Measuring income and wealth at the top using administrative and survey data

Jesse Bricker, Board of Governors of the Federal Reserve System
Alice Henriques, Board of Governors of the Federal Reserve System
Jacob Krimmel, University of Pennsylvania
John Sabelhaus, Board of Governors of the Federal Reserve System

Jesse Bricker
Federal Reserve Board

Alice Henriques
Federal Reserve Board

Jacob Krimmel
University of Pennsylvania

John Sabelhaus
Federal Reserve Board

Measuring Income and Wealth at the Top Using Administrative and Survey Data[1]

ABSTRACT   Most available estimates of US wealth and income concentration indicate that top shares are high and rising in recent decades, but there is some disagreement about specific levels and trends. Household surveys are the traditional data source used to measure top shares, but recent studies using administrative tax records suggest that those survey-based top share estimates may not be capturing all of the increasing concentration. In this paper we reconcile the divergent top share estimates, showing how the choice of data sets and methodological decisions affect the levels and trends. Relative to the new and most widely-cited top share estimates based on administrative tax data alone, our preferred estimates for both wealth and income concentration are lower and rising less rapidly in recent years.

Understanding the determinants and effects of wealth and income inequality are mainstays of political economy. Within the general topic of inequality, the study of *top* wealth and income shares garners particular interest. Measuring and explaining wealth and income concentration has challenged economists at least since Pareto (1896) and Kuznets (1953), and the recent availability of high-quality micro-level administrative tax data is generating renewed interest in the share of resources controlled by top wealth and income groups. Indeed, the striking trends in U.S. top wealth and income shares reported in the most widely-cited studies based on those newly-available administrative data sets are now accepted as facts to be embraced and potentially addressed by policy makers. These observations about levels and trends in top wealth and income shares have begun to transcend academic debates, entering the mainstream political arena through best sellers such as Rajan (2010), Stiglitz (2012), and Piketty (2014), and through political movements such as *Occupy Wall Street*.

Despite the political controversies generated by estimated top wealth and income shares, relatively little attention is paid to the sensitivity of those estimates to data and methodology.[2] For example, using administrative income tax data, Saez and Zucman (2016) estimate that the top 1 percent (by wealth) had a wealth share of 42 percent in 2013, up from 29 percent in 1992. However, the Survey of Consumer Finances (SCF), which combines administrative and survey data, shows less than half the increase in the top 1 percent wealth share, rising from 30 percent in 1992 to 36 percent in 2013.[3] Similarly, Piketty and Saez (2003, updated) show that the top 1 percent (by income) had a 23 percent income share in 2012, an increase of 10 percentage points since 1992. The SCF shows a 20 percent income share for the top 1 percent in 2012, an increase of 8 percentage points since 1991.[4] Differences in levels and trends in top wealth and income shares at higher fractiles, such as the top 0.1 percent, are even more striking.[5]

The goals of this paper are to investigate why the various data and approaches are giving different answers about top wealth and income shares, and to provide preferred estimates that reflect what can best be gleaned from all of the available data. The two main sources of micro

---

[2] Notable exceptions include Burkhauser, Larrimore, and Simon (2012), Burkhauser, Feng, Jenkins, and Larrimore (2012), and Smeeding and Thompson (2011) for top income shares, and Kopczuk (2015) for top wealth shares.
[3] Bricker, et al. (2014) describe the results from the latest SCF, conducted in 2013. A slow rise in top wealth shares is also consistent with estimates derived from administrative *estate tax* data (Kopczuk and Saez, 2004).
[4] SCF income values are for the year preceding the survey.
[5] These issues are not unique to the U.S. See, for example, Atkinson, Piketty, and Saez (2011) who provide a multi-national and longer-run view of rising income inequality.

data used here are administrative tax records and the SCF household survey, but the framework for reconciling and generating preferred estimates of wealth and income concentration relies importantly on available macro data. The macro data are key for understanding top wealth and income shares, because changes in the aggregate composition of income and balance sheets over time affect *for whom* the micro data are comprehensively capturing resources. Further, the two micro data sources measure income and wealth differently and rely on different income and wealth concepts. The biases that arise from different measurement and concepts can be quite large. Overall, the top share estimates derived in this paper show much lower and less rapidly increasing top shares than the widely-cited values from the Saez and Zucman (2016) and Piketty and Saez (2003) studies mentioned above.[6]

This paper begins with the goal of producing new and *improved* estimates for the share of wealth and income held at the top of their respective distributions. From an economic point of view, the preferred concept of wealth includes all assets over which the family has a legal claim that can be used to finance present and future consumption. This concept mirrors the household wealth concept used by the Financial Accounts of the United States (FA) as it includes a family's liabilities, financial and nonfinancial assets, as well as rights to defined benefit (DB) pensions. The preferred income concept includes all income received by the family, whether or not it is fully taxed, partially taxed, or untaxed. This concept mirrors Personal Income in the National Income and Product Accounts (NIPA). Both the FA and NIPA are aggregate data, though, and micro data sets are needed for distributional analysis.

Several challenges must be confronted when estimating wealth and income distributions with micro data, such as the SCF or the administrative tax data. The first is that micro data sets do not include every FA wealth concept nor every NIPA income concept. Untaxed income, like the value of employer-provided health insurance and some government transfer income, is never collected in the income tax data and only sometimes collected in a survey. The SCF wealth estimate typically does not include defined benefit pensions, while most forms of consumer debt cannot be estimated when wealth is inferred from income tax data.

---

[6] The top share estimates from Piketty and Saez (2003) and Saez and Zucman (2016) are regularly updated and published in *World Wealth and Income Database* maintained by Facundo Alvaredo and Tony Atkinson along with Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. The database is accessible at www.wid.world/.

A second estimation challenge concerns differences in population coverage and measurement between these micro datasets. Household surveys are generally thought to reliably cover the full income and wealth distribution, save perhaps the very top. Administrative tax data can reliably cover the top, but coverage suffers at the bottom of the distribution because many families are not required to file tax returns.

Differences in measurement also arise in the unit of analysis, which are *tax units* in the income tax data and the *family* in a household survey. There are many more tax units (161 million) than families (122 million). 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 includes more observations than counting the top 1 percent (1.22 million) of families in a survey.

In addition to the conceptual, coverage, and unit of analysis difficulties that plague efforts to measure either income or wealth concentration, estimating top wealth shares using administrative tax data introduces yet another potential sources of error. Wealth can only be measured indirectly in income tax data—meaning wealth is inferred mainly by "capitalizing" income flows—which is the approach at the heart of Saez and Zucman (2016).[7] In a survey like the SCF, wealth is measured directly by querying families about their balance sheets.

Accounting for these measurement differences by constraining the SCF to match administrative tax data concepts resolves the discrepancies between the various top wealth share estimates. In particular, evidence here and in Kopczuk (2015) shows the sensitivity of wealth inferred from income tax data. For example, the top 0.1 percent wealth share decreases by nearly 20% (3.5 percentage points) when the rate of return on interest income *alone* is varied slightly.

Estimated income shares in the administrative tax data (Piketty and Saez, 2003) and SCF are actually quite similar, so conceptual mismatch in that sense is not a serious problem. However, there is trend bias in both the SCF and the administrative tax-based top income shares because of the increasing share of personal income not being measured in either of those micro data sources.

---

[7] Greenwood (1983), among others, provided the foundational work for the capitalization approach. Capitalization is used in conjunction with other approaches in the SCF sampling procedure (see online appendix to this paper and Kennickell and Woodburn, 1999).

The central goal of this paper, though, is to go beyond reconciliation and provide *preferred* top share estimates. These preferred estimates marry the concepts from the macro data to the micro data and cover the full target population, which is all US families. We provide evidence that augmenting the SCF gets us close to this ideal. We first demonstrate that the SCF represents the full family income and wealth distribution, save for the Forbes 400. By augmenting the SCF household survey along these lines and aligning the preferred wealth and income concepts and measurement laid out above, we derive preferred top share estimates.

The reasons for focusing on both wealth and income in one paper are mostly practical. Wealth and income are strongly correlated, so the decisions about how to measure top wealth shares are not neatly separated from the decisions about how to measure top income shares. Indeed, the principle of "capitalizing" specific income flows forms the basis of wealth inferences in the administrative income tax data and is also used to infer who should be surveyed in the SCF (as described in the on-line appendix). This process ties top wealth and income share estimates together in an important way.

In addition to the statistical issues, there is also an important conceptual reason for considering both wealth and income concentration in the same paper. Neither income nor wealth concentration tells us everything we want to know about key questions in political economy, but together, the two tell us most of what we want to know. Top income shares are interesting because changes in the *flow* of returns from current production suggests that something may be amiss in how factor payments are being determined. Top wealth shares are interesting above and beyond top income shares because disproportionate and/or increasing control over the *level* of economic resources may reflect increasing and persistent income concentration—assuming the rich are saving more of their increased incomes—but it could also be driven by trends in relative asset prices and heterogeneous returns on assets. Though dynastic wealth may be less important today than in the past in determining the wealthiest (Kopczuk, 2015b), both wealth and income concentration may reflect and shape inequality of opportunity (Yellen, 2014).

Some distributional shifts in income might be attributable to fundamental economic factors such as skill biased technological change, but that probably does not explain increased income concentration *within* the top 1 percent. Institutional factors may be having an impact across factors of production generally (capital vs labor) and within factors (managerial vs production labor) such that those with the highest incomes are able to capture even higher shares

going forward. On the other hand, changes in the way that labor is compensated may be mechanically affecting measured top income shares, if (unmeasured) health and retirement costs are disproportionately pushing down incomes for the non-rich.

One specific concern is that wealth concentration may feed on itself, if undue political influence is being exercised by those who can (sometimes independently) finance election campaigns, and generate an even more favorable tax or regulatory environment for themselves in subsequent periods. The primary concerns about the effects of rising wealth inequality involve investment and economic growth. Rising wealth concentration may intensify financing constraints for the non-wealthy, affecting investment in education, entrepreneurship, and other risk-taking for those with diminished resources. As with incomes, however, it is important to consider what may be driving estimates of top wealth shares, before recommending policies to address those trends.

Identifying the potential biases in top wealth and income share estimates begins with a comprehensive discussion of data and concepts, which is the subject of section one of the paper. Section two of the paper then focusses on deriving the preferred estimates for top wealth shares, and section three focusses on top income shares. For both wealth and income, in the course of generating the preferred top shares, we also show how to reconcile the existing SCF and administrative tax data top share estimates. The reconciliation shows that the first-order divergence between the SCF and administrative tax data is basically conceptual in nature, and not a problem of population coverage. The reconciliations generally involve the differences between micro and macro concepts, unit of analysis, whether and how certain groups are represented in the micro data, and potential survey reporting for different types of incomes.

Our preferred estimates for the wealth share at the top are lower and growing more slowly than in the widely-cited "capitalized" administrative tax data from Saez and Zucman (2016), but that is mostly for methodological reasons, especially the specific capitalization factors used to estimate certain types of wealth. Indeed, our preferred top wealth share estimates are quite similar to the published SCF values, because one adjustment (adding the Forbes 400) pulls up the SCF top wealth shares, and another adjustment (distributing defined benefit pension wealth) pushes top shares down by a similar amount.

Our preferred estimates for top income shares are also lower and rising less rapidly than the recent and widely-cited estimates from Piketty and Saez (2003, updated) derived from

administrative tax data. However, those administrative tax data income shares are similar (on an equivalent basis) to SCF top shares, and thus the preferred income top shares are also lower and growing more slowly than published estimates based on the SCF. The differences in levels for incomes at the top (by income) are affected to some extent by the choice of measuring incomes for tax units versus families, but in the end, the wedge in the trends between our preferred and the available top income share estimates is largely driven by the failure of the available micro data to capture cash and in-kind transfers, which are growing rapidly as a share of total income over time.

I. Measuring Wealth and Income Concentration: Concepts and Data Sources

Our starting points for measuring top wealth and income shares are the aggregate concepts and estimates of household sector net worth and income built into the Financial Accounts of the United States (FA) and the National Income and Product Accounts (NIPA). The distributional analysis itself is based on two distinct (but related) micro data sets. Top income and wealth shares are first estimated using the Survey of Consumer Finances (SCF), a household survey micro data set collected by the Federal Reserve Board. Top income and wealth shares are then estimated from an administrative income tax data produced by the Statistics of Income (SOI) Division at the Internal Revenue Service. These SOI administrative micro tax data are the direct source of top income shares in Piketty and Saez (2003), the indirect source of top wealth shares in Saez and Zucman (2016), and the basis for drawing the sample of SCF high-end respondents.

