average, fiscal income accounts for only about 45 percent of national income in both data sources. This is significantly less than in the US (average 66 percent in survey data and 63 percent in tax data) and Germany (average 74 percent in survey data and 72 percent in tax data). In this paper, we do not attempt to distribute the entire national income as reflected in the NA, an exercise that has been attracting a growing amount of attention (Alvaredo et al. 2016). The objective of this latter methodological approach is to address the "large and growing gap between the income recorded in the datasets traditionally used to study inequality—household surveys, income tax returns— and the amount of national income recorded in the national accounts (Saez and Zucman 2020)". In this paper, we are simply concerned with comparing the two above-mentioned "datasets traditionally used to study inequality".

---

[10] The source for national income data is the WID database.

Table 3: Comparison of Aggregate (Control) Incomes of the Distribution in DE, FR, & US

	National Accounts	Survey Data		Tax Data		Ratio of fiscal income
	Total National Income (NA)	Total Fiscal Income	% of NA	Total Fiscal Income	% of NA	Survey/Tax
	(1)	(2)	(3)	(4)	(5)	(6)
DE						
1989	3,781	1,599	42.3	2,276	60.2	0.70
2001	3,168	2,682	84.7	2,326	73.4	1.15
2004	2,919	2,387	81.8	2,157	73.9	1.11
2006	2,803	2,215	79.0	2,077	74.1	1.07
2007	2,751	2,179	79.2	2,047	74.4	1.06
2010	2,508	1,913	76.3	1,964	78.3	0.97
FR						
1984	2,602	1,051	40.4	1,165	44.8	0.90
1989	2,635	1,039	39.4	1,126	42.7	0.92
1994	2,453	1,439	58.7	1,044	42.6	1.38
2000	2,636	1,378	52.3	1,072	40.7	1.29
2005	2,514	1,273	50.6	1,105	43.9	1.15
2010	2,339	899	38.4	1,206	51.6	0.75
US						
1979	6,563	4,377	66.7	4,310	65.7	1.02
1986	7,757	5,374	69.3	4,974	64.1	1.08
1991	8,768	5,996	68.4	5,727	65.3	1.05
1994	9,656	6,575	68.1	5,984	62.0	1.10
1997	10,927	7,395	67.7	6,865	62.8	1.08
2000	12,521	8,551	68.3	8,030	64.1	1.06
2004	13,518	8,708	64.4	8,186	60.6	1.06
2007	14,494	9,226	63.7	9,182	63.3	1.00
2010	14,292	8,867	62.0	8,609	60.2	1.03
2013	15,317	9,338	61.0	9,095	59.4	1.03

Source: National accounts income is obtained from the World Inequality Database (WID). Survey-data fiscal income is based on our calculations using the LIS Database. See Table A.1 for tax data sources.

Notes: This table reports total national income from national accounts (NA) and total fiscal income from survey and tax data sources. For the definition of fiscal income, refer to Table 2. Column 6 reports the ratio of survey-based total income (column 2) to tax-based total income (column 4). All income is reported in billions of 2013 USD.

3.6 Definition of top income groups

Once we align the units and fiscal income concepts between the two data sources, our next step is to locate the cut-off percentile(s) in the income distribution where the discrepancy between the two data sources becomes sufficiently persistent. To achieve this, we divide the top decile of the income distribution into three income subgroups: the top income percentile (Top 1%), the next four income

percentiles (Top 5-1%), and the bottom five percentiles of the top income decile (Top 10-5%).[11] Our selection of these three groups is motivated by their widespread use in the growing literature on top income shares. In the remainder of this paper, all references to top income subgroups use this tripartite classification.

4. Results

Our results are presented as follows. First, we report similarities and differences in total and mean incomes of each of the top income groups in the three countries. Then, we discuss similarities and differences in income shares held by each of the top three income groups, also in all three countries. We close by focusing on the US case, where a more detailed disaggregation exercise can be carried out, to assess similarities and differences in the income composition of the top income groups.[12] Here we analyze the role of each income category in driving observed differences of top incomes between tax and survey data.

4.1 Survey data versus tax data: comparison of total and mean incomes of the top income groups

Table 4 presents total and mean fiscal income of each of our three groups for the United States. During our study period, the survey-to-tax ratio of total fiscal income (column 3) for the Top 5-1% is, on average, 1.03, and mean fiscal income ratio (column 6) is 1.01. Likewise, the ratio for the Top 10-5% group is 1.09 and 1.06 respectively. This implies that total and mean fiscal income of the bottom nine percentiles (i.e., the Top 5-1% and the Top 10-5%) of the top decile is slightly higher in survey data than the tax data.

In the US, there are marked differences, however, in both total and mean fiscal incomes within the Top 1% income group. Survey data largely indicate total and mean incomes, for this group, which are lower than those estimated based on tax data. For instance, in 2013, total and mean fiscal incomes in the survey-based data are more than 40 percent lower than in the tax data. A worrying trend is that the difference between the two data sources in total and mean fiscal incomes has progressively become larger for this

---

[11] Top 5-1% group represents the income of the percentiles 96 through 99. Likewise, Top 10-5% group represents income of the percentiles 91 through 95.

[12] The survey data from LIS would allow us to carry out this disaggregation analysis in FR and DE as well. However, a more granular assessment of fiscal income components in the tax-based estimates is only possible for the US. Different income components are available in the WID database but mainly for pre-tax and post-tax national income definitions, not for fiscal income which is the object of analysis in this paper.

income group. Whereas in 1979 the mean fiscal income from survey data accounted for 73 percent of mean fiscal income from tax data, by 2013 the same share had declined to 59 percent.

Table 4: Comparison of Total and Mean Fiscal Incomes of Top Income Groups in US

	US: Total Fiscal Income (in millions USD)			US: Mean Fiscal Income (in USD)		
	Survey (1)	Tax (2)	Survey/Tax (3)	Survey (4)	Tax (5)	Survey/Tax (6)
	Top 1%			**Top 1%**		
1979	90,029	118,585	0.76	88,285	121,679	0.73
1986	179,720	225,263	0.80	153,826	203,520	0.76
1991	249,782	412,166	0.61	197,311	342,179	0.58
1994	399,290	489,064	0.82	311,710	392,143	0.79
1997	503,576	684,846	0.74	379,134	529,651	0.72
2000	623,904	941,949	0.66	447,765	700,474	0.64
2004	669,361	1,042,832	0.64	456,414	724,280	0.63
2007	726,752	1,430,513	0.51	483,004	954,471	0.51
2010	719,920	1,339,124	0.54	465,826	857,495	0.54
2013	871,841	1,504,542	0.58	541,763	923,043	0.59
	Top 5-1%			**Top 5-1%**		
1979	192,821	188,935	1.02	47,271	48,466	0.98
1986	367,199	332,103	1.11	78,574	75,012	1.05
1991	506,428	483,132	1.05	100,011	100,274	1.00
1994	612,502	557,512	1.10	119,539	111,756	1.07
1997	736,073	699,220	1.05	138,544	135,192	1.02
2000	894,810	857,979	1.04	160,548	159,508	1.01
2004	1,013,693	981,170	1.03	172,801	170,363	1.01
2007	1,169,715	1,211,179	0.97	194,350	202,031	0.96
2010	1,217,066	1,249,343	0.97	196,877	200,001	0.98
2013	1,377,602	1,439,530	0.96	214,011	220,790	0.97
	Top 10-5%			**Top 10-5%**		
1979	175,279	170,009	1.03	34,377	34,889	0.99
1986	325,381	295,604	1.10	55,700	53,414	1.04
1991	445,313	404,845	1.10	70,354	67,220	1.05
1994	514,586	460,211	1.12	80,343	73,802	1.09
1997	612,912	550,844	1.11	92,290	85,203	1.08
2000	741,951	662,620	1.12	106,497	98,551	1.08
2004	848,441	761,841	1.11	115,704	105,825	1.09
2007	985,978	922,766	1.07	131,058	123,138	1.06
2010	1,022,811	968,412	1.06	132,363	124,023	1.07
2013	1,138,770	1,088,053	1.05	141,527	133,505	1.06

Source: See Table A.1; LIS Database.

Notes: This table reports total fiscal income (columns 1-3) and mean fiscal income (columns 4-6) of each top income group in the US. Column 3 reports the ratio of survey-based total fiscal income to tax-based total fiscal income, and column 6 does the same for mean fiscal income. All income is reported in current USD.

Tables 5 and 6 present comparisons of total and mean fiscal incomes respectively for Germany and France. As in the US case, total and mean fiscal incomes of Germany are similar whether we use survey or tax data for the bottom nine percentiles (of the top decile). For Germany, the average survey-to-tax ratio of total fiscal income for the Top 5-1% group is, on average, 0.98, while the same ratio for the Top 10-5% group is 1.03. The ratio of mean fiscal incomes for the Top 5-1% group is 1.05 and the ratio for the Top 10-5% group is 1.10. The discrepancy between the two data sources for France is larger for the bottom nine percentiles. The survey-to-tax ratio of total fiscal income is 1.10 for the Top 5-1% group while it is 1.15 for the 10-5% group. Likewise, the ratio of mean fiscal income is 1.22 and 1.27 for the two income groups in France respectively. Thus, again like in the case of the United States, survey data give higher total and mean fiscal incomes than do tax data for nine out of the ten top percentiles.