This section describes how the various wealth concepts, income concepts, population coverage, and unit of analysis compare and contrast across these four data sets, and thus sets the stage for developing preferred estimates of top wealth shares in section II, and top income shares in section III.

I.A. Wealth Concepts and Data

Our starting point for measuring wealth concentration is the concept of net worth owned by the household sector, as embodied in the Financial Accounts of the United States (FA).[8] FA household sector net worth includes the net worth of non-profit institutions serving households. The FA does make it possible to separate non-profit real estate holdings, but financial assets owned by non-profits will be included in the overall net worth measure in the FA.[9] Net worth is generally calculated as the difference between households' total assets (financial and non-financial) less their total liabilities (debts to other sectors). However, because households effectively "own" the other *private* sectors (such as corporations) through ownership of equities and debt, household sector net worth effectively represents all private net worth claims.

There is little difference in the conceptual measure of wealth across the micro data (SCF and administrative tax) and macro data (FA). There are, however, key differences in how various balance sheet items are estimated in the two micro data sets, as shown in Table 1. The most notable difference is that income-generating financial and business assets are estimated in the administrative tax data by applying "gross capitalization" to the observed income flows, while those assets are estimated directly in the SCF through the survey questionnaire. A key assumption in gross capitalization is that all assets of a given type earn a single rate of return, and thus there is a direct relationship between the stock and the flow.[10]

Implementing the gross capitalization approach also requires choosing a gross capitalization factor for each asset type, which in Saez and Zucman (2016) is solved for using the ratio of a given FA asset balance to the corresponding aggregate administrative tax data flow. This approach generates micro-level wealth totals that, by construction, match the macro-level wealth totals. However, any mismatch between the micro and macro data concepts will lead to

---

[8] Most of the discussion here is focused on concepts in FA Table B.101, though the reconciliation between SCF and FA aggregates also involves details on pensions from sub-tables such as Table L.117. For details on the SCF and FA reconciliation, see the on-line appendix, Henriques and Hsu (2014), and Dettling et al. (2015). For details on the FA, see Board of Governors of the Federal Reserve System (2015).

[9] Wealth held by most non-profits is appropriately excluded from the household balance sheet, though non-profits are included in the FA aggregate household wealth (because the FA data are the residual of total net worth and everything else). However, charitable foundations controlled by a family is an interesting grey area, as the family does not consume goods and services from the assets in the foundation but may be able to consume reputational benefits. The SCF collects information on the value of such charitable trusts and foundations, and wealth held in these entities. Including these assets along with SCF household wealth would have only marginal impacts on top share estimates. In the 2010 SCF, for example, the wealth share held by the top 1 percent would increase from 34.5% to 34.7%. Further, these assets only constitute about 9% of the total assets held by non-profits (McKeever, 2015).

[10] Fagereng et al, (2016) test this assumption and reject it. Families at the upper tail of the wealth distribution have much higher rates of return than other families. Tabulations from the SCF are consistent with this finding as well.

bias in capitalization factors and a misallocation of wealth. For example, if the FA aggregate for some asset includes holdings of non-profit institutions, while the micro income flows do not, then too much wealth will be assigned (per dollar of income) at the micro level. Similarly, if the micro data misses small income flows—say the modest interest earned on checking and savings accounts in a low interest rate environment—the corresponding FA assets will be assigned only to those families with large and reported interest flows. These possibilities are more than theoretical, as we show later in the paper that implausible capitalization factors are the key to understanding differences between the survey and administrative tax data estimates for top wealth shares.

Assets that do not generate observable income flows, such as housing and pension wealth, are allocated in the gross capitalization framework using correlations with other observables in the administrative tax data, such as property taxes and wages, and are benchmarked to available external sources, such as the SCF or published IRS statistics. Again, those assets are measured directly in the SCF, along with non-mortgage liabilities for which there are no useful correlates in the tax data that can be used for distribution. The one asset category that requires inference in the SCF is defined benefit (DB) pension wealth. The approach for distributing future DB claims in our preferred top share estimates involves using the survey reports of wages, current DB coverage, and years in plan for those still working, and current benefits for those already receiving benefits.[11]

I.B. Income Concepts and Data

Our starting point for estimating top income shares is the concept of Personal Income (PI) as measured in the National Income and Product Accounts.[12] PI is a very broad concept, and meant to capture all forms of income received by individuals, non-profit institutions serving households, private noninsured welfare funds, and trust funds. It includes income that is taxed, partly-taxed (such as Social Security benefits), and untaxed (mostly transfers, whether cash or in-kind). In this section we discuss the conceptual differences between administrative tax data, the

---

[11] The algorithm for distributing SCF DB pension wealth is described in the on-line appendix and in greater detail in Devlin-Foltz, Henriques, and Sabelhaus (2016).

[12] Most of the discussion here is focused on broad income concepts in NIPA Table 2.1, though a comprehensive reconciliation with the micro data also involves details from other parts of the NIPA, such as Tables 1.12, 3.12, 7.9, 7.10, 7.11, and 7.20. For a detailed reconciliation of NIPA and SCF incomes, see Dettling et al. (2015).

SCF, and NIPA, thereby establishing the underpinnings for our preferred top shares estimates presented later in the paper. Although our starting point for measuring top income shares is PI, we acknowledge there are some irreconcilable differences between the micro and macro data, a key timing adjustment, and one notable addition on the micro side, for realized capital gains.[13] These differences are highlighted in Table 2.

In many ways the SCF and administrative tax data are closely related, and generally consistent with the concept of NIPA PI. Most forms of income from current production, including wages and salaries, business income, interest and dividends paid directly to persons, and other smaller types of "market" income are conceptually (and empirically) similar in the two micro data sources. To some extent this is by construction, because the SCF income module invites respondents to refer to their income tax returns when answering those questions. The two micro data sets are in turn mostly consistent with the NIPA in those categories, though NIPA makes adjustments for underreporting of proprietors' incomes and imputes certain incomes such as the rental value of owned housing and value of financial services provided by banks.

The two micro data sets both count realized capital gains as part of the core income measure, while NIPA does not count capital gains in PI. The NIPA exclusion is based on fundamental national income accounting principles. That is, capital gains are not tied directly to current production, nor do they constitute a transfer from one sector to another. However, for the purpose of measuring top income shares, we choose to include realized gains because they do constitute a flow of current resources over which the family has control.

The treatment of retirement incomes is also different in the micro and the macro data. In the NIPA, and again, based on the principle that incomes should be derived from current production or arising from transfers across sectors, retirement "income" occurs when employers contribute to retirement plans on their employees behalf, or when the retirement assets generate interest and dividends. The actual payment of retirement benefits is a mixed bag in the NIPA, with withdrawals and benefits paid from private plans not included, and payments from

---

[13] One aspect of income concentration we do not (and cannot) address in this paper is the conceptual issue of what frequency should be used to measure top shares. Wealth is generally more straight-forward, because concentration is measured at a point in time, though we will see frequency also plays a role there in terms of what can and cannot be measured. One can argue that income concentration should be measured at lower frequencies, in order to sort out transitory income effects, and also to address some of the conceptual issues we raise, such as measuring retirement income when the claim is established versus when the income is actually received. The decision here to focus on annual measures is largely driven by what data is available over long periods.

government plans showing up as transfer income. In the micro data, employer contributions and capital income earned by retirement plans are generally unobserved, but withdrawals are (though to a differing degree in the SCF and administrative tax data) generally observed.

To some extent the appropriate treatment of retirement income cannot be separated from the frequency over which incomes are being measured. On a lifetime basis it would not matter whether the [private retirement] income was counted as it was accrued or when it was paid out, but the distinction does matter when using annual data. Given the availability of cash-flow oriented micro data at an annual frequency, the top shares estimates we present are based on realized benefits, which implicitly adjusts the NIPA PI concept for a portion of "net saving" in retirement plans, where net saving is new contributions plus interest and dividends earned on plan assets, less pensions benefits paid. However, the fact that some new employee contributions (employee paid Social Security taxes) to retirement plans are still counted (in the micro data) as part of non-retirement income means the adjustment is only partial.

The more substantial conceptual differences between our preferred income top share estimates and those available in the micro data are associated with non-taxable government transfers and in-kind compensation. In principle, the SCF captures government cash transfers, but the administrative tax data by construction do not, and the rising share of transfers in NIPA PI means that less total income is being distributed over time when using either micro dataset.[14] Neither the SCF nor the administrative tax data make any adjustment for in-kind compensation and transfers, which, especially through employer-provided health and the major government health programs, have roughly doubled as a share of total NIPA PI since 1988. Our conceptually preferred measure for top income shares allocates these missing income pieces, which brings our overall income concept close to NIPA PI. The remaining conceptual differences are in the imputations and retirement income timing discussed above.

I.C. Coverage and Unit of Analysis

The population of interest in our analysis of top wealth and income shares is all U.S. households. In some ways this is a simplistic statement, because households are the ultimate claimants on all private incomes and wealth. However, there is substantial private income

---

[14] The evolving differences in the concept of income in administrative versus survey data is also emphasized by Burkhauser, Larrimore, and Simon (2012), and Armour et al (2014).

received and wealth owned by non-profit institutions that should be excluded, and that is not completely feasible to sort out given the available macro data. In addition to these sectoral coverage issues, there are also differences in population coverage and measurement across the distribution of households, with administrative income tax data generally perceived to be more accurate at the top of the distribution, and household surveys like the SCF thought to provide better coverage at the bottom. These comparisons are further confounded by the differences in unit of observation across the micro data, with the administrative data collected for tax units, and the survey data collected for households.

Table 3 summarizes the differences in coverage and unit of analysis across the four data sets we are working with. The first key difference between the two micro data sets is unit of analysis. In the U.S. income tax data, observations are tax filing units, not families. The number of tax units (about 161 million in 2012) is approximately 30 percent higher than the number of families (122 million in the SCF).[15] Most of the tax units at the very top are also families, meaning that many of those observed as a single family in survey data but multiple tax units in the tax data are found in the bottom 99 percent of the wealth and income distribution. In the 2010 SCF, for example, less than 3 percent of coupled families in the top 1 percent filed separately, while about 17 percent of couples in families in the bottom 99 percent filed separately. The implication, then, is that any top share fractile estimate is effectively based on a population that may include 30 percent more family units than the fractile suggests.

There are many reasons to prefer the household (or family, which is close to household) as the unit of analysis for measuring top wealth and income shares. Many of the tax units residing in multiple tax unit families are dependent filers, with very low incomes, and therefore effectively sharing resources with the other members of the household (usually their parents) who are able to claim them on their taxes. The same can be argued for unmarried partners sharing living arrangements and resources but filing taxes separately. It makes sense to pool their resources when characterizing their share or income or wealth. One can argue that roommates who are not sharing resources could be treated as separate units, but in the end, the issue is really

---

[15] Statistics on tax units here and later in the paper are from Emmanuel Saez's website, in the regularly updated file http://eml.berkeley.edu/~saez/TabFig2014prel.xls.The actual unit of observation in the SCF is the "Primary Economic Unit," or PEU, which is somewhere between the Census "family" and "household" concepts. See the appendix to Bricker et al (2014) for a precise definition. The number of families in the SCF is benchmarked to that found in the Current Population Survey. The number of tax units includes an estimate of non-filers.

about what one means when measuring the wealth or income shares of "the" top 1 percent. Is it the top 1.22 million families in 2012, or the top 1.61 million tax units? Our preferred estimate is based on families, and the substantial difference between the total counts of families and tax units will turn out to be a key drivers of the wedge between existing estimates of the levels of top wealth and income shares.

Sector coverage matters when comparing the SCF to administrative tax data, and between the two micro data sets and the two macro data sets. The micro data sets do not attempt to measure wealth and income received by non-profit institutions, and the only available adjustment on the macro side is in the FA balance sheet measure, which separates real estate holdings of non-profit institutions. This sectoral overlap becomes important when thinking about the total income or wealth in the denominator of the concentration measures, and whether (for example) a given income flow or asset holding should be allocated to a given top shares group or spread more evenly throughout the distribution. In particular, the "capitalization" approach to estimating top wealth shares relies on administrative income tax data flows calibrated to FA levels. That approach will assign non-profit non-housing asset holdings across groups based on measured incomes, exacerbating any differences in actual wealth holdings.

There is also a key difference between the micro data sets in terms of population coverage, and that has a potentially first-order bearing on estimated top shares. The goal of the SCF is to survey the entire non-institutional population, using a standard nationally-representative area probability (AP) sample along with the "list" sample derived from administrative tax returns, designed to correct for low survey response among wealthy families.[16] The members of the Forbes 400 in the year the sample is drawn are explicitly excluded from the SCF sample.[17] In our preferred top wealth and income share estimates we add in the Forbes 400, but there is some question as to whether the SCF captures the rest of the top of the (just-below-Forbes) distribution (more on this in the next section).

The population coverage for administrative income tax data is necessarily limited to the income tax filing population. Although there are many more tax units than there are families,

<hr>

[16] See the on-line appendix for a detailed discussion of the SCF sampling strategy. See Sabelhaus, et al (2015) for direct estimates of the relationship between income and unit non-response. O'Muircheartaigh et al. (2002) provide a comprehensive description of the NORC national AP sample.

[17] The sampling frame technically excludes other "public" figures as well, but assuming those families have observational equivalents who not public figures, there is no bias in the estimated wealth distribution.

there are many families (low-income and retired) where no individual or couple is required to file a tax return. Indeed, of the 161 million estimated tax units in 2012, only 145 million actually filed tax returns. Using other household survey data, Piketty and Saez (2003) supplement the tax-based income concentration measures by increasing the denominator (total income) to account for non-filers.[18]

Both the SCF and the administrative income tax data face challenges in terms of population coverage. The coverage challenge for the administrative tax data is mostly about non-filers, and to some extent, the coverage problems cannot be cleanly separated from the concept of income being measured, because the income composition of non-filers is very different than the income composition of filers. The SCF also faces issues in terms of capturing certain types of income, but the more immediate concern is whether the SCF actually captures the top of the distribution, as the sampling strategy is designed to accomplish.

I.D. Does the SCF Capture the Top End?

It is difficult to argue with the presumption that administrative tax data *should* provide better estimates of top wealth and income shares, because "participation" in the administrative data is required by law, and traditional household surveys are well-known to suffer from underrepresentation of very wealthy families.[19] In addition, administrative tax data are subject to audit, and thus (again) one presumes that income and other reporting will be more accurate in those data. Unlike most other household surveys, the SCF is designed to overcome the underrepresentation problem, because administrative tax data are used to select the sample, and rigorous targeting and accounting for wealthy family participation assures those families are properly represented. Also, SCF cases are reviewed for internal consistency (to some extent guided by the administrative sampling data) but that review process may fail to capture all reporting errors. In this section we show that the SCF does a very good job identifying and

---

[18] They estimate that non-filers have 20 percent of the average income of filers, where income is defined using the same taxable income concepts of the filers.

[19] Sabelhaus et al. (2015) show this is the case for the Consumer Expenditure Survey and Current Population Survey (CPS). Burkhauser, Feng, Jenkins, and Larrimore (2012) show that at least some of the divergence between CPS and administrative incomes is also due to top-coding of very high incomes in the CPS. Attanasio, Hurst, and Pistaferri (2015) use household budget data to study inequality, and in addition to the non-response issues, they find that reporting problems further confound consumption-based inequality estimates.

surveying wealthy families, and there may be some downward bias in capturing certain types of income at the very top.

The SCF strategy begins with the view that a combination of survey and administrative data is better than either in isolation. The benefit of the survey component is straight-forward, as the data collector can control the population being studied and the specific wealth and income concepts being measured. However, for the purposes of studying top wealth and income shares, that benefit can be dwarfed by failure to survey wealthy families. Measuring top wealth and income shares by expanding on simple random sampling in a traditional household survey is not a viable solution, because thin tails at the top lead to enormous sampling variability, and disproportional non-participation at the top biases down top share estimates.

The SCF effectively overcomes the problems of thin tails and differential non-participation by oversampling at the top, relying on administrative data derived from tax records, and by verifying that the top is represented using targeted response rates in several high end strata.[20] The SCF "list" sample is actually comprised of seven strata, where the first basically overlaps the address-based random sample, and the remaining strata identifying increasingly wealthy groups of families up to (but not including) the Forbes 400. In very general terms, the top four strata in any given year, made up of roughly 1,000 SCF families, effectively represents the top 1 percent of all families. The targeted response rates in the list sample do vary across strata in an expected manner, with participation rates falling as predicted wealth rises. The response rate in the wealthiest SCF stratum is around 12 percent, increasing to about 25 percent in the second-wealthiest stratum, 30 percent in the third-wealthiest stratum, 40 percent in the fourth- and fifth-wealthiest and then about 50 percent in the two least-wealthy strata. These high-end response rates are considerably lower than the roughly 70 percent response rate observed in the SCF AP sample.

The fact that participation rates are lower for very wealthy SCF families does not mean the sample is biased by underrepresentation at the very top, however, it just reflects the fact that very wealthy families are much more difficult to contact and then less likely to participate in the survey, given contact. Sample weights are systematically varied across the top strata in order to

---

[20] The online appendix has extensive details about the SCF sampling process. At the time the list sample is drawn, the most recent complete administrative data are those from two years prior to the survey year. The sample includes individual and sole proprietorship tax filings from the IRS administrative tax data (see Statistics of Income, 2012).

correct for the differential non-response. The important question is whether the families who eventually participate in the survey, thus representing their respective wealth stratum, are statistically distinguishable from sampled non-participants.[21] Indeed, a regular step in the SCF quality control involves comparing and contrasting participants and non-participants within stratum, in order to identify those sorts of potential biases. These comparisons are based on comparing administrative data incomes in the years prior to the survey.[22]

The administrative data underlying the SCF sampling are consistent with participants being representative of non-participants within each high end strata. The distributions of total incomes for SCF participants are similar to those of sampled non-respondents (Figure 1A). Moving from the fourth-highest stratum to the highest stratum, one sees the substantial non–linearity of incomes that characterize the top end, as each successive log scale for income shifts to the right in dramatic fashion. The range of incomes in the top four SCF strata completely cover the top 1 percent in an overlapping way, meaning, for example, that the top of the fourth-highest stratum overlaps with the bottom of the third highest stratum, and so on. The capital income distribution of SCF respondents are also similar to those of non-respondents (Figure 1B), and the non-linearity in incomes as one moves from the fourth-highest to the highest stratum is even more dramatic.[23]

In general, statistical tests confirm the visual indication that participants and sampled non-participants within strata have very similar income distributions. The null hypothesis is that the two distributions come from the same underlying distribution, and the test statistics generally fail to the reject the null using either the Kolmogorov-Smirnov or Wilcoxon Rank-Sum tests. The specific results vary by year and across strata, but in the 2013 sample, the null was rejected for only the second highest stratum for total income.[24]

Focusing on the means of the distributions across strata, average total incomes for both participants and sampled non-participants in the fourth highest stratum are generally around $500,000, whereas the average total incomes in the highest stratum are above $50 million

---

[21] See, for example, the discussion in Kennickell and Woodburn (1999).

[22] One would perhaps like to compare respondent and non-respondent incomes in the survey year itself, or to compare respondent-reported and administrative incomes for the survey year, but any such comparison would involve an implicit audit and thus violate the explicit agreement the SCF has with respondents to NOT audit their data.

[23] Capital income here includes taxable and non-taxable interest, dividends, Schedule C and Schedule E business income, Schedule F farm income, and capital gains.

[24] Results across income concepts, strata, and for earlier years are available upon request.

(Figure 2A, shown again on a log scale). The averages for total income versus capital income only differ noticeably for the fourth-highest and third-highest stratum (Figure 2B). In the top two stratum, average total income is dominated by and effectively equivalent to capital income. As with differences in the distributions, one can test for differences in the means by income measure, stratum, and year, and in general the tests fail to reject the null that the means for participants and sampled-non-participants are the same.[25]

In addition to average levels, one can also compare SCF respondents and non-respondents in terms of observable pre-survey income volatility. This metric also shows that SCF participants are similar to non-respondents for both total income (Figure 3A) and capital income (Figure 3B). Income at the top is known to be much more volatile than in the rest of the income distribution, and the trend seems to be towards higher relative volatility at the top.[26] In the SCF sampling data, for the top four strata covering the top 1 percent, about one-fourth of 2013 families experienced income changes below -50 percent or above + 50 percent. The similarity between SCF respondents and non-respondents means that potential distortionary effects from sampling families with very high or very low transitory income shocks is not a problem.

Although it would violate SCF protocol to directly evaluate the accuracy of any given SCF respondent's reported income, it is possible to get an estimate of reported income accuracy on average using two distributional comparisons against the entire SOI data set for a given survey year. The first approach is to compare the *growth distribution* of incomes reported by SCF respondents to the *growth distribution* observed in the SOI administrative data for families with comparable income levels. The second approach involves looking at how many SCF families report incomes above published SOI thresholds, and how much income in total is reported by those in a given top income group.[27]

High income and high wealth families typically have volatile income. For example, in the complete 2011 SOI data set, about 60 percent of the families with AGI greater than $500,000 realized a decline in income (AGI) in their 2012 tax filing (Figure 4, red bars). At the tails, about 22 percent of the families in 2011 with AGI greater than $500,000 had a decline in income of 50

---

[25] In 2013, the differences for the second-highest stratum were significant at the 5 percent level. Again, results for other years, income measures, and stratum are available upon request.

[26] See, for example, DeBacker et al (2013), Guvenen, Kaplan, and Song (2014), and Parker and Vissing-Jorgenson (2010).

[27] We are grateful to the IRS Statistics of Income Division for the unpublished growth rate distributions and threshold comparisons described here.

percent or more, and about 11 percent had an increase in income of 50 percent or more. However, of the 2011 SOI families with AGI greater than $500,000 and that responded to the SCF, about 74 percent reported an annual income decline (survey-reported income relative to the last year of administrative sampling income) and nearly 32 percent reported a decline in income of 50 percent or more (Figure 4, blue bars). Thus, although the patterns of income change in Figure 4 are broadly similar, some high income SCF respondents may be, on net, underreporting 2012 income, and the SCF data editing process does not correct for that underreporting. One possible explanation is that many high income SCF families haven't filed their taxes at the time of interview so they may be unaware of their actual 2012 income during the interview.[28]

In addition to comparing growth rate distributions, it is possible to look at whether the SCF is capturing the very top of the SOI income distribution in any given year. One of the (now regular) tables published in the *SOI Bulletin* shows income thresholds for various top share groups, along with the amount of income earned above those thresholds.[29] Thus, it is possible to look at various SOI cutoffs (for the top 10 percent, top 1 percent, and top 0.1 percent) and investigate whether the SCF finds the right number of families above those cutoffs, and the right amount of total income above the threshold. These comparisons are far from perfect, because the SCF is on a family basis while SOI is in tax units, and (although SCF respondents are asked to refer to their tax returns) the value of income they report may differ from the Adjusted Gross Income (AGI) concept in the SOI tables.[30] Indeed, the modest biases one expects show up clearly: the SCF has more families above any given threshold and generally more income (additional family income will increase a given tax unit's income, which pushes a few more families over the threshold) except for the top 0.1 percent, for which the SCF finds roughly the same total income (the tax unit vs family distinction is less important as one gets closer to the

---

[28] Almost 19 percent of SCF families in the top two sampling strata have not yet filed their taxes as of interview date but plan to do so; only 4 percent of all other SCF families have not yet filed taxes. Many high-wealth families file the taxes late in the year, after getting an extension.

[29] The published series is "Individual Income Tax Shares," with the latest (by Adrian Dungan) published in the Spring 2015 *SOI Bulletin*, and available on-line at www.irs.gov/uac/SOI-Tax-Stats-Individual-Income-Tax-Return-Form-1040-Statistics. We are grateful to SOI for providing thresholds and counts in the early SCF years not covered in the published time series.