As in the US case, we find significantly lower total and mean fiscal income in the survey data relative to tax data for the Top 1%. Whereas this difference in Germany is comparable to that in the United States (the survey-to-tax ratio of mean incomes is, on average, 0.65 for the US compared to 0.70 for Germany), the difference in France is much smaller. Indeed, the survey-to-tax ratio of mean fiscal incomes for the Top 1% group in France is, on average, 1.04, indicating higher mean fiscal income in survey data compared to tax data. On the other hand, the survey-to-tax ratio of total fiscal income for France is on average, 0.93. Perhaps a factor contributing to the larger discrepancy in France, is that some key income items from the LIS Database are only available at the household level. For instance, whereas, in Germany and the US, entrepreneurial income is available at the person-level, it is only available for France at the household-level and hence only the head-of-households are assigned the income. Whereas in the US and German cases, two or more tax units within a household would receive their respective income, in the French case only the tax unit associated with the household head would receive the entire income. This could potentially increase the income of the top income groups. The trends in France, similarly, are not as consistent as in the US and German case. One explanation for this year-to-year variation in the trend could be the availability of some income item for France in some years and not others -- for instance, household private pensions (LIS income item *i33*) is not available prior to 2000.[13]

---

[13] For a list of income items available for France in all years, see Table A.2.

Table 5: Comparison of Total Fiscal Incomes of Top Income Groups in DE and FR

| | DE: Total Fiscal Income | | | | FR: Total Fiscal Income | | |
| | Survey | Tax | Survey/Tax | | Survey | Tax | Survey/Tax |
	(1)	(2)	(3)		(4)	(5)	(6)
	Top 1%				**Top 1%**		
1989	69,830	144,967	0.48	1984	40,269	42,751	0.94
2001	147,023	201,099	0.73	1989	49,018	55,956	0.88
2004	141,623	183,136	0.77	1994	65,781	55,837	1.18
2006	146,185	211,756	0.69	2000	61,020	68,356	0.89
2007	148,690	225,330	0.66	2005	61,705	83,030	0.74
2010	125,850	218,981	0.57	2010	90,794	92,972	0.98
	Top 5-1%				**Top 5-1%**		
1989	120,175	164,917	0.73	1984	88,976	76,258	1.17
2001	263,474	246,113	1.07	1989	95,860	89,285	1.07
2004	256,239	243,639	1.05	1994	124,719	95,585	1.30
2006	263,032	255,032	1.03	2000	125,094	110,226	1.13
2007	272,353	263,322	1.03	2005	127,847	125,977	1.01
2010	251,620	268,452	0.94	2010	137,939	151,679	0.91
	Top 10-5%				**Top 10-5%**		
1989	110,387	140,978	0.78	1984	80,784	66,589	1.21
2001	228,056	208,235	1.10	1989	83,917	75,722	1.11
2004	225,750	206,073	1.10	1994	109,571	83,079	1.32
2006	228,810	210,091	1.09	2000	114,131	93,960	1.21
2007	230,784	214,101	1.08	2005	115,129	104,981	1.10
2010	222,912	217,674	1.02	2010	118,180	125,616	0.94

Source: See Table A1; LIS database.

Notes: This table reports the total fiscal income of each top income group for Germany (columns 1-3) and France (columns 4-6). Columns 3 and 6 reports the ratio of survey total fiscal income to tax total fiscal income. All income is reported in millions of 2013 USD.

Table 6: Comparison of Mean Fiscal Incomes of Top Income Groups in DE and FR

| | DE: Mean Fiscal Income | | | | FR: Mean Fiscal Income | | |
	Survey (1)	Tax (2)	Survey/Tax (3)		Survey (4)	Tax (5)	Survey/Tax (6)
		Top 1%				Top 1%	
1989	208,249	421,702	0.49	1984	157,504	173,980	0.91
2001	339,969	429,680	0.79	1989	191,537	204,519	0.94
2004	324,555	395,217	0.82	1994	238,942	185,887	1.29
2006	329,817	441,696	0.75	2000	215,036	207,623	1.04
2007	330,017	466,552	0.71	2005	213,775	233,414	0.92
2010	281,819	445,153	0.63	2010	285,108	251,532	1.13
		Top 5-1%				Top 5-1%	
1989	89,598	119,934	0.75	1984	87,004	77,586	1.12
2001	152,311	131,465	1.16	1989	93,642	81,583	1.15
2004	146,804	131,447	1.12	1994	113,256	79,552	1.42
2006	148,360	132,991	1.12	2000	110,458	83,700	1.32
2007	151,122	136,304	1.11	2005	110,731	88,536	1.25
2010	140,865	136,430	1.03	2010	108,287	102,589	1.06
		Top 10-5%				Top 10-5%	
1989	65,840	82,019	0.80	1984	63,195	54,199	1.17
2001	105,440	88,986	1.18	1989	65,581	55,352	1.18
2004	103,430	88,943	1.16	1994	79,601	55,316	1.44
2006	103,202	87,645	1.18	2000	80,440	57,078	1.41
2007	102,377	88,660	1.15	2005	79,773	59,024	1.35
2010	98,821	88,499	1.12	2010	74,221	67,969	1.09

Source: See Table A.1; LIS Database.

Notes: This table reports mean fiscal income of each top income group for Germany (columns 1-3) and France (columns 4-6). Columns 3 and 6 reports the ratio of survey-based mean fiscal income to tax-based mean fiscal income. All income is reported in 2013 USD.

4.2 Surveys data versus tax data: comparison of income shares of the top income groups

The similarities and differences in fiscal income shares follow the same pattern as with the total and mean fiscal incomes discussed above. Tables 7, 8, and 9 report the income shares of the top income groups for each data source. These tables also report, in columns 3 and 6, the percentage points difference ("gap") in income shares between the two data sources for the US, Germany, and France. We calculate the income shares using two income controls. First, we present results, in columns 1-3, using fiscal income totals

based directly on the respective data sources (referred to as *own fiscal* income control), and next, we present results, in columns 4-6, using a *common* national income control based on national accounts (NA), which is defined as the Gross Domestic Product (GDP) minus capital depreciation plus net income received from abroad. We present the results from the former (i.e., those in columns 1 and 2) in Figures A.1 (for Germany), A.2 (for France), and A.3 (for the US).

<u>United States.</u>

We begin with our results for the US. As expected, in the US, we find large differences in income shares between estimates based on survey versus tax data for the Top 1%, and minimal difference for the bottom nine percentiles of the top income decile. In 2013, for instance, the Top 10-5% group accounts for 12.7 percent of the total fiscal income in the tax data, while it accounts for 12.9 percent of the total fiscal income in the survey data using own fiscal income control, and 7.5 and 7.9 percent respectively using a common NA income control. The average differences in income shares for the years of study of the Top 10-5% group are less than half a percentage point. (Incidentally, in all years, income shares based on the survey data are larger than the income shares in the tax data).

Unlike in the Top 10-5% group, the gap in income shares for the Top 5-1% group has increased in recent years. Whereas the gap in 1979 was 0.1 percentage points using the own income control and 0.2 percentage points using the common income control (both in favor of survey data), by 2013, it has grown to 1.1 percentage points using own income control and 0.4 percentage points using common income control (both in favor of tax data). Despite this trend, the average difference in income shares of the bottom nine percentiles of the top income decile is minimal (0 percentage points using own income control and 0.5 percentage points in favor of survey data using common income control).

As for the Top 1% in the US, not only is there a greater gap in income shares between the two data sources, but this gap has increased over time. Whereas the difference, using the own income controls, between the tax and survey data was 2.0 percentage points in 1979, it increased to 5.1 percentage points in 1991, and further increased to 7.6 percentage points in 2013. Similarly, using a common income control, the gap has increased from 1.3 percentage points in 1979 to 4.4 percentage points in 2013 in favor of income shares calculated from tax data. As acknowledged in the existing empirical literature, the Tax Reform Act (TRA) of 1986, may partially explain this finding: "Series excluding capital gains display a sharp increase from 1986 to 1988 due to the Tax Reform Act of 1986 which resulted [in] (a) a shift from corporate income toward individual business income, [and] (b) a surge in top wage incomes" (as recalled in Atkinson et al.

23

2011, page 30, footnote to Figure 5). In particular, the TRA 1986 created substantial incentives for closely held businesses to shift from corporate to pass-through entities which are taxed at the individual level (e.g., business profits of an S-corporation are passed through the owners each year so that business income falls under the individual income tax).

Similarly, strong incentives remained even for those businesses incorporated as traditional C-corporation (subject to corporate income tax) to shift to the personal tax base by increasing royalty, interest, and rent payments or paying higher wages to entrepreneurs (Gordon and Slemrod 2000; Atkinson et al. 2011). The rise of pass-through businesses in recent decades is unequivocal when "a massive conversion of C-corporations to S-corporations" occurred (Kopczuk and Zwick 2020, page 9). "In 1960, the US economy had about 1 million C-corporations, 1 million partnerships, and almost no S-corporations. By 1980, the number of C-corporations had risen to 2.2 million, the number of partnerships to 1.4 million, and S-corporations had reached about 500,000. But by 2012, the number of C-corporations had declined to 1.6 million, while partnerships had climbed to 3.4 million and S-corporations to 4.2 million." (Kopczuk and Zwick 2020, page 4).