[30] One subtle point about negative incomes affects the very top end in an important way. A taxpayer experiencing a capital loss may have that loss limited in a given tax year, but (for example) a business loss may be fully deductible against other positive incomes. Thus, if an SCF respondent accurately reports a loss, but misreports the type of loss, they could be misclassified based on "total" income. The analysis here is based on the SCF "total income" measure, which is, at the end of the day, the respondent's best estimate as to what they actually received during the year.

very top). Importantly, we do not observe any trend in how well the SCF captures top income over time.

Though the SCF covers the top end of the income distribution, other comparisons of SCF and SOI incomes by source suggest there may be more general income reporting challenges for capital income – such as interest, dividend, and business income – that are likely affecting top families. For example, Moore and Johnson (2008) show that aggregate total income in the SCF generally matches total aggregate income published by SOI, but the aggregate of some forms of capital income in the SCF appear to be understated while wages and other types of income are overstated relative to the tax data. Saez and Zucman (2016) also state that the capital income *concentration* in the SCF is lower than the capital income concentration in the income tax data, and argue that this is evidence that the SCF is not capturing the top of the distribution.

How can the SCF capture the top of the income distribution and match total taxable income but have understated capital income shares? We argue that understated capital income in the SCF is mainly due to *classification* of income. Wages as a share of total income of the wealthiest SCF families has grown more than in the tax data since 2001.[31] We concede that some of what respondents call wages may, in fact, be business income, as the two could be thought of interchangeably to business owners. Business income is the largest source of capital income in both the SCF and the income tax data.[32]

The question posed at the beginning of this section is whether the SCF accomplishes its goal of identifying and surveying high end families. The answer is basically yes, though given the restriction on auditing respondents, there will always be some uncertainty about exactly who is being included and whether their reported incomes are accurate. The importance of showing that the SCF captures families at the very top is in one sense first order for our purposes here, but in another sense, it is just a corollary to the fact established later in the paper that, *after being made conceptually equivalent*, top wealth and income shares in the SCF and administrative tax data are effectively the same. Given that the populations in the two micro data sets are

---

[31] The wage share of income of the top 1 percent of SCF families was 47% in the 2001 SCF and was 49% in 2013 (author's calculations). In the tax data, comparable wage share of families reporting more than $200,000 in AGI (roughly comparable to the top 1 percent) was 45% and decreased to 44% (SOI table 1.4).
[32] We also show in the on-line appendix that the income tax data may be missing some forms of capital income for lower income families in recent years, which would lend an upward bias to capital income concentration estimates in the income tax data in figures III and X of Saez and Zucman (2016). Further, the shares reported in the final year of these figures is undoubtedly biased up because 2012 was a year than many wealthy families *chose* to realize capital income (Wolfers, 2015).

effectively aligned, the more salient questions involve *what* we should be measuring conceptually, and *how* should we should be measuring those desired concepts.

II. Top Wealth Shares in Administrative and Survey Data

Wealth concentration is at the center of recent media discussions (Feldstein, 2015, Harwood, 2015, Wolfers, 2015) and academic discussions (Auerbach and Hassett, 2015, Mankiw, 2015, Piketty, 2015, and Weil, 2015). In addition to concerns about causes and effects of rising wealth concentration, some of the debate exists because different wealth concentration estimates paint contrasting pictures about what is actually happening. Published SCF household survey estimates indicate that wealth concentration at the top is high but increasing slowly (Bricker et al, 2014) with a trajectory similar to that in estate tax data (Kopczuk and Saez, 2004), though the level of wealth concentration is higher in the SCF. The inferences about top wealth shares using capitalized income tax data (Saez and Zucman, 2016) indicate much higher and more rapidly growing wealth shares at the top of the wealth distribution.

In this section we present our preferred estimates of top wealth shares, and we show how those preferred estimates compare and contrast to both published SCF and the gross capitalization estimates. Our preferred top share estimate is constructed by starting with SCF wealth measures, adding estimated wealth of the Forbes 400, and then distributing the value of defined benefit (DB) pensions as measured in the FA. The preferred measure shows slower *growth* in wealth concentration than in Saez and Zucman (2016), and in fact, the preferred top shares growth rate is very similar to the SCF.[33] The differential growth in wealth concentration has led to a substantial widening between *levels* of estimated wealth concentration in recent years. We also investigate the source of divergence in growth rates and levels, by constraining the SCF to conceptually match Saez and Zucman (2016). Using that approach, we are able to confirm that the differentials in wealth concentration are not attributable to the wealth concept per se, nor to population coverage or survey-reporting errors and are, in fact, attributable to assumptions and methodology.

---

[33] The slower growth of top shares in the SCF is also consistent with patterns in top shares derived from estate tax data, as in Kopczuk and Saez (2004). Saez and Zucman (2016) include updates of the estate tax estimates in their 2016 paper, but those estimates are very sensitive to assumptions about mortality differentials for decedents affected by the estate tax.

II.A. Preferred Estimates of Top Wealth Shares

In all of the estimates discussed here, top wealth shares in the United States are very high and increasing over time. Figure 5A shows the estimated share of wealth owned by the top 1 percent for the period 1989 to 2013 based on three different measures, and Figure 5B shows the same for the top 0.1 percent wealth shares. In general, the estimated top wealth shares using the gross capitalization method applied to administrative tax data (the grey lines) produced by Saez and Zucman (2016) are higher and growing more rapidly than the top wealth shares in published SCF estimates (the red lines) and also higher than those based on our preferred measure (the black lines).

Our preferred measure for top wealth shares begins with the published SCF *Bulletin* concept and estimates, next adds the wealth known to be missing because the Forbes 400 is excluded from the SCF sample, and then adds the value of DB pensions.[34] With those two adjustments, the preferred measure is conceptually equivalent to household sector net worth in the Financial Accounts of the United States (FA) but excludes non-profit institutions.[35] Thus, the measure encompasses all the private resources available to families for present and future consumption. Most of that wealth is "marketable" in the sense of being available to trade for current consumption, with the exception of DB wealth, but that reflects private claims to future consumption.

Estimates of top wealth shares for both the top 1 percent and the top 0.1 percent were closer across the methods in the early years of the SCF than they are now, but differential growth rates have led to very different levels in recent years. In the most recent period, the preferred estimate for the top 1 percent wealth share is about 33 percent of total wealth, while the

---

[34] 'Bulletin' wealth derives its name from the fact that this is the consistent series published in the *Federal Reserve Bulletin* after each triennial survey. For the most recent survey, see Bricker, et al, (2014). Our estimate of Forbes 400 wealth by finding the sum of wealth of these families from the Forbes list, which was (for example) $2.021 trillion in 2013, or about 3 percent of total household wealth. We add this total to the total wealth in the SCF to create a new estimate of total US family wealth. To compute a new top 1 percent estimate, we remove from the SCF top 1 percent those families that represent the 400 lowest wealth families (weighted) and add the total Forbes wealth, then divide by the new estimate of total US family wealth (which includes Forbes wealth). Alternatively, we can estimate the top shares after including the Forbes families by using inferences from a Pareto distribution (see Vermeulen, 2016). The answers are qualitatively similar, though we prefer to use the data rather than make the inherent assumptions necessary for the Pareto distribution.

[35] There are a few minor differences between the preferred measure and FA household sector net worth, described in the on-line appendix, and introduced to make the estimates more consistent with Saez and Zucman (2016). Primarily, we start with SCF *Bulletin* net worth, subtract vehicles, miscellaneous financial and nonfinancial assets, cash value of whole life insurance, and miscellaneous debt.

capitalized income value is nearly 42 percent. In a proportional sense, the divergence in the most recent years is even larger for the top 0.1 percent, with the preferred measure showing a share just under 15 percent of total wealth, and the capitalized income value over 22 percent. The different measures all agree that wealth concentration is increasing *within* the top 1 percent, though the gross capitalization estimates are the most extreme in that regard.

II.B. Reconciling the Wealth Concentration Estimates

If the SCF sampling strategy does a good job capturing the top end of the wealth distribution, and SCF respondents do a good job reporting the value of the assets and liabilities, what is causing the substantial divergence between estimated top wealth shares in the SCF-based preferred and gross capitalization measures? Our approach to answering this question involves constraining the SCF to be conceptually and empirically similar to the gross capitalization estimates, and showing that most of the divergence is eliminated. In particular, when we measure top wealth shares after constraining SCF totals to match FA aggregates and adjusting the number of families in the top fractile to be consistent with tax unit counts, most of the recent level differences are eliminated, or at least brought within the range of SCF statistical confidence.

The effects of constraining the SCF-based preferred top wealth share estimates to be conceptually and empirically equivalent to the gross capitalization estimates are shown in Figure 6A for the top 1 percent, and Figure 6B for the top 0.1 percent. The first adjustment, which involves moving from the preferred (black) lines to the orange lines labelled "Preferred Wealth, FA Concepts and Values" is based on calibrating the sum of SCF values to match FA values across asset and liability categories. In general, the SCF and FA aggregates track very well over long periods of time.[36] There are some notable differences in levels and trends, however. Most importantly, the SCF finds a higher and (since 2001) more rapidly rising estimate for the value of owner-occupied housing, which pushes up the ratio of SCF to FA net worth in recent years.[37]

---

[36] See Dettling, et al., (2015) for a comparison of aggregate SCF and FA balance sheets for the 1989 through 2013 period. Also, Brown, et al., (2011) show that SCF debt by category generally tracks Equifax aggregates very well, though some categories such as credit cards are difficult to compare because of point-in-time versus revolving balance accounting for debt outstanding.

[37] The differences in SCF and FA housing stock valuations are driven by the very different methodological approaches. In the aggregate FA data, the housing stock is valued using a perpetual inventory that involves new investment, depreciation, and a national house prices index. In the SCF, house values are owner-reported. Henriques and Hsu (2014) discuss how house values in the SCF compare favorable to other micro-based estimates, such as the American Housing Survey, and Henriques (2013) provides evidence that SCF respondent house valuations generally track local area house price indexes quite well. See the on-line appendix for more details.

Thus, when the SCF house values (and other asset and liability categories) are scaled to match the corresponding FA aggregates, owner-occupied housing is disproportionately scaled *down*. The differential rescaling is important, because the divergence in owner-occupied housing aggregates implies that benchmarking administrative data to FA instead of the SCF *lowers* wealth more *below* the top fractiles than *above* the top fractiles, and more so for the top 0.1 percent than even the top 1 percent.

The second set of constraints imposed on the SCF adjustment involves shifting the top fractile cutoffs to be on a tax return instead of a household basis.[38] The shift from the orange lines in Figure 6A and Figure 6B reflect the impact of imposing this constraint, and the lines labelled "FA Concepts and Values, Tax Units" are again noticeably shifted up. We also add the shaded area around the second constrained top share estimates, which represents the 95 percent confidence interval.[39] Indeed, all of the differences in recent top 1 percent wealth shares are effectively eliminated when we constrain the SCF, and all but the most recent periods are reconciled for the top 0.1 percent. The exercise does raise questions about why, for example, the SCF top 1 percent wealth shares are above the capitalized values in the early years of the survey, and why the top 0.1 percent shares are growing much more rapidly in recent years, but the magnitude of the adjustments and range of the confidence intervals makes it clear that top wealth shares are very sensitive to the specific data and methods being used.

II.C. Gross Capitalization for Fixed Interest Assets

Much of the difference between our preferred estimates and the capitalized income top shares can be reconciled by trivial changes to the data, meaning whether or not to calibrate to the FA aggregates or whether to count the top 1 percent versus the top 1.3 percent of families. The remaining difference in top wealth shares is more about trends than levels, as both the top 1 percent and top 0.1 percent wealth shares are rising more rapidly in the gross capitalization estimates, relative to even our constrained SCF-based estimates. It turns out that the gross

---

[38] In practice this constraint is imposed by simply changing the target counts of families in a given fractile to match the estimated number of tax units in a given fractile, which is the same as saying that every household at the top is also a tax unit. As noted earlier in the paper, there were about 30 percent more tax units than families in 2013, so one can think of the constrained "top 1 percent" as really representing the top 1.3 percent of families. The on-line appendix has details about the distributions of tax units versus families.