Whereas these changes in the tax code affected income reporting in the tax data, they do not seem to have substantially affected the way households report their income in surveys. The income share of the richest 1 percent of tax units, as computed from survey data, changed much less between 1986 and 1991. What is also interesting to note is that the increasing discrepancy in income shares between survey and tax data after 1986 is limited to the Top 1%. Figure 2 presents the trends in fiscal income shares in the US (using own income controls) between survey and tax data for the Top 1% and the Top 5-1%. Whereas the survey and tax-data income share of the Top 1% diverge after 1986, the trends in income share of the next four percent of top earners do not change after 1986 relative to before.
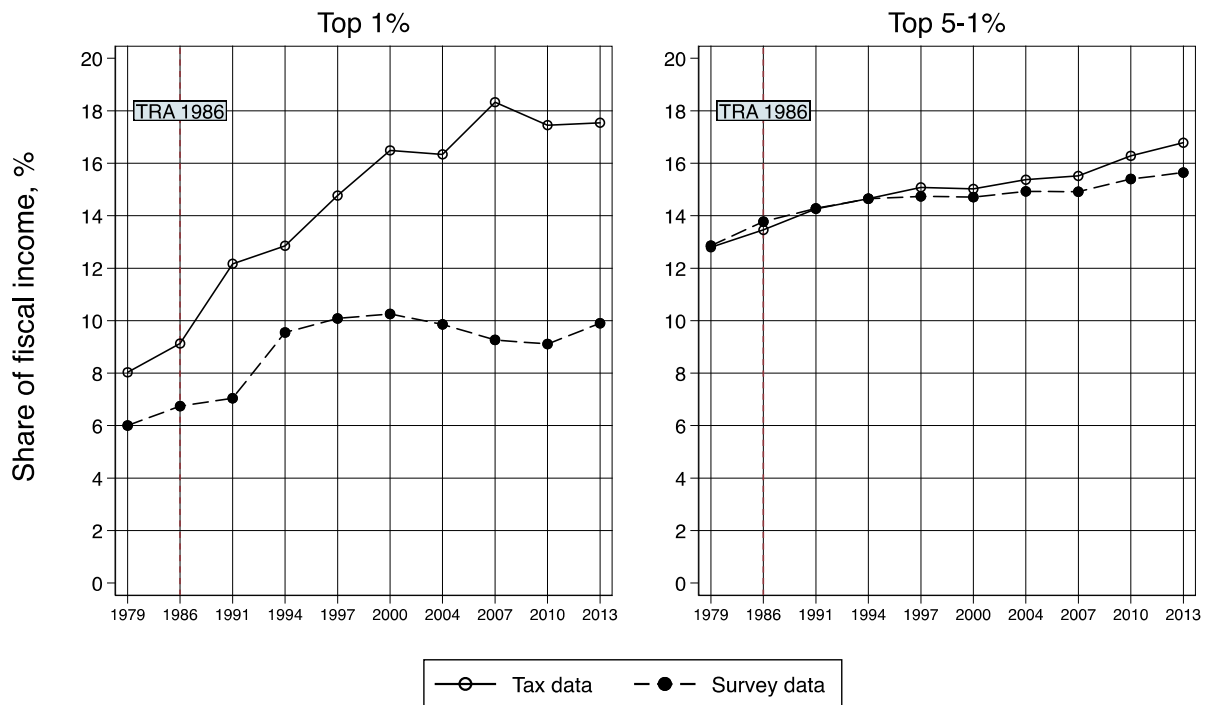
Figure 2: Trends in Fiscal Income Share (in %) in the US

Source: Columns 1 and 2 of Table 7.

Notes: This figure shows the trend in the shares of the fiscal income held by the Top 1% and the next 4 percentile of top earners (Top 5-1%) in the US using the survey and tax data. The vertical line at 1986 highlights the year that the Tax Reform Act of 1986 (TRA 1986) was passed in the US.

This empirical finding is consistent with the understanding that shifts in income reporting resulting from changes in tax laws may not necessarily reflect real changes in economic behavior or the welfare of households. Smith et al. (2019) estimate that approximately 30 percent of the increase in business incomes following the 1986 Tax Reform does not reflect businesses' real economic growth. As echoed in Kopczuk and Zwick (2020), the reorganization of businesses to pass-through "continued through the 1990s and 2000s and even accelerated after the 2001 tax cuts during the Bush administration" (page 10).[14]

---

[14] More recent empirical work by Piketty, Saez, and Zucman (2018) and Auten and Splinter (2019), have suggested alternative ways to go beyond fiscal income data, that is, reconciling income definitions with those found in the National Accounts. Such exercises include, among other things, the full allocation of C-corporations' retained earnings and corporate taxes to individuals, in order to remove the bias introduced by the incentives to a pass-

In the same period during which changes in US tax law affected the share of reported fiscal income captured at the very top of the income distribution, the capacity of survey data to accurately capture the upper income brackets declined. Indeed, as documented by Morelli and Munoz (2020), unit non-response rates have been rapidly increasing in the Current Population Survey (CPS) data (the source of the US dataset in the LIS Database). The unit non-response rate was slightly below 4 percent in 1977 and rose to above 8 percent around 2007 and then above 14 percent by 2018. This has likely affected the representativeness of the top of the income distribution disproportionately, because non-response rates are positively associated with income level (Korinek, Mistiaen, and Ravallion 2006, and 2007; and Hlasny, 2020).

We thus draw the following conclusions. The growing discrepancy between estimates based on tax data versus survey data may be due, on the one hand, to the simple reclassification of income flows from corporate to personal (which may lead to overstating top incomes in the tax data) and, on the other hand, to the declining capacity of household survey data to capture income at the very top of the income distribution. We are unable to estimate the relative importance of each factor.

Germany

Next, we turn our attention to Germany. The results for Germany are similar to those for the US. Again, the bottom nine percentiles (of the top decile) display minimal differences in fiscal income shares between the two data sources. On average, income shares in the survey data are 0.1 percentage points lower using own denominator and 0.1 percentage points higher using the common national income denominator. The shares of the Top 5-1% group are, on average, 0.4 percentage points higher in the tax data using the own denominator and 0.1 percentage points higher using the common denominator. However, there is a larger gap at the top of the distribution. The average gap for the Top 1% is 4.0 percentage points using the own denominator and 2.8 percentage points using the common denominator. What is different from the US case, however, is that, in Germany, the discrepancy in income shares, between the two sources has been relatively stable over time.

---

through reorganizing of businesses. Using this approach (e.g., the distributional national accounts approach), all corporate income is allocated to shareholders. In the work of Piketty, Saez, and Zucman (2018), the portion of corporate profit that is not paid out as dividends (undistributed profits) is distributed, in equal 50 percent shares, proportionally to dividends and realized capital gains. It is worth noting that such adjustments do not affect the accuracy of the increase in top income shares (based on pre-tax national income) from 1980 to later years.

Finally, we take a close look at France. The difference in income shares between the tax and survey data of all the top income groups in France is relatively small, but it fluctuates from year-to-year. On average, the gap is 1 percentage point for the bottom nine percent of the top income decile using our own income control and 0.6 percentage points using the common income control (both in favor of survey data). In 2010, the survey-based shares are higher for all top income groups by more than 2.5 percentage points using our own income denominator, but less than 0.6 percentage points higher using the common income denominator. As discussed above, the fluctuation in the trends of the differences between the survey and tax data series in France could be due to the difficulty in matching the income concepts in France as compared with the US and German cases.

In conclusion, using the common national income denominator, the Top 1% income share for the most recent period is lower, using survey data, by 4.4 percentage points in the United States, 3.9 percentage points in Germany, and 0.1 percentage point in France. While in Germany and France there is no clear indication of an increase in this gap, in the case of the US, the gap has risen almost continuously from a little over 1 percentage point in the 1980s, to more than double that in 2013. This is evident when comparing Figures A.1 and A.2, for Germany and France respectively, with Figure A.3, for the US. We next disaggregate the observed gap in the US by income components.