[39] The on-line appendix and SCF website have details about how to use replicate weights and bootstrapping for generating confidence intervals consistent with the dual-frame sample design.

capitalization implied rate of return on just one asset class (fixed-income) is responsible for *all* of the differential growth in wealth concentration at the very top. That is, when (more realistic) alternative rates of return are used in gross capitalization, the growth at the very top looks very much like the SCF-based top wealth share growth rates shown above.

The analysis of the biased gross capitalization factor begins with the actual (SCF) versus derived (gross capitalization) portfolios of the top 0.1 percent, as shown in Figures 7A and 7B. Assets of the top wealth holders are broken down into four broad categories: housing, pensions, equity plus business, and fixed income. The shares of the first three are very similar, and the share of fixed interest assets are also similar through 2001 or so. Indeed, all of the differential growth in wealth of the top 0.1 percent occurs in the fixed interest asset class, consisting mostly of bonds, CDs, call accounts, money market accounts, and other savings instruments. As of 2013, the gross capitalization approach implies that nearly half of assets owned by the top 0.1 percent are in the fixed interest class.

Is this dramatic shift in portfolio composition plausible, or just an artifact of the gross capitalization approach implemented by Saez and Zucman (2016)? To answer this, we consider the implied gross capitalization factor underlying those estimates, and compare those to the implied capitalization factors if one instead uses a market rate of interest or an alternative based on estate tax filings. The result of those comparisons is shown in Figure 8. The current low-interest rate environment has led to increases in capitalization factors based on 10 year Treasury yields, the Moody's AAA bond yield, or the ratio of prior year interest income to estate tax fixed-interest assets, any of which may be on the high end of plausible values. However, the implied gross capitalization factor solved for using the ratio of FA assets to administrative tax data interest income is much higher, and has clearly reached implausible levels.[40] Based on that estimate, for every one dollar of observed interest income, gross capitalization is currently generating nearly 100 dollars of wealth.[41]

---

[40] For reference, the gross-capitalization model used in the SCF sampling exercise (see the on-line appendix) uses the Moody's AAA rate to capitalize SOI interest income. It is also worth noting that the bond series in the B.101 table of the FA has been subject to downward revision as new source data become available.

[41] The rate of return on these sorts of assets does appear to vary across the wealth distribution in the SCF. In the 2013 SCF, the average rate of return on fixed-income assets (found by the ratio of SCF interest income to SCF fixed-income assets) across all households is about 1 percent, but the average rate of return for the top 1 percent of families is almost 6 percent. Fagereng et al, (2016) also show that families at the upper tail of the wealth distribution have much higher rates of return than other families.

Figure 9 makes the point clearly that there is basically no remaining unexplained difference in top 0.1 percent wealth shares when the constrained SCF is compared to gross capitalization when even a slightly more reasonable rate of return is used. Lowering the implied capitalization factor at the top to be consistent with (the still very conservative) ten year Treasury rate, the top 0.1 percent wealth share (shown by the blue line) lies almost completely within the confidence interval for the constrained SCF estimates. The re-estimated top 0.1 percent wealth share under the alternative gross capitalization parameters falls to just under 19 percent in the most recent period, which is still well above our preferred estimate of about 15 percent, but those differences are completely explained by the other constraints imposed above.

What is driving the implausible capitalization factors in the Saez and Zucman (2016) estimates? Our discussion of data and methods in Section I indicates that a few things can go awry when using the ratio of the estimated FA asset value to measured income flows. The FA asset totals includes holdings by non-profits, while the taxable income flow does not, so the gross capitalization factor is biased up. The household sector of the FA tries to separate out direct holdings from pension and other tax-preferred asset holdings, but any misclassification towards direct holdings will also bias up the numerator of the gross capitalization ratio. The household sector of the FA is also a residual claimant on asset holdings, so any sectoral misallocation of a given asset holding towards households will introduce bias. It is also likely that in the current low-interest environment, the much lower interest earnings on checking and savings deposits are going unmeasured in the tax data, and to the extent those are more relevant for families outside the top 1 percent, their share of fixed-interest assets is being allocated to the top wealth families who have (quantitatively observable) interest. Ultimately, though, given the available data, we cannot point to any one explanation with certainty.[42]

III. Top Income Shares in Administrative and Survey Data

Income concentration and wealth concentration are both contentious issues, and many see the two measures as strongly correlated. Everyone seems to know that the rich are getting richer, whether we categorize them as rich by their income or their wealth. In some ways income

---

[42] Some of these issues may impart serious bias to the capitalization factors. The online appendix describes these issues in more detail, and some back-of-the-envelope calculations suggest that substantial biases in capitalization factors are likely introduced by these inconsistencies between micro income and macro balance sheet estimates.

concentration is a more straight-forward measure, because we can look directly at administrative data to gauge how top income shares are evolving over time, rather than (as in gross capitalization for wealth shares) requiring additional assumptions about the relationship between income and the value of assets generating that income. However, in another sense, the *concept* of income itself has changed in fairly dramatic ways over the period that top income shares are rising, and we will show that those conceptual changes are having a first-order impact on estimated top shares.

In this section we present our preferred estimates of top income shares, and, as with top wealth shares, we show how those preferred estimates compare and contrast to both published SCF and the administrative tax-based estimates. Our preferred top income share estimate is constructed by starting with SCF income measures, then adding components of NIPA personal income (PI) that are not measured in the SCF. The preferred measure shows slower *growth* in income concentration than the estimates in Piketty and Saez (2003, updated) based on administrative tax data, but unlike the top wealth shares, our preferred top income shares are also (modestly) lower and rising more slowly than published SCF estimates. We investigate the source of divergence in top income growth rates and levels, by once again constraining the SCF to conceptually match the administrative tax based estimates. Using that approach, we are able to confirm that the differentials in income concentration are not (at least to a first approximation) attributable to lack of population coverage at the very top or survey under-reporting in the SCF.

III.A. Preferred Estimates of Top Income Shares

In all of the estimates discussed here, top income shares in the United States are high and increasing over time. Figure 10A shows the estimated share of income received by the top 1 percent for the period 1988 to 2012 based on three different measures, and Figure 10B shows the same for the top 0.1 percent income shares. In general, the estimated top income shares based on administrative tax data (the grey lines) from Piketty and Saez (2003, updated) are higher and rising more rapidly than the top income shares in published SCF estimates (the red lines) and are also higher than those based on our preferred measure (the black lines).

The differences between the various estimated top income shares are, as with wealth shares, first-order. In 2012, our preferred estimate of the top 1 percent income share is just under 18 percent, while the administrative tax-based estimate is nearly 23 percent. The gap is

26

proportionally larger for the top 0.1 percent, and both gaps are increasing over time, though as with wealth, much of the increase in the top 1 percent income share is accounted for the top 0.1 percent income share. That is, the substantial income gains are occurring *within* the top 1 percent and not just for the 1 percent as a whole.

Our preferred measure for top income shares begins with the published SCF *Bulletin* concept and estimates. As with top wealth shares, the first adjustment on the income side is needed because the Forbes 400 is excluded from the SCF sample.[43] Although the Forbes 400 account for about 3 percent of total household sector net worth, the relationship between income and wealth is such that the Forbes groups accounts for a much smaller fraction of income, and thus adding them generally increases the average incomes of the top groups by a more modest amount. Thus, the estimated *shares* of income received by the top income groups are pushed up, but the effects are much more muted than for the top wealth shares.

The more substantial adjustments are to the SCF income concepts, and involve adding the in-kind transfers included in NIPA PI but not measured in the SCF survey. In particular, we add the value of employer-provided health insurance, the value of in-kind government transfers such as SNAP, and the value of Medicaid, Medicare, and other government health programs. Together, these incomes amounted to about 7 percent of NIPA PI in 1988, but had roughly doubled as a share of PI by 2012. That increasing share of total PI interacts with the casual observation that these forms of income are much less concentrated than the measured incomes, and that pulls down the preferred top shares in every year, but disproportionally more in recent years.[44] This is seen most clearly in the gaps between the published SCF income measure and our preferred measure: the modest but rising Forbes income share is pulling the two together, but the addition of in-kind incomes is larger, and on net, pushing the two apart.

III.B. Reconciling the Income Concentration Estimates

---

[43] The Forbes 400 is based on estimated wealth holdings, and there is no attempt to produce estimates of the income those wealth families earn during the year. We estimate their incomes using information on income and wealth for the top 0.1 percent of families in the SCF sample, for whom we know both income and wealth. For those top families, we compute the median ratio of income to wealth, and then we apply that ratio to the estimated Forbes wealth. Although the Forbes 400 account for about 3 percent of total wealth, our approach suggests they account for less than 1 percent of income.

[44] The distribution of the in-kind transfers is, as with our wealth imputations, driven by the available data in the SCF. Employer-provided health benefits are distributed across families based on their reported employer-sponsored health coverage, Medicare is distributed equally for eligible families, and the means-tested transfers are all distributed to the bottom 99 percent by income.

We approach the income shares reconciliation with the same basic starting point as we used for wealth shares. If the SCF sampling strategy does a good job capturing the top end of the income distribution, and SCF respondents do a good job reporting their incomes, what is causing the substantial divergence between estimated top income shares in the SCF-based preferred and administrative tax-based measures? Again, we constrain the SCF to be conceptually and empirically similar to the tax-based, and show that most of the divergence is eliminated. In particular, when we measure top income shares after constraining the SCF income concept to match the tax-based concept and we adjust the number of families in the top fractile to be consistent with tax unit counts, most of the level differences are eliminated, or at least brought within the range of SCF statistical confidence.

The effects of constraining the SCF-based preferred top income share estimates to be conceptually and empirically equivalent to the administrative tax-based estimates are shown in Figure 11A for the top 1 percent, and Figure 11B for the top 0.1 percent. The first adjustment, which involves moving from the preferred (black) lines to the orange lines labelled "Market Income, Families" is based on restricting the SCF income concept to match what is available in the tax data (see Table 2). That basically involves removing cash transfers—most notably Social Security benefits, but also other cash transfers—from the SCF income concept. Because those forms of income are disproportionately received by families in the bottom 99 percent by income, removing those income sources shifts the concentration numbers up. Because those sources of income are becoming increasingly important, the effects are larger in recent years. The quantitative effect of moving from the SCF *Bulletin* income measure to the more restrictive Market Income measure is to move the income concentration estimates further away from the preferred income measure, and for the same reasons.

The second reconciliation, as with the wealth shares, also uses the constrained Market Income concept, and further involves redefining how many families the top fractiles represent. Again, there are thirty percent more tax units than families in 2012, and thus the top 1 percent on a tax unit basis represents about 1.6 million families instead of the 1.2 million families in the top 1 percent using the SCF and preferred distributional measures. Adding the extra 400,000 families to the top 1 percent, and the extra 40,000 families to the top 0.1 percent, increases the top share estimates in a predictable and sizable way. The remaining differences between the top income shares in the constrained SCF (the blue lines) and administrative tax data are mostly about

28

volatility, and not levels per se. Further, the width of the confidence intervals shows how income variability and sampling interact, especially at the very top, to generate a wide confidence interval on estimated top shares.[45] Indeed, the point estimates for the constrained SCF top 1 percent income shares are actually above the administrative tax-based estimates, and basically the same for the top 0.1 percent.

III.C. Even More Comprehensive Incomes?

The steps taken to reconcile our preferred top income shares with the administrative tax-based estimates are suggestive of a broader question. What else is missing from an even more comprehensive income measure, and what might be the result of incorporating those other missing pieces into the analysis of top income shares? Figure 12 reinforces the fact that the more comprehensive income measure in our preferred top income shares diverge from the narrow administrative tax-based measures and the SCF *Bulletin* measure and that even our preferred measure is not complete. Even though the three income measures in the micro data all include something the PI measure does not—realized capital gains—even our most comprehensive income estimate is still less than the NIPA total.