Table 7: Comparison of Income Shares of Top Income Groups in US

	Own Income Denominator			Common Income (NA) Denominator		
	Tax (%) (1)	Survey (%) (2)	Tax - Survey (pp) (3)	Tax (%) (4)	Survey (%) (5)	Tax - Survey (pp) (6)
		Top 1%			Top 1%	
1979	8.0	6.0	2.0	5.3	4.0	1.3
1986	9.1	6.7	2.4	5.9	4.7	1.2
1991	12.2	7.0	5.1	7.9	4.8	3.1
1994	12.9	9.5	3.3	8.0	6.5	1.5
1997	14.8	10.1	4.7	9.3	6.8	2.5
2000	16.5	10.3	6.2	10.6	7.0	3.6
2004	16.3	9.9	6.5	9.9	6.4	3.5
2007	18.3	9.3	9.1	11.6	5.9	5.7
2010	17.5	9.1	8.3	10.5	5.7	4.9
2013	17.5	9.9	7.6	10.4	6.0	4.4
		Top 5-1%			Top 5-1%	
1979	12.8	12.9	-0.1	8.4	8.6	-0.2
1986	13.5	13.8	-0.3	8.6	9.5	-0.9
1991	14.3	14.3	0.0	9.3	9.8	-0.4
1994	14.7	14.6	0.0	9.1	10.0	-0.9
1997	15.1	14.7	0.3	9.5	10.0	-0.5
2000	15.0	14.7	0.3	9.6	10.0	-0.4
2004	15.4	14.9	0.4	9.3	9.6	-0.3
2007	15.5	14.9	0.6	9.8	9.5	0.3
2010	16.3	15.4	0.9	9.8	9.6	0.3
2013	16.8	15.6	1.1	10.0	9.5	0.4
		Top 10-5%			Top 10-5%	
1979	11.5	11.7	-0.2	7.6	7.8	-0.2
1986	12.0	12.2	-0.2	7.7	8.5	-0.8
1991	12.0	12.6	-0.6	7.8	8.6	-0.8
1994	12.1	12.3	-0.2	7.5	8.4	-0.9
1997	11.9	12.3	-0.4	7.5	8.3	-0.8
2000	11.6	12.2	-0.6	7.4	8.3	-0.9
2004	11.9	12.5	-0.6	7.2	8.0	-0.8
2007	11.8	12.6	-0.7	7.5	8.0	-0.5
2010	12.6	12.9	-0.3	7.6	8.0	-0.4
2013	12.7	12.9	-0.2	7.5	7.9	-0.4

Source: See Table A.1; LIS Database.

Notes: Income shares in this table are calculated using either our own income control or a common income control. Here, our own income control refers to total fiscal income reported in columns 1 and 2 of Table 4 for survey and tax data respectively. The common income control refers to income from national accounts reported in column 1 of Table 3. Columns 3 and 6 of this table report the gap in percentage points between income shares from tax and survey data.

Table 8: Comparison of Income Shares of Top Income Groups in Germany

	Own Income Denominator			Common Income (NA) Denominator		
	Tax (%) (1)	Survey (%) (2)	Tax - Survey (pp) (3)	Tax (%) (4)	Survey (%) (5)	Tax - Survey (pp) (6)
	Top 1%			Top 1%		
1989	10.5	7.2	3.3	6.3	3.1	3.3
2001	11.0	7.0	4.0	8.1	5.9	2.2
2004	10.1	7.1	3.1	7.5	5.8	1.7
2006	11.5	7.4	4.0	8.5	5.9	2.6
2007	12.0	7.5	4.6	9.0	5.9	3.0
2010	11.7	6.9	4.8	9.2	5.3	3.9
	Top 5-1%			Top 5-1%		
1989	12.0	12.4	-0.4	7.2	5.3	2.0
2001	13.5	12.5	1.0	9.9	10.6	-0.7
2004	13.5	12.8	0.7	10.0	10.5	-0.5
2006	13.8	13.3	0.5	10.2	10.5	-0.3
2007	14.1	13.7	0.4	10.5	10.8	-0.4
2010	14.4	13.8	0.5	11.3	10.6	0.7
	Top 10-5%			Top 10-5%		
1989	10.2	11.4	-1.2	6.2	4.8	1.3
2001	11.4	10.8	0.6	8.4	9.1	-0.8
2004	11.4	11.3	0.1	8.4	9.2	-0.8
2006	11.4	11.6	-0.2	8.4	9.2	-0.7
2007	11.4	11.6	-0.1	8.5	9.2	-0.7
2010	11.7	12.1	-0.5	9.1	9.3	-0.1

Source: See Table A.1; LIS Database.

Notes: Income shares in this table are calculated using either our own income control or a common income control. Here, our own income control refers to total fiscal incomes reported in columns 1 and 2 of Table 5 for survey and tax data respectively. The common income control refers to income from national accounts reported in column 1 of Table 3. Columns 3 and 6 of this table reports the difference in percentage points between income shares from tax and survey data.

Table 9: Comparison of Income Shares of Top Income Groups in France

	Own Income Denominator			Common Income (NA) Denominator		
	Tax (%) (1)	Survey (%) (2)	Tax - Survey (pp) (3)	Tax (%) (4)	Survey (%) (5)	Tax - Survey (pp) (6)
		Top 1%			Top 1%	
1984	7.0	7.3	-0.3	3.1	3.0	0.2
1989	8.2	7.8	0.4	3.5	3.1	0.4
1994	7.7	6.6	1.1	3.3	3.9	-0.6
2000	8.3	5.8	2.5	3.4	3.0	0.4
2005	8.7	5.6	3.1	3.8	2.8	1.0
2010	8.1	10.6	-2.5	4.2	4.1	0.1
		Top 5-1%			Top 5-1%	
1984	12.5	16.2	-3.7	5.6	6.6	-0.9
1989	13.1	15.2	-2.1	5.6	6.0	-0.4
1994	13.2	12.5	0.7	5.6	7.3	-1.7
2000	13.4	11.8	1.6	5.4	6.2	-0.7
2005	13.2	11.6	1.6	5.8	5.9	-0.1
2010	13.2	16.1	-2.9	6.8	6.2	0.6
		Top 10-5%			Top 10-5%	
1984	11.0	14.7	-3.8	4.9	5.9	-1.0
1989	11.1	13.3	-2.2	4.7	5.3	-0.5
1994	11.5	11.0	0.5	4.9	6.4	-1.6
2000	11.4	10.8	0.6	4.6	5.6	-1.0
2005	11.0	10.5	0.5	4.8	5.3	-0.5
2010	11.0	13.8	-2.9	5.6	5.3	0.3

Source: See Table A1; LIS Database.

Notes: Income shares in this table are calculated using either our own income control or a common income control. Here, our own income control refers to the total fiscal incomes reported in columns 4 and 5 of Table 5 for survey and tax data respectively. The common income control refers to income from national accounts reported in column 1 of Table 3. Columns 3 and 6 of this table reports the difference in percentage points between income shares from tax and survey data.

4.3 Income composition of top income groups in the US

Labor income forms a major share of income for most parts of the income distribution. However, as we move higher up within the top decile, non-labor income becomes more relevant. As the share of the non-labor component of income increases, so does (as we shall see below) the discrepancies between the two types of data. To understand this more fully, we divide the total fiscal income of each income group into three income components: labor, business, and capital.

Table 10 disaggregates total fiscal income for the top income groups in the US into the three core income components. We find that, on average, 90.1 percent of the total income of the Top 10-5% group is derived from labor in the tax data, and 88.3 percent is derived from labor in the survey data. This percentage decreases slightly when we move up to the Top 5-1% group. For this group, on average, the share from labor is 82.5 percent in the survey data and 82.2 percent in the tax data. The differences in income composition between the two data sources for both the Top 10-5% and Top 5-1% groups are minimal.

Where the non-labor component, and by symmetry the labor component, is vastly different between the two data sources is within the top percentile. Here, we find 78.7 percent of fiscal income to be composed of labor income in the survey data, compared to only 59.0 percent in the tax data. This also implies that we find larger differences (relative to the two other subgroups) in the non-labor components. Business income accounts for 12.1 percent in survey data and 24.6 percent in tax data, while capital income accounts for 9.2 percent in the survey data and 16.4 percent in the tax data.

Figure 3 compares the composition of income of the top three income groups for the year 2013. The composition of income is approximately equal for the bottom nine percentiles of the top decile. The results for the bottom nine percentiles of the top decile are striking in that (i) the tax units are similar, (ii) the mean and total incomes of these groups are similar, (iii) the income shares held by these groups are similar, and (iv) the composition of income is also similar -- whether we use fiscal income derived from tax data or from survey data. This is an important finding.

There are, however, large discrepancies in the Top 1% group, and these arise because the non-labor component constitutes a larger share in the tax data compared to the survey data. Figure 4 presents the income components for the Top 1% group in 1986 and in 2013. It is interesting to note that the shares of these income components in the Top 1% group differ less between the two data sources in 1986 than in 2013. In 2013, the differences due to non-labor income, and in particular business income, is much larger. If we take business and capital income together and look at its share against labor, the tax data show an approximately equal division between the two. If we do the same using our survey data, we note an increasing share of labor income because both capital income, and even more so, business income shrink.

In Figure A.7, we present the ratio of each component of survey total fiscal income over the corresponding component of tax total fiscal income, for each of the top income groups. A ratio of 1 would mean that the corresponding top income groups in the two data sources hold equal income from a particular income component. We find that the ratio of total fiscal labor income is close to 1 for all top income groups, and

31

importantly, there are no increasing or decreasing trends. This is, however, not the case for the non-labor income components. In recent years, we find that the survey data capture more capital income but less business income for the bottom 9 percentiles of the top decile. For the Top 1%, the tax data capture more of both capital and business incomes. What is interesting, and consistent with the TRA 1986 explanation above, is that the business income significantly increases in tax data compared to survey data in the years after 1986. While the ratio of business income for the Top 1% is 1.0 and 1.1 in 1979 and 1986 respectively, we find that the ratio falls to 0.3 in 1991. The ratio of business income has stayed at that level or further decreased since. TRA 1986 may explain this finding, as a growing share of total income at the very top is recorded as pass-through personal income.