The remaining divergence between NIPA PI and our preferred income measure involves a mix of imputations, known and unknown underreporting, and unreconciled conceptual discrepancies. It might be feasible in principle to produce distributional estimates for incomes such as imputed rent on owner-occupied housing or the value of in-kind financial services, using a data set like the SCF. One could also imagine rescaling the SCF reported incomes in categories for known underreporting for (say) proprietor's income, but that underreporting is also known to have a distributional component (small proprietors are worse when it comes to underreporting) that would have to be considered. There are also some tax-basis versus economic profit and rent adjustments incorporated into the NIPA, and one would have to work through those in order to align to the comprehensive PI measure. Although these various adjustments might affect estimated top shares, it is not clear in what direction. What is clear is that further adjustments

---

[45] The working paper version of this paper (Bricker, et al., 2015) has more details on the variability of top incomes, particularly with respect to the capital income shares. Saez and Zucman (2016) emphasize that the failure of the SCF to capture top capital incomes is indicative that the survey is missing top wealth holders, but we show there that most of the capital income at the top is captured as well after doing the same reconciliation exercise we do here for total incomes, and the remaining modest differences are likely associated with some of the reporting issues discussed in section I of this paper.

such as those should be done very carefully, and simply scaling available data to match the aggregates could bias the final answer.

IV. Concluding Remarks

Rising top wealth and income shares are often cited as a call to action by those who believe government can and should do more about inequality in terms of taxation, spending, regulation, and other market interventions. Rising inequality raises obvious normative concerns, and there is growing belief that recent macroeconomic instability and slow growth may be additional symptoms of the same underlying phenomenon.[46] Economists disagree about the fundamental causes of rising inequality, as some argue that the trends are associated with free market prices adjusting to equate supply and demand, while at the other extreme some argue that influence wielded by the already wealthy improves their market shares by changing the rules of the game.[47]

The preferred estimates for top shares presented here reflect what we think can be gleaned from the best available data sources, including administrative tax data, the SCF, and macro aggregates. The estimates agree with the widely-held view that inequality, at least as reflected in top wealth and income shares, has been rising in recent decades. However, the levels and trends in our preferred top share estimates are more muted than in recent studies which are based directly on administrative income tax data (Piketty and Saez, 2003, updated, and Saez and Zucman, 2016) but the levels and trends for top wealth shares are a bit larger than estimates based on estate tax data (Kopczuk and Saez, 2004).

Although the SCF makes it possible to inform and improve on direct estimates of top wealth and income shares derived from administrative tax data, the survey is still far from capturing comprehensive wealth and income measures. The SCF adds some government transfers into the tax-oriented income measures, but still misses employer-provided benefits, government in-kind (especially health) transfers, and other forms of income that are both substantial and growing over time. There are also direct analogs in terms of shortcomings in the wealth measures, as (for example) the value of most families' key retirement asset—Social

---

[46] For a somewhat contrary position on the economic stability effects, see Bordo and Meissner (2012).

[47] The view that markets underlie rising inequality is well described by Kaplan and Rauh (2010, 2013). See also Jones (2015) for a discussion of how competition among innovators affects top shares.

Security—is not measured as part of household net worth.[48] The effect of these omissions is important for understanding top shares, and even more important when looking at inequality across the entire distribution.

The reconciliations here cannot be extended back in time before the development of the modern SCF household survey, but the specific issues raised draw attention to how changes in government policies and market practices are affecting the measurement of top shares over time. In particular, although the administrative tax data makes it possible to show that top share families are getting increasingly large slices of a particular pie, the overall size of the pie being measured in those data is shrinking relative to more economically-meaningful concepts of wealth and income. The increasingly unmeasured part of the pie is not disappearing, but it is evolving. It may be difficult or even impossible to allocate the missing pieces in the very long historical series, thus any very long trends should also be viewed with an eye towards the conceptual divergence being driven by evolving government policy and economic institutions.

Building on the conceptual measurement theme, the reconciliation of top shares presented here speaks directly to the underlying impetus for—and possible approaches to—public policy towards wealth and income distribution. The failure to properly measure the effects of government policies and market practices that disproportionately benefit families in the middle and bottom of the wealth or income distribution leads directly to overstatement of top wealth and income shares. Policies and practices such as social insurance and government investment in human capital generate real benefits and the debate is thus properly focused on the distribution of those benefits. If we measure only the costs of such policies and practices, without measuring the benefits, it becomes more difficult to make the case in favor of such policies in policy debates.

---

[48] The Social Security actuaries estimate that the present value of future Social Security benefits for current workers is currently about $58 trillion, which is nearly the size of conventionally measured household sector net worth. Social Security wealth is also rising faster than other forms of wealth. Devlin-Foltz, Henriques, and Sabelhaus (2016) show how the distribution of Social Security wealth for near-retirees interacts with other forms of retirement wealth. Not surprisingly, given the progressive nature and cap on earnings in the benefit formula, Social Security wealth is disproportionately important for the bottom half of the wealth distribution.

References

Armour, Philip, Richard V. Burkhauser, and Jeff Larrimore. 2014. "Levels and Trends in U.S. Income and its Distribution: A Crosswalk from Market Income towards a Comprehensive Haig-Simons Income Approach," *Southern Economic Journal*, 81(2): 271-293.

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. (March)

Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri. 2015. "The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980-2010," Forthcoming in *Improving the Measurement of Consumer Expenditures*, Christopher D. Carroll, Thomas Crossley, and John Sabelhaus (Eds.), Studies in Income and Wealth, Volume 74. Cambridge, MA: National Bureau of Economic Research.

Auerbach, Alan J., and Kevin Hassett. 2015. "Capital Taxation in the Twenty-First Century." *American Economic Review*, 105(5): 38-42.

Board of Governors of the Federal Reserve System. 2015. "Financial Accounts of the United States". Retrieved from http://www.federalreserve.gov/releases/z1/current/z1.pdf.

Bordo, Michael D., and Christopher M. Meissner. 2012. "Does Inequality Lead to a Financial Crisis?" National Bureau of Economic Research Working Paper 17896. (March)

Bricker, Jesse, Lisa J. Dettling, Alice Henriques, Joanne W. Hsu, Kevin B. Moore, John Sabelhaus, Jeffrey Thompson, and Richard A .Windle. 2014. **"**Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances*,"* *Federal Reserve Bulletin*, 100(4): 1-40. (September)

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus. 2015. "Measuring Income and Wealth at the Top Using Administrative and Survey Data," Federal Reserve Board, *Finance and Economics Discussion Series*, 2015-30. (April)

Brown, Meta, Andrew Haughwout, Donghoon Lee, and Wilbert van der Klaauw. 2011 (revised 2013). "Do We Know What We Owe? A Comparison of Borrower- and Lender-Reported Consumer Debt," *Federal Reserve Bank of New York Staff Reports*, 523: 1-54.

Burkhauser, Richard V., Jeff Larrimore, and Kosali I. Simon. 2012. "A 'Second Opinion' on the Economic Health of the American Middle Class." *National Tax Journal,* 65(1): 7-32.

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): 371-388.

Debacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. 2013. "'Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns," *Brookings Papers on Economic Activity*, 67-122. (Spring)

Dettling, Lisa J., Sebastian Devlin-Foltz, Jacob Krimmel, Sarah J. Pack, and Jeffrey P. Thompson. 2015. "Comparing Micro and Macro Sources for Household Accounts in the United States: Evidence from the Survey of Consumer Finances," Federal Reserve Board, *Finance and Economics Discussion Series*, 2015-86. (June)

Devlin-Foltz, Sebastian, Alice M. Henriques, and John Sabelhaus. 2016. "Is the U.S. Retirement System Contributing to Rising Wealth Inequality?" *Russell Sage Journal of the Social Sciences*. (Forthcoming)

Fagereng, Andrea, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2016. "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality" *American Economic Review*, 106(5) (forthcoming).

Feenberg, Daniel, and Elizabeth Coutts. 1993. "An Introduction to the TAXSIM Model" *Journal of Policy Analysis and Management* (Winter) pp 189-194.

Feldstein, Martin. (2015) "The Uncounted Trillions in the Inequality Debate" http://www.wsj.com/articles/the-uncounted-trillions-in-the-inequality-debate-1450046600. Accessed December 13, 2015.

Greenwood, Daphne. 1983. "An Estimation of US Family Wealth and its Distribution from Microdata" *Review of Income and Wealth* (March) pp 23-44.

Guvenen, Fatih, Greg Kaplan, and Jae Song. 2014. "How Risky are Recessions for Top Earners," *American Economic Review,* 104(5): 148–153. (May)

Harwood, John. (2015) "10 questions with Bernie Sanders" http://www.cnbc.com/2015/05/26/10-questions-with-bernie-sanders.html (May 26, 2015).

Henriques, Alice M. (2013). "Are Homeowners in Denial about their House Values? Comparing Owner Perceptions with Transaction-Based Indexes," Federal Reserve Board, *Finance and Economics Discussion Series*, 2013-79.

Henriques, Alice M., and Joanne W. Hsu. 2014. "Analysis of Wealth using Micro and Macro Data: A Comparison of the Survey of Consumer Finances and Flow of Funds Accounts," in Jorgenson, Dale W., J. S. Landefeld and Paul Schreyer eds., *Measuring Economic Sustainability and Progress*, Studies in Income and Wealth, Volume 72. Cambridge, MA: National Bureau of Economic Research.

Jones, Charles I. 2015. "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality," *Journal of Economic Perspectives*, 29(1): 29-46. (Winter)

Kaplan, Steven N., and Joshua Rauh. 2013. "It's the Market: The Broad-Based Rise in The Return to Top Talent," *Journal of Economic Perspectives*, 27(3): 35-56. (Summer)

Kaplan, Steven N., and Joshua Rauh. 2010. "Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes?" *Review of Financial Studies,* 23(3): 1004-1050.

Kennickell, Arthur B., and R. Louise Woodburn. 1999. "Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth," *Review of Income and Wealth,* 45(2): 193-215.

Kopczuk, Wojciech. 2015. "What Do We Know About the Evolution of Top Wealth Shares in the United States?" *Journal of Economic Perspectives*, 29(1): 47-66. (Winter)

Kopczuk, Wojciech. 2015b. "Recent Evolution of Income and Wealth Inequality: Comments on Piketty's 'Capital in the Twenty-First Century'" *Tax Law Review*, 68(3): 545-562.

Kopczuk, Wojciech, and Emmanuel Saez. 2004. "Top Wealth Shares in the United States, 1916-2000" Evidence from Estate Tax Returns," *National Tax Journal*, 52:445-487.

Kuznets, Simon. 1953. "Shares of Upper Income Groups in Income and Savings." New York: National Bureau of Economic Research, 1953.

Mankiw, N. Gregory. 2015. "Yes, *r* > *g*. So What?" *American Economic Review*, 105(5): 43-47.

McKeever, Brice. 2015. "The Nonprofit Sector in Brief in 2015: Public Charities, Giving, and Volunteering." Washington DC: Urban Institute.

Moore, Kevin and Barry Johnson. 2008. "Differences in Income Estimates Derived from Survey and Tax Data," In *Proceedings of the Survey Research Methods Section, American Statistical Association.* American Statistical Association, pp. 1495-1503.

O'Muircheartaigh, Colm, Stephanie Eckman, and Charlene Weiss. 2002. "Traditional and Enhanced Field Listing for Probability Sampling," *American Association for Public Research 2002: Strengthening Our Community - Social Statistics Section*, 2563-2567.

Pareto, Vilfredo. 1896. *Cours d'Économie Politique*. Geneva: Droz.

Parker, Jonathan A., and Annette Vissing-Jorgensen. 2010. "The Increase in Income Cyclicality of High-Income Households and Its Relation to the Rise in Top Income Shares." *Brookings Papers on Economic Activity*, 1-55. (Fall)

Piketty, Thomas. 2015. "About *Capital in the Twenty-First Century*." *American Economic Review*, 105(5): 48-53.

Piketty, Thomas. 2014. *Capital in the 21st Century*. Harvard University Press. (April)

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118(1): 1-39.

Rajan, Raghuram G. 2010. *Fault Lines: How Hidden Fractures Still Threaten the World Economy*, Princeton University Press. (May)

Sabelhaus, John, David Johnson, Stephen Ash, Thesia Garner, John Greenlees, Steve Henderson, and David Swanson. 2015. "Is the Consumer Expenditure Survey Representative by Income?" Forthcoming in *Improving the Measurement of Consumer Expenditures*, Christopher D. Carroll, Thomas Crossley, and John Sabelhaus (Eds.), Studies in Income and Wealth, Volume 74. Cambridge, MA: National Bureau of Economic Research.

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," Forthcoming, *Quarterly Journal of Economics*. Accessible version: National Bureau of Economic Research Working Paper 20625. (October 2014)

Smeeding, Timothy M., and Jeffrey P. Thompson. 2011. "Recent Trends in Income Inequality: Labor, Wealth and More Complete Measures of Income," in Immervoll, Herwig, Andreas Peichl and Konstantinos Tatsiramos eds., *Who Loses in the Downturn? Economic Crisis, Employment and Income Distribution. Research in Labor Economics, Vol. 32.* Bingley, U.K.: Emerald, pp. 1-50.

Statistics of Income. 2012. *Individual Income Tax Returns.* Washington, DC: Internal Revenue Service.

Stiglitz, Joseph E. 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future*, W.W. Norton. (June)

Vermeulen, Phillip. 2014. "How Fat is the Top Tail of the Wealth Distribution?" Directorate General Research, Monetary Policy Research Division, European Central Bank. (September)

Weil, David N. 2015. "Capital and Wealth in the Twenty-First Century." *American Economic Review*, 105(5): 34-37.

Wolfers, Justin. 2015. "The Gains from the Economic Recovery are Still Limited to the Top One Percent" http://www.nytimes.com/2015/01/28/upshot/gains-from-economic-recovery-still-limited-to-top-one-percent.html.

Yellen, Janet. 2014. "Perspectives on Inequality and Opportunity from the Survey of Consumer Finances". Speech at the Conference on Economic Opportunity and Inequality, Federal Reserve Bank of Boston

Table 1. Measuring Household Wealth in the SCF and Capitalized Administrative Tax Data

	Survey of Consumer Finances (SCF)	Administrative Tax Data
Owner-Occupied Housing	Direct report on value of primary residence	Allocate FA housing total by capitalizing property tax paid on 1040 form (among itemizers)
+ Businesses	Direct report on value of businesses	Allocate FA total by capitalizing business income on form 1040
+ Non-retirement Financial	Direct report on value of checking accounts, savings accounts, CDs, mutual funds, directly-held stocks, annuities, trusts, managed accounts	Allocate FA total by capitalizing interest, non-taxable interest, dividend income on form 1040
- Mortgage Liabilities	Direct report on value of mortgage balances	Allocate FA outstanding mortgages by capitalizing mortgage interest deduction reported on 1040 form
- Other Liabilities	Direct report on value of lines of credit, car loans, education debt, credit cards, other consumer debt	Unallocated
+ DC Retirement	Direct report on value of IRAs, defined-contribution pensions on current job and past jobs	Allocate FA pension total using wages & pension payments (DB and DC not separated)
= Marketable Net Worth	**SCF _Bulletin_ concept**	
+ DB Retirement	Allocate FA DB total using wages and direct report on plan participation and benefits	Allocate FA pension total using wages & pension payments (DB and DC not separated)
= Private Net Worth	**Preferred estimate**	
+ Unallocated liabilities		**Saez and Zucman (2016)**

Table 2. Income Concepts and Data Sources

|	Survey of Consumer Finances (SCF)	Administrative Tax Data	National Income and Product Accounts
Wages and salaries, business income, interest and dividends paid directly to persons, other "market" incomes	Concepts generally consistent with income tax based reporting	Concepts generally consistent with income tax based reporting	Concepts generally consistent with income tax based reporting Adjusts for under reporting of proprietors' income, various rental and other capital income imputations
+ Realized capital gains	Concepts consistent with income tax based reporting	Concepts consistent with income tax based reporting	Capital gains not included in NIPA PI
+ Retirement income cash flow timing adjustment	Excludes employer contributions to and earnings on pension balances and Social Security Includes withdrawals and payments from retirement plans	Excludes employer contributions to and earnings on pension balances and Social Security Includes taxable withdrawals and payments from retirement plans	Adjust timing to match micro data concepts Effectively subtracts part of "net saving" in retirement plans from NIPA PI
= Market income		**Piketty and Saez (2003, updated)**	

Table 2. Income Concepts and Data Sources (Continued)

+ Government cash transfers	Social Security collected separately in work and pensions module and as a component of total in income module	No information on non-taxable cash transfers	Includes all government cash transfers
	SSI, TANF, other cash transfers collected in income module (known to be somewhat underreported, as in other surveys)		
= Total cash income	**SCF *Bulletin* concept**		
+ In-kind transfers and benefits	No direct information on employer-provided health, government-provided health, or other in-kind benefits	No direct information on employer-provided health, government-provided health, or other in-kind benefits	Includes all employer-provided health, government health, and other government in-kind benefits
	Distribute between top shares using proportionality		
= Total cash and in-kind income	**Preferred measure**		PI less imputations and partially adjusted for retirement income timing

**Table 3. Coverage and Unit of Analysis Across Income and Wealth Data Sets**

	Survey of Consumer Finances (SCF)	Administrative Tax Data	National Income and Product Accounts (NIPA)	Financial Accounts (FA)
Unit of Analysis	Families	Tax Units	Aggregate	Aggregate
Coverage	Entire non-institutional population	Tax-filing population only	Households and non-profit institutions	Households and non-profit institutions
	Corrects for low participation at high end using list sample	Supplement with information on non-filers from other data sources		Possible to separate out non-profit holdings of real estate
	Excludes Forbes 400			

Figure 1. Income Densities for Top Strata SCF Respondents and Non-Respondents

A. Total Income, 2009-2011
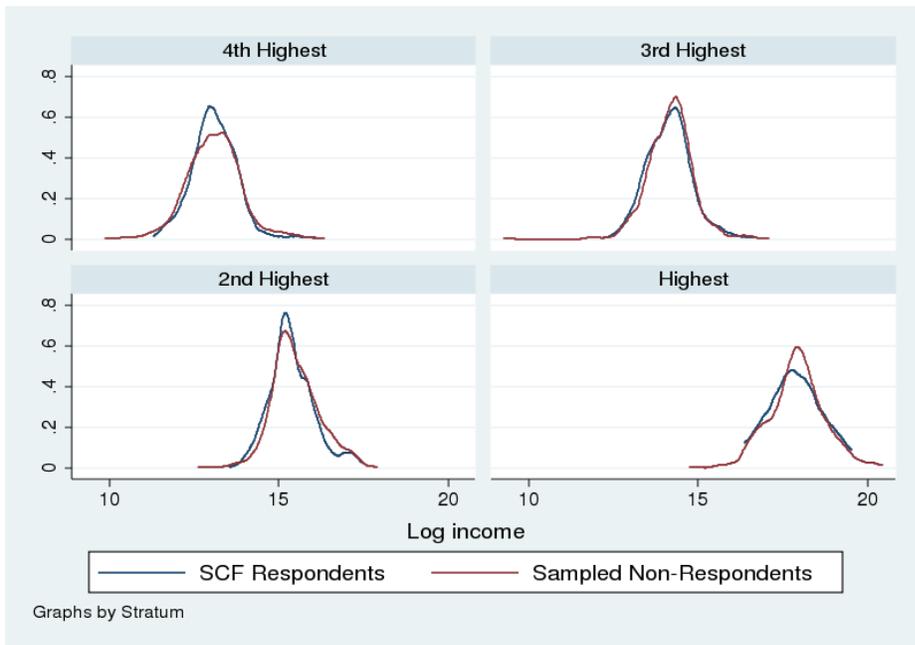
B. Capital Income, 2009-2011
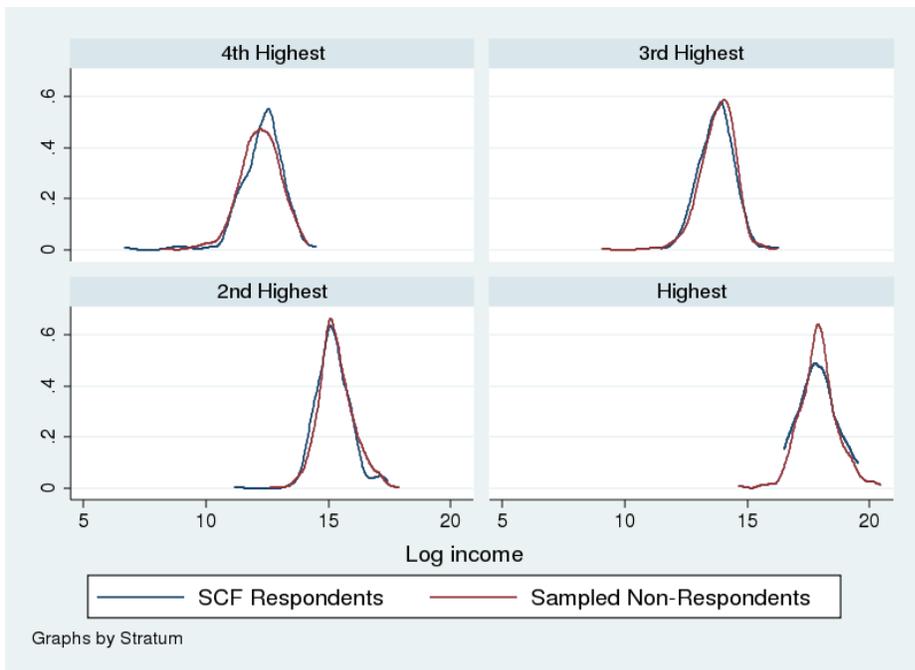
Note: Incomes are 3-year averages and include capital gains. Sample includes the 4 highest strata, which fully encompasses the top 1% of the predicted wealth distribution. Incomes include capital gains. Data for the calendar years 2009-2011 are associated with the sampling for the 2013 SCF. Data source: Statistics of Income, Individual Sole Proprietorship (INSOLE).

Figure 2. Mean Incomes for Top Strata SCF Respondents and Non-Respondents

A. Total Income, 2009-2011
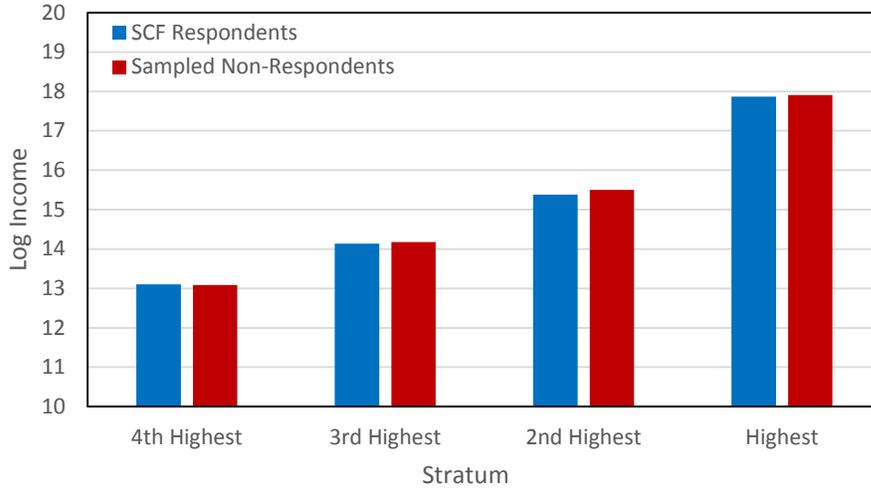
B. Capital Income, 2009-2011
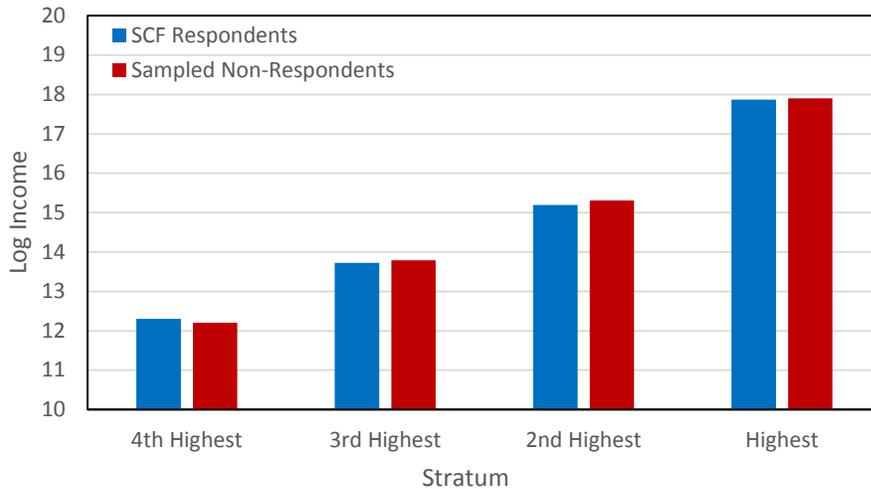
Note: Incomes are 3-year averages include capital gains. Sample includes the 4 highest strata, which fully encompasses the top 1% of the predicted wealth distribution. Incomes include capital gains. Data for the calendar year 2009-2011 are associated with the sampling for the 2013 SCF. Data source: Statistics of Income, Individual Sole Proprietorship (INSOLE) data.