Table 10: Comparison of Income Composition of Top Income Groups in US

	Survey Data			Tax Data		
	Labor (1)	Business (2)	Capital (3)	Labor (4)	Business (5)	Capital (6)
Top 1%						
1979	66.1	22.9	11.0	59.0	17.0	24.0
1986	73.1	14.9	12.0	65.7	11.1	23.1
1991	71.3	12.7	16.0	57.4	23.0	19.7
1994	81.6	10.0	8.4	59.1	26.8	14.1
1997	76.5	12.5	10.9	60.3	26.7	13.0
2000	79.5	12.8	7.7	63.0	24.7	12.3
2004	84.8	8.9	6.3	58.4	28.4	13.2
2007	82.9	10.0	7.1	54.3	27.7	18.1
2010	85.3	9.5	5.2	57.2	29.2	13.6
2013	85.5	7.0	7.5	55.8	31.3	12.9
Top 5-1%						
1979	81.6	12.2	6.2	78.4	11.5	10.1
1986	83.2	8.8	7.9	83.3	7.3	9.4
1991	79.6	10.6	9.8	80.8	10.2	9.0
1994	82.7	8.3	8.9	82.9	10.8	6.3
1997	79.2	7.9	12.8	82.0	11.0	7.1
2000	80.6	8.5	10.9	82.3	11.0	6.7
2004	84.0	7.3	8.8	83.2	11.5	5.3
2007	82.4	6.9	10.7	80.1	12.0	7.9
2010	86.5	5.9	7.6	85.2	10.1	4.7
2013	85.4	5.4	9.3	83.4	11.7	5.0
Top 10-5%						
1979	88.7	6.7	4.6	89.1	5.2	5.7
1986	87.7	6.1	6.2	90.1	3.8	6.0
1991	86.9	6.7	6.5	89.4	4.7	5.9
1994	89.2	5.1	5.7	91.1	5.0	3.9
1997	86.4	5.6	8.0	91.0	4.8	4.2
2000	87.4	5.5	7.0	89.7	5.6	4.7
2004	89.5	4.8	5.7	89.4	6.8	3.8
2007	87.4	5.4	7.2	88.6	6.0	5.4
2010	89.6	4.8	5.6	91.2	5.5	3.3
2013	90.3	3.8	5.9	91.2	5.7	3.1

Source: See Table A.1; LIS Database.

Notes: This table reports the composition of fiscal income from labor, business, and capital components for the US. For a definition of each fiscal income component, refer to Table 2.

Table reading: Columns 1-3 in each row add up to 100 percent; this represents a breakdown of the total survey-based fiscal income reported in column 1 of Table 4. Likewise, columns 4-6 represent a breakdown of total tax-based fiscal income reported in column 2 of Table 4.
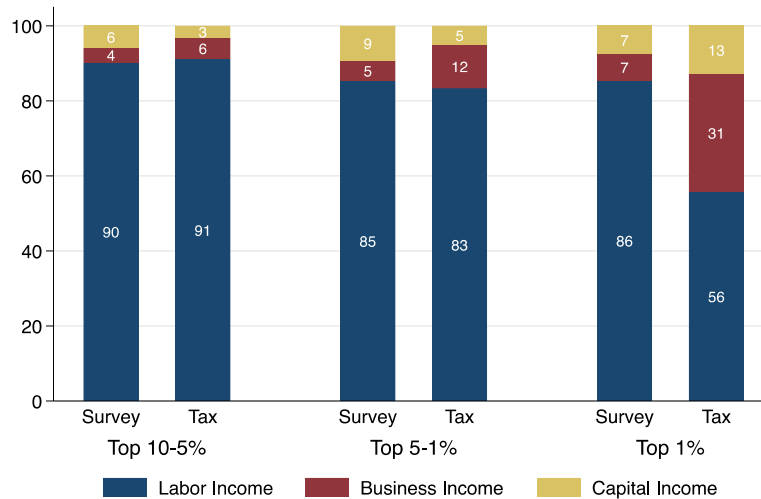
Figure 3: Comparison of Fiscal Income by Income Component for the Top Income Groups in US in 2013

Source: See Table A.1; LIS Database.

Notes: This figure shows the disaggregation of total fiscal income (in percent) by income components (labor, business, and capital) held by each income group for the year 2013. Income groups represent the top percentile (Top 1%), the next four percentiles (Top 5-1%), and the bottom five percentiles of the top decile (Top 10-5%).
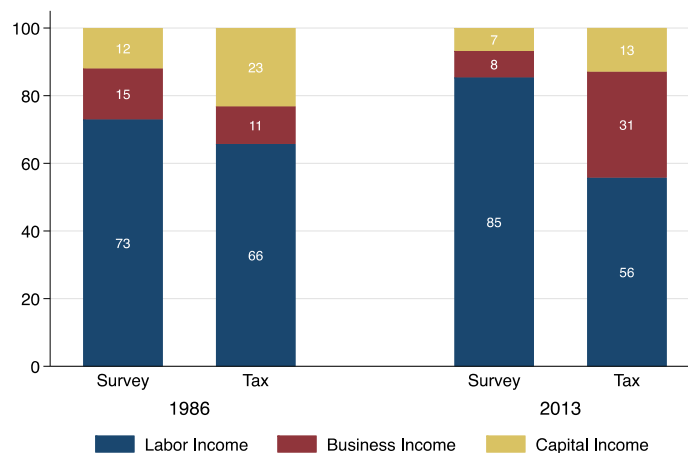

Figure 4: Comparison of Fiscal Income by Income Component for the Top 1% in US, 1986 and 2013

Source: See Table A.1; LIS Database.

Notes: This figure shows the disaggregation of total fiscal income (in percent) by income components (labor, business, and capital) held by the Top 1% group in 1986 and 2013. For other years, see Figures A.4, A.5, and A.6.

4.4 Survey data versus tax data: comparison of income composition of top groups in the US

In Table 11, we further disaggregate the gap in income shares for the top income groups in the US into the labor, business, and capital income components. Our objective is to find out which income source, or sources, are responsible for the overall discrepancy in the income shares of the Top 1%. Column 1 of this table is equivalent to column 3 in Table 7, i.e., the difference in income shares between tax and survey data derived using the own income control. We find that the gap for the Top 1% income group is driven by the non-labor components of income. In 2013, for instance, of the difference in the Top 1% share of 7.6 percentage points, less than one-fifth was due to labor income, while more than four-fifths was due to business and capital incomes combined. While the percentage point gap due to labor and capital income have stayed relatively stable during our period of study, the gap due to business income has dramatically increased. The gap in business income has been growing continuously for the past three decades. Recently, the gap in business income has accounted for more than half of the gap in total income shares.

Given that a substantial portion of pass-through income has been reclassified for tax purposes, this may be an overestimation of the true role of business income as a driver of the observed gap between tax and survey data at the very top. Respondents in household surveys did not have an incentive to reclassify this income source and we believe they continued to treat it as corporate income or, less likely, to include it in labor income.

Table 11: Difference between Tax and Survey Data Income Shares by Income Component, US

| | Tax - Survey (pp) (Own Income Denominator) | Tax - Survey by Income Categories (pp) | | |
| | | Labor | Business | Capital |
	(1)	(2)	(3)	(4)
Top 1%				
1979	2.0	0.8	0.0	1.3
1986	2.4	1.1	0.0	1.3
1991	5.1	2.0	1.9	1.3
1994	3.3	-0.2	2.5	1.0
1997	4.7	1.2	2.7	0.8
2000	6.2	2.2	2.8	1.2
2004	6.5	1.2	3.8	1.5
2007	9.1	2.3	4.1	2.7
2010	8.3	2.2	4.2	1.9
2013	7.6	1.3	4.8	1.5
Top 5-1%				
1979	-0.1	-0.5	-0.1	0.5
1986	-0.3	-0.3	-0.2	0.2
1991	0.0	0.2	-0.1	-0.1
1994	0.0	0.0	0.4	-0.4
1997	0.3	0.7	0.5	-0.8
2000	0.3	0.5	0.4	-0.6
2004	0.4	0.3	0.7	-0.5
2007	0.6	0.1	0.8	-0.4
2010	0.9	0.6	0.7	-0.4
2013	1.1	0.6	1.1	-0.6
Top 10-5%				
1979	-0.2	-0.1	-0.2	0.1
1986	-0.2	0.1	-0.3	0.0
1991	-0.6	-0.2	-0.3	-0.1
1994	-0.2	0.0	0.0	-0.2
1997	-0.4	0.2	-0.1	-0.5
2000	-0.6	-0.3	0.0	-0.3
2004	-0.6	-0.5	0.2	-0.3
2007	-0.7	-0.5	0.0	-0.3
2010	-0.3	-0.1	0.1	-0.3
2013	-0.2	-0.1	0.2	-0.4

Source: See Table A.1; LIS Database.

Notes: This table reports the composition of the "gap" between fiscal incomes shares from tax and survey data. Column 1 is equivalent to column 3 of Table 7. Columns 2-4 present the values in column 1, broken out by income component. See also Table 7.

5. Conclusion

It is widely believed that household survey data do not accurately capture top incomes. Yet, this statement is rarely verified. We investigate this crucial issue by comparing tax data, based on the studies listed in Table A.1, to household survey data, available from the LIS Database, for France, Germany, and the US. Exploiting the flexibility of survey data allows us to align (define the same way) both income categories and units of analysis. We can thus construct, in the survey data, fiscal income equivalent to what is reported in the tax data and reconfigure household survey units into tax units.

Our main results confirm that average fiscal incomes of the richest percentile are in some cases substantially lower in the survey data than as estimated by tax data. In Germany and the US, this occurs almost entirely above the 99th percentile. In the US, in 2013, the top income share of the richest percentile, estimated using survey data, is 7.6 percentage points less than what is observed in tax data. In Germany, in 2010, that difference is 4.8 percentage points. In France, in 2010, in contrast, the difference is in the opposite direction (the share is higher in the survey data), by 2.5 percentage points. The observed income gap is increasing over time in the case of the US, while it is largely stable in Germany, and is volatile in France. About three-fourths of the survey-versus-tax-data gap in the US is attributable to the non-labor portion of income. Half of the gap is due to business income, and one-quarter to capital income.

The findings suggest that the observed growing gap in the top percentile group in the US may be driven by tax changes that provide incentives to shift corporate income into the personal income. We cannot, however, quantify the extent to which US tax changes are responsible for the growing gap.

Several other findings may be mentioned. First, we find that the estimated annual aggregate fiscal incomes based on the two data sources are more or less equal. Both account for about 65 percent of national income in the US, about 70 percent in Germany, and about 45 percent in France. Second, the overall fiscal income of the income groups above the 90th percentile and below the richest percentile, align well across the two sources, for all three countries and in all periods under investigation. Third, mean labor incomes in the Top 5-1% and Top 10-5% income groups are also similar in the two types of data. There is also little difference in the mean non-labor income of these two income groups.

Our results indicate that differences between household survey data and tax data regarding the top 1 percent cannot be generalized across national contexts, implying that thorough country-specific

assessments remain indispensable. On a more positive note, for 99 percent of national distributions, we find no systemic or significant differences between the two sources.

Reference:

Alvaredo, Facundo, and Thomas Piketty. 2015. "France Preliminary Estimates of Top Income Shares 2006-2012: Note on Methods." World Inequality Lab.

Alvaredo, Facundo, Anthony Atkinson, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2016. "Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in WID.world." WID.world working paper No 2016/2.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2018. *World Inequality Report 2018*. Belknap Press of Harvard University Press.

Atkinson, Anthony Barnes, and Thomas Piketty. 2007. *Top Incomes over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries.* OUP Oxford.

_____. 2010. *Top Incomes: A Global Perspective*. Oxford University Press.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature*, 49 (1): 3-71.

Auten, Gerald, and David Splinter. 2019. "Top 1 Percent Income Shares: Comparing Estimates Using Tax Data." In *AEA Papers and Proceedings*, 109:307–11.

Bartels, Charlotte. 2018. "Top Incomes in Germany, 1871-2014." IZA Discussion Paper.

Bartels, Charlotte, and Katharina Jenderny. 2015. "The Role of Capital Income for Top Incomes Shares in Germany." halshs-02654332. https://halshs.archives-ouvertes.fr/halshs-02654332/document.

Bartels, Charlotte, and Maria Metzing. 2019. "An Integrated Approach for a Top-Corrected Income Distribution." *The Journal of Economic Inequality* 17 (2). Springer: 125–43.

Blanchet, Thomas, Ignacio Flores, and Marc Morgan. 2018. "The Weight of the Rich: Improving Surveys Using Tax Data." *WID. World Working Paper Series*, no. 2018/12.

Bourguignon, François. 2018. "Simple Adjustments of Observed Distributions for Missing Income and Missing People." *The Journal of Economic Inequality* 16 (2). Springer: 171–88.

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus. 2016. "Measuring Income and Wealth at the Top Using Administrative and Survey Data." *Brookings Papers on Economic Activity* 2016 (1). Brookings Institution Press: 261–331.

Burkhauser, Richard V., Shuaizhang Feng, Stephen P Jenkins, and Jeff Larrimore. 2012. "Recent Trends in Top Income Shares in the United States: Reconciling Estimates from March CPS and IRS Tax Return Data." *Review of Economics and Statistics* 94 (2). MIT Press: 371–88.

Burkhauser, Richard V., Nicolas Hérault, Stephen P. Jenkins, and Roger Wilkins. 2017. "Top Incomes and Inequality in the UK: Reconciling Estimates from Household Survey and Tax Return Data." *Oxford Economic Papers* 70 (2). Oxford University Press: 301–26.

Canberra Group. 2011. *Canberra Group Handbook on Household Income Statistics*. 2nd Edition. Geneva: United Nations.

Dell, Fabien. 2007. "Top Incomes in Germany Throughout the Twentieth Century: 1891–1998." *Top Incomes over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. Oxford University Press Oxford, 365–425.

_____. 2011. "Update and Extension of Germany Series." WID.world Technical Notes Series No. 2011/1. https://wid.world/document/wid_methodology_notes_2011_1_germany/.

Gordon, Roger H., and Joel B. Slemrod. 2000. "Are 'Real' Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?" *In Does Atlas Shrug? The Economic Consequences of Taxing the Rich,* Ed. Joel B. Slemrod, 240–80. New York: Russell Sage Foundation; Cambridge; London: Harvard University Press.

Hlasny, Vladimir 2020. "Nonresponse Bias in Inequality Measurement: Cross-Country Analysis Using Luxembourg Income Study Surveys." *Social Science Quarterly*.

Hlasny, Vladimir, and Paolo Verme. 2018. "Top Incomes and Inequality Measurement: A Comparative Analysis of Correction Methods Using the EU SILC Data." *Econometrics* 6 (2). Multidisciplinary Digital Publishing Institute: 30.

Kopczuk, Wojciech, and Eric Zwick. 2020. "Business Incomes at the Top." National Bureau of Economic Research (No. W27752).

Korinek, A., J. Mistiaen, and M. Ravallion. 2006. "Survey Nonresponse and the Distribution of Income." *Journal of Economic Inequality* 4(1): 23.

Korinek, A., J. A. Mistiaen, and M. Ravallion. 2007. "An Econometric Method of Correcting for Unit Nonresponse Bias in Surveys." *Journal of Econometrics* 136: 213–235.

Kuznets, Simon. 1955. "Economic Growth and Economic Inequality." *American Economic Review*, XLV (1955), 1-28.

Landais, Camille. 2008. "Top Incomes in France: Booming Inequalities?" *Paris School of Economics, Mimeo*.

Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org [lisdatacenter.org] (multiple countries; 10/17/2020 – 10/20/2020). Luxembourg: LIS.

Lustig, Nora. 2020. "The Missing Rich in Household Surveys: Causes and Correction Approaches." *Paris School of Economics, Mimeo*. https://doi.org/10.31235/osf.io/j23pn.

Morelli, Salvatore, and Ercio Munoz. 2020. "Unit Nonresponse Bias in the Current Population Survey." Unpublished manuscript, City University of New York.

Piketty, Thomas. 2001. "Income Inequality in France 1901-98." CEPR Discussion Paper.

_____. 2003. "Income Inequality in France, 1901–1998." *Journal of Political Economy* 111 (5). The University of Chicago Press: 1004–42.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *The Quarterly Journal of Economics* 118 (1). MIT Press: 1–41.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *The Quarterly Journal of Economics* 133 (2). Oxford University Press: 553–609.

Saez, Emmanuel. 2015. "Striking it Richer: The Evolution of Top Incomes in the United States (Updated with 2013 preliminary estimates)." WID.world Technical Note Series No. 2015/1. https://wid.world/document/wid_methodology_notes_2015_1_us/.

Saez, Emmanuel, and Gabriel Zucman. 2020. "The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts". *Journal of Economic Perspectives*—Volume 34, Number 4—Fall 2020—Pages 3–26.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *The Quarterly Journal of Economics* 134 (4). Oxford University Press: 1675–1745.

A. Appendix

Table A.1: Sources for Tax Data

Country	Tax Data Sources	Years
DE	Dell (2007)	1989
	Dell (2011)	2001, 2002, 2004, 2006, 2010
	Bartels & Jendery (2015)	2010
FR	Piketty (2003)	1984, 1989, 1994
	Landais (2008)	2000, 2005
	Alvaredo & Piketty (2015)	2010
US	Piketty & Saez (2003)	1979, 1986, 1991, 1994, 1997
	Saez (2015)	2004, 2007, 2010, 2013

Note: This table reports the sources for fiscal income from tax data used in this study.