Figure 3. Pre-Survey Income Volatility of Top Strata SCF Respondents, Non-Respondents

A. Percent Change in Total Income, 2010-2011
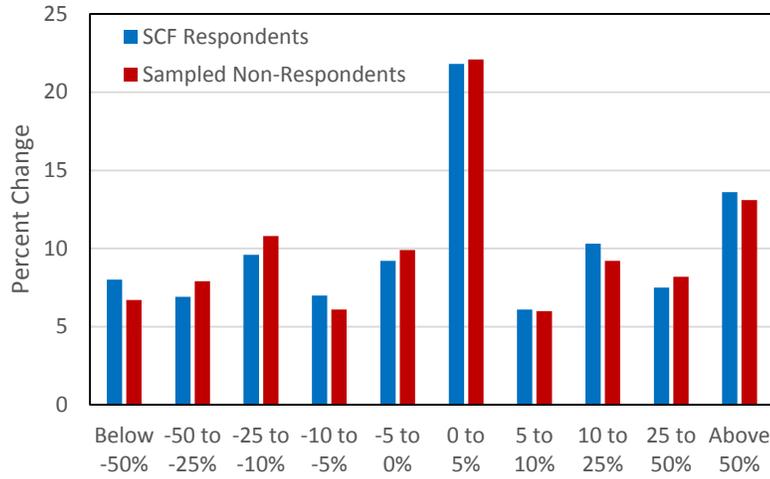
B. Percent Change in Capital Income, 2010-2011
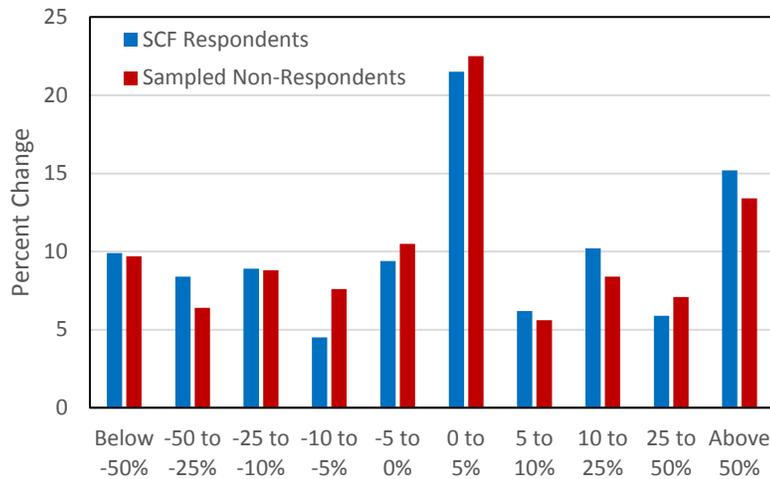
Note: Sample includes the 4 highest strata, which fully encompasses the top 1% of the predicted wealth distribution. Incomes include capital gains. Data for the pre-survey calendar years 2010 and 2011 are associated with the sampling for the 2013 SCF. Data source: Statistics of Income, Individual Sole Proprietorship (INSOLE) data.

Figure 4. Income Change for Families with AGI Greater than $500,000, 2011-2012
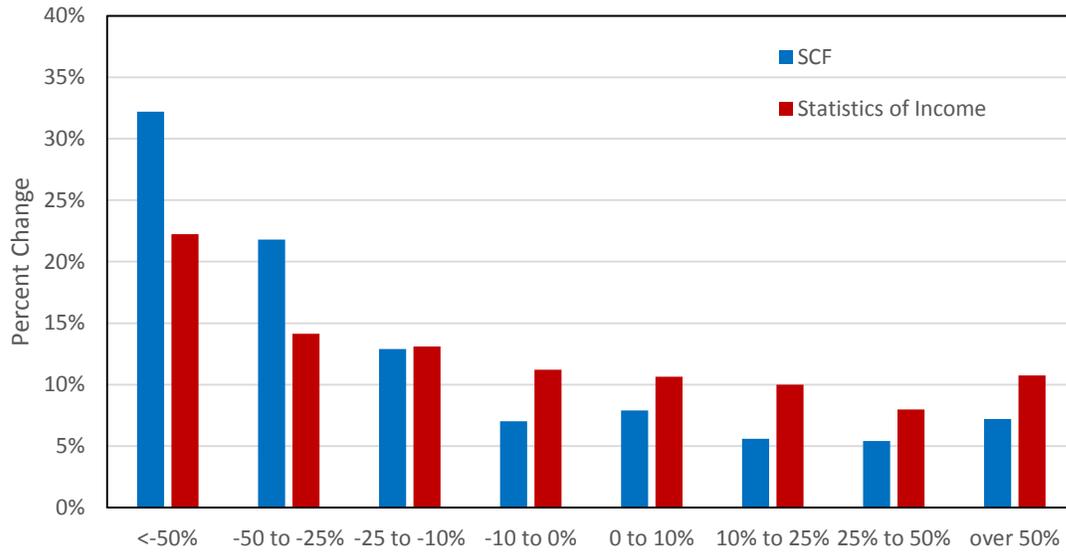
Notes: The red bars show the change in Adjusted Gross Income (AGI) from 2011 to 2012 among all tax returns with AGI over $500,000 in 2011 (according to unpublished SOI tabulations). The blue bars show the change in AGI from 2011 to 2012 among sampled SCF households with AGI over $500,000 in the INSOLE data. For SCF households, changes are computed using AGI provided by SOI in 2011 and AGI computed with NBER TAXSIM (Feenberg and Coutts, 1993) using household income from the 2013 SCF. Data sources: Federal Reserve Board, 2013 Survey of Consumer Finances (SCF); Statistics of Income, 2011-2012 Individual Sole Proprietorship (INSOLE; tabulations by Michael Parisi).

Figure 5. Top Wealth Shares, 1989-2013

A. Top 1% Wealth Shares

Legend:
- Preferred Wealth Measure
- SCF Bulletin Wealth
- Income Tax Data

B. Top 0.1% Wealth Shares

Legend:
- Preferred Wealth Measure
- SCF Bulletin Wealth
- Income Tax Data

Sources: Survey of Consumer Finances (SCF) and Saez and Zucman (2016). Preferred wealth measure is SCF Bulletin, augmented by DB pension estimate and Forbes 400 adjustment. See Section II and on-line appendix for details on each estimate .

Figure 6. Reconciling Top Wealth Shares

A. Top 1% Wealth Share

- Preferred Wealth
- Preferred Wealth, FA Concepts & Values
- Preferred Wealth, FA Concepts & Values, Tax Units
- Capitalized Tax Income

B. Top 0.1% Wealth Share

- Preferred Wealth
- Preferred Wealth, FA Concepts & Values
- Preferred Wealth, FA Concepts & Values, Tax Units
- Capitalized Tax Income

Sources: Survey of Consumer Finances (SCF) and Saez and Zucman (2016). See Sections I, II and on-line appendix for more details on wealth concepts. Shaded area represents 95% confidence interval based on sampling and imputation variance.

Figure 7. Top 0.1% Wealth Composition, SCF vs. Capitalized Administrative Income Data

A. Top 0.1% in SCF

Legend: ■ Housing ■ Pension ■ Equity + Business ■ Fixed Income Assets

B. Top 0.1% in Capitalized Income

Legend: ■ Housing ■ Pension ■ Equity + Business ■ Fixed Income Assets
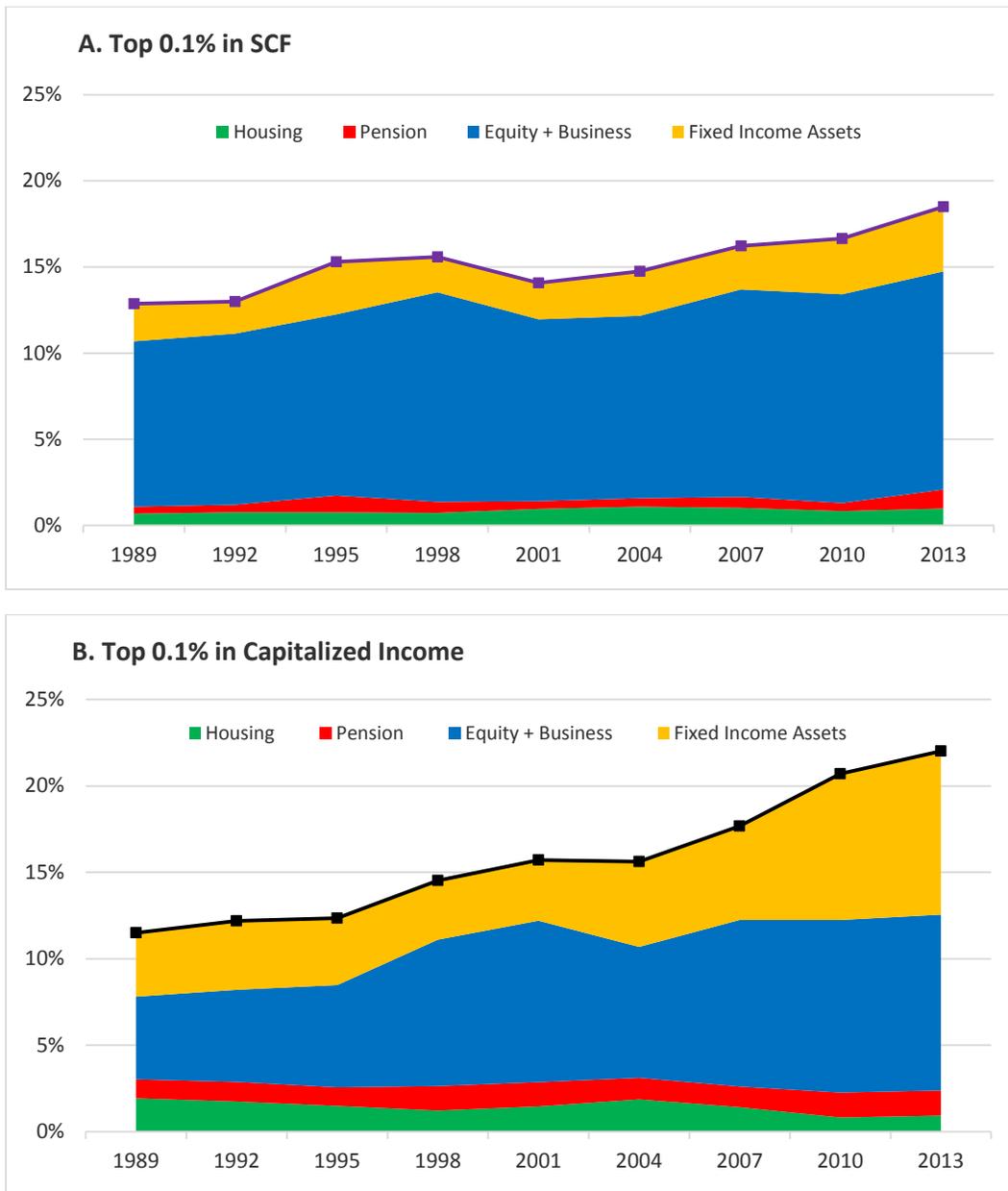
Notes: In panel A, we assume that the assets of Forbes 400, omitted from the SCF, are split proportional to the assets of the top 0.01% according to Saez and Zucman (2016). Administrative data are through 2012, though labelled as 2013. For each year on the x-axis, share of wealth held by the top 0.1 percent of families is broken into four general types of wealth: wealth from housing, from pensions, from corporate equities and private businesses, and from fixed income assets. Fixed income assets are bonds, CDs, savings accounts, and money market funds. Equities and businesses include the net worth of corporate equities, S-Corps, partnerships, and sole proprietorships. The cumulative height of the SCF top 0.1 percent is the SCF net worth benchmarked to FA values, adjusted for tax-units, and including an estimate of the Forbes