Table A.2: Variables used from LIS database

Variable Name	Description	Germany						France						United States									
		1989	2001	2004	2006	2007	2010	1984	1989	1994	2000	2005	2010	1979	1986	1991	1994	1997	2000	2004	2007	2010	2013
pi11	Wage Income, person	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
pi32	Public Contributory Pensions, person	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
hi32	Public Contributory Pensions, household							x	x	x	x	x	x										
pi33	Private Pensions, person	x	x	x	x	x	x							x	x	x	x	x	x	x	x	x	x
hi33	Private Pensions, household										x	x	x										
pi42	Unemployment Benefits, person	x	x	x	x	x	x							x	x	x	x	x	x	x	x	x	x
hi42	Unemployment Benefits, household							x	x	x	x	x	x										
hi43	Sickness and Work Injury Pay, household									x	x	x	x										
pi12	Self-Employment Income, person	x	x	x	x	x	x							x	x	x	x	x	x	x	x	x	x
hi12	Self-Employment Income, household							x	x	x	x	x	x										
hicapital	Capital Income, household							x	x	x													
hi21	Interest and Dividends, household	x	x	x	x	x	x				x	x	x	x	x	x	x	x	x	x	x	x	x
hi22	Rental Income, household	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Source: METIS, downloaded from the LIS website, 10/20/2020. Source of LIS surveys - Germany (all years): German Socio Economic Panel Survey; France (all years): Household Budget Survey; United States (all years): Current Population Survey - March Supplement.
Notes: This table reports the variables used to construct fiscal income using the survey data contained in the LIS Database. An "x" indicates the variables in each LIS dataset that we used to construct fiscal income. Also see Table 2.
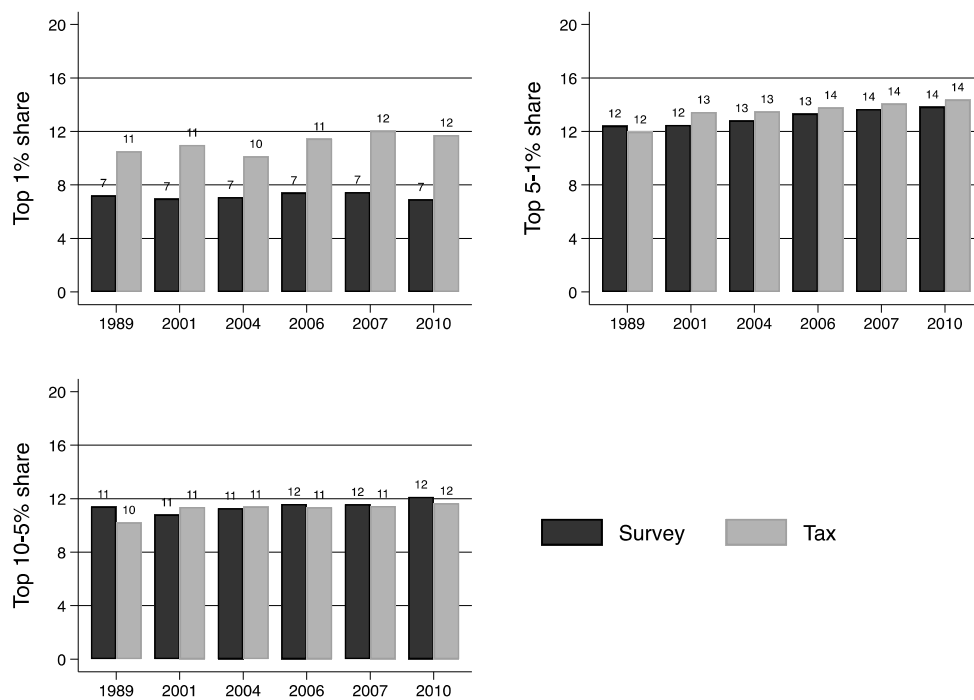
Figure A.1: Fiscal Income Share (in %) of Top Income Groups in Germany

Source: See Tables 7, 8, and 9.

Notes: This figure compares the fiscal income shares of the top income groups in Germany using the survey-data and tax-data fiscal income series. The incomes shares reported here are equivalent to those presented in columns 1 and 2 of Tables 7, 8, and 9.
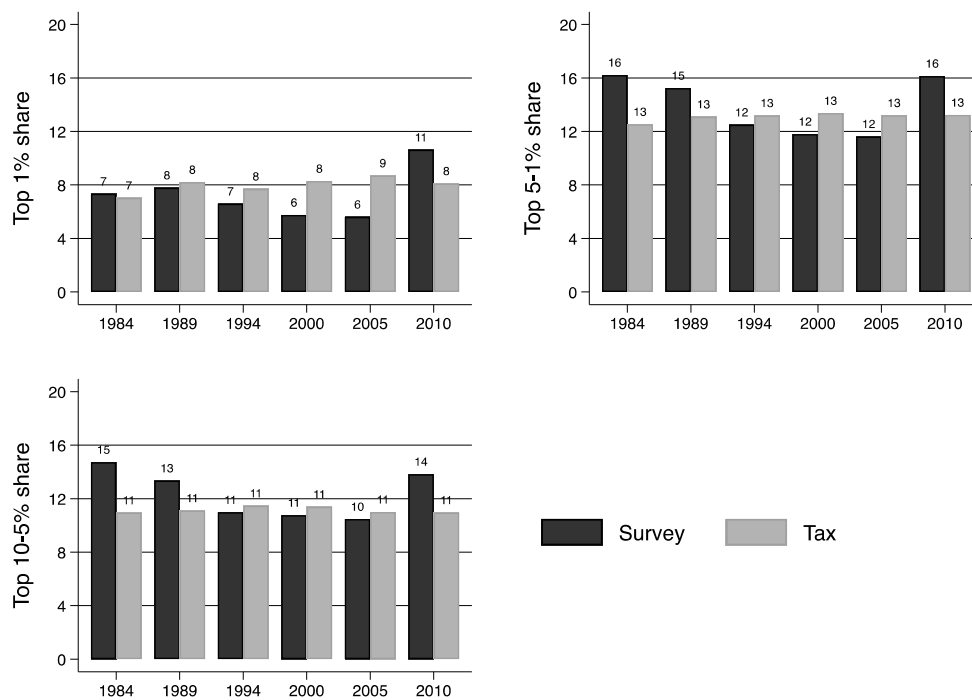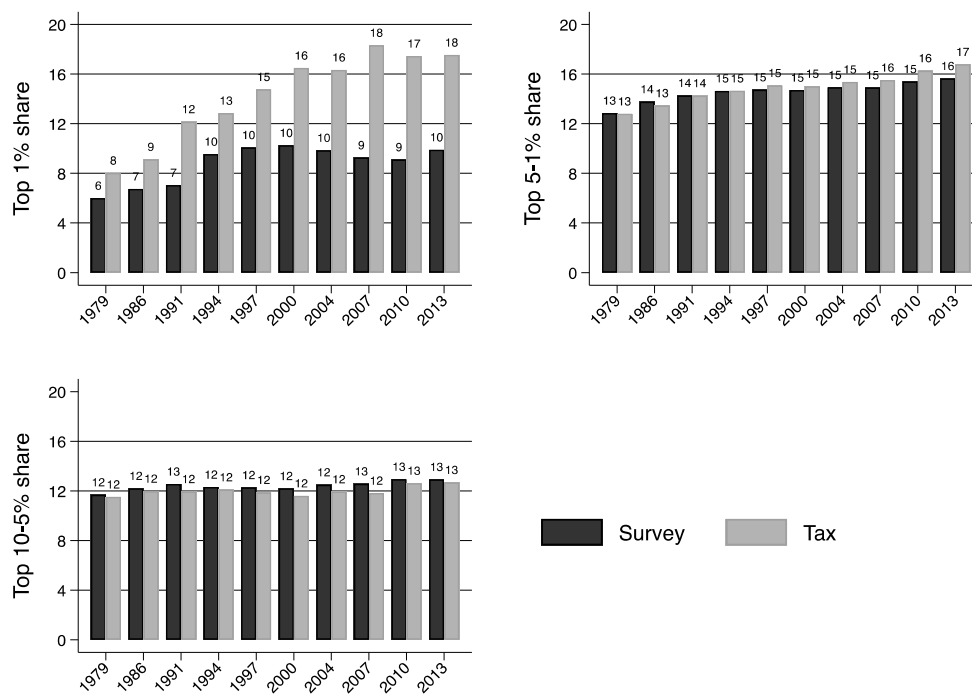
Figure A.2: Fiscal Income Share (in %) of Top Income Groups in France

Source: See Tables 7, 8, and 9.

Notes: This figure compares the fiscal income shares of the top income groups in France using the survey-data and tax-data fiscal income series. The incomes shares reported here are equivalent to those presented in columns 1 and 2 of Tables 7, 8, and 9.

Figure A.3: Fiscal Income Share (in %) of Top Income Groups in the US

Source: See Tables 7, 8, and 9.

Notes: This figure compares the fiscal income shares of the top income groups in the US using the survey-data and tax-data fiscal income series. The incomes shares reported here are equivalent to those presented in columns 1 and 2 of Tables 7, 8, and 9.

Figure A.4: Comparison of Fiscal Income by Income Component for the Top 1% in the US

Source: See Table A.1; LIS Database.

Notes: This figure shows the disaggregation of total fiscal income (in percent) by income components (labor, business, and capital) held by each of the top income groups for all available years for the US.
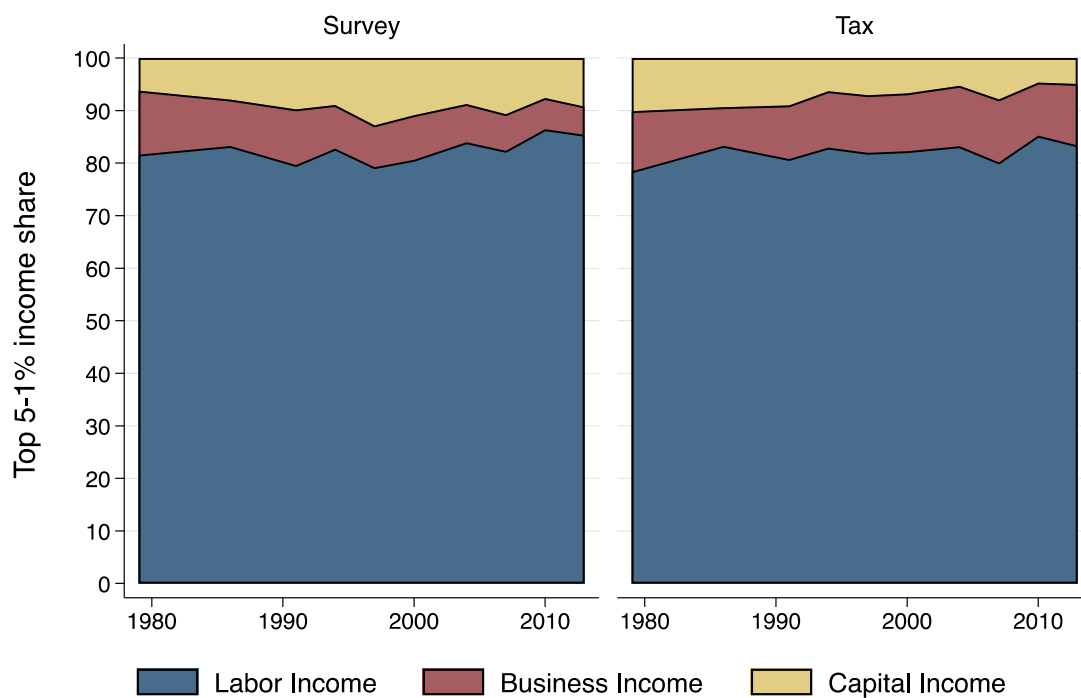
Figure A.5: Comparison of Fiscal Income by Income Component for the Top 5-1% in the US

Source: See Table A.1; LIS Database.

Notes: This figure shows the disaggregation of total fiscal income (in percent) by income components (labor, business, and capital) held by each of the top income groups for all available years for the US.
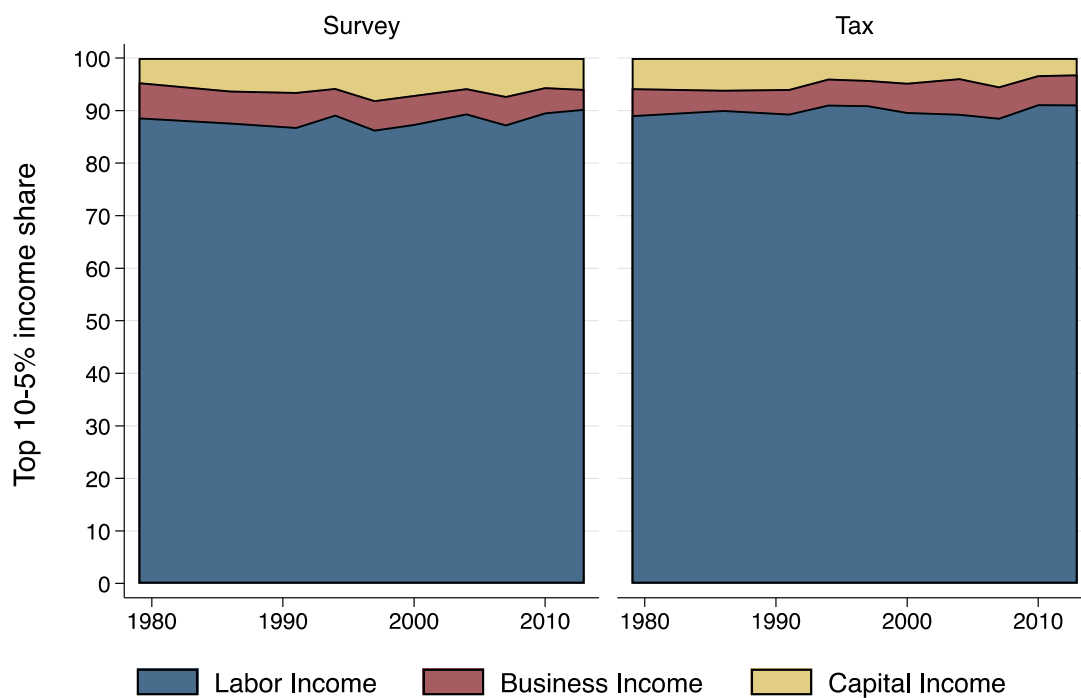
Figure A.6: Comparison of Fiscal Income by Income Component for the Top 10-5% in the US

Source: See Table A.1; LIS Database.

Notes: This figure shows the disaggregation of total fiscal income (in percent) by income components (labor, business, and capital) held by each of the top income groups for all available years for the US.
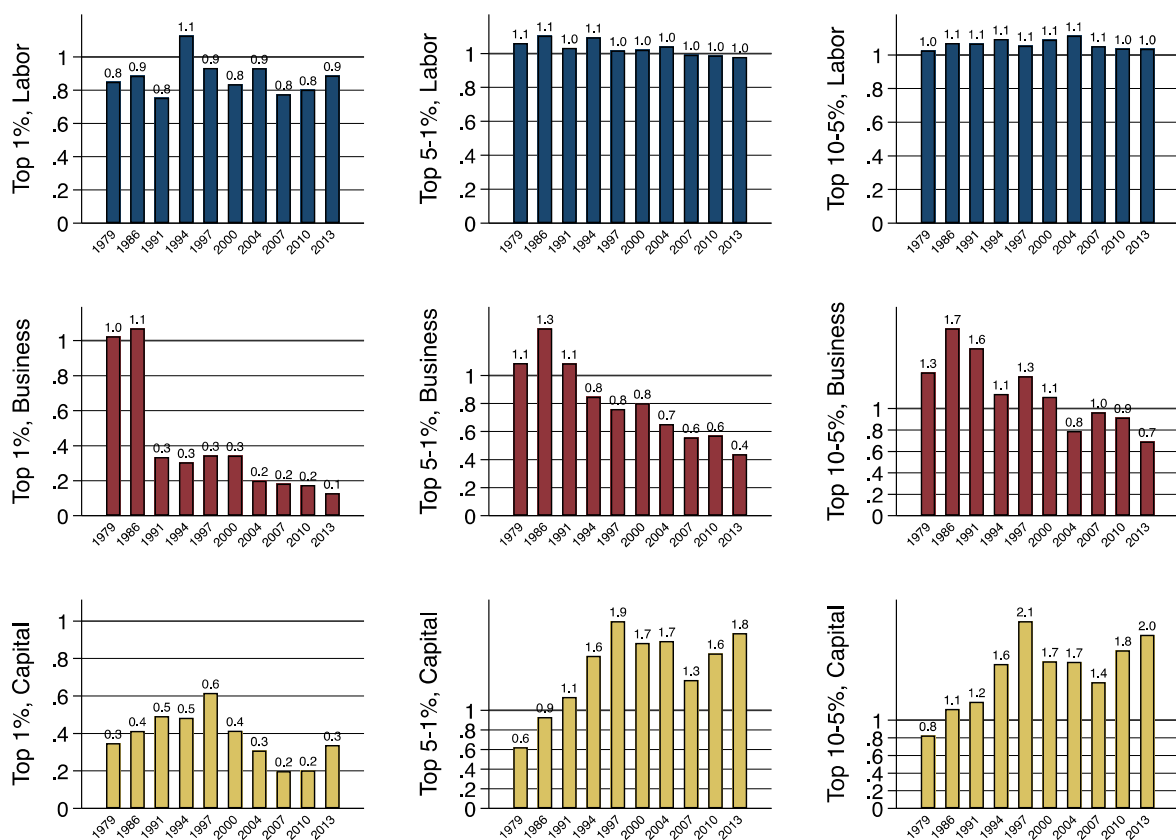
Figure A.7: Ratio of Survey Total Fiscal Income over Tax Total Fiscal Income by each Income Component in the US

Source: Authors' calulation using Table A.1 and LIS Database.

Notes: This figure shows the ratio of survey total fiscal income over tax total fiscal income held by top income groups, disaggregated by income components (labor, business, and capital), in the US. A ratio equal to 1 would imply equal income held by the corresponding income groups in the two sources of data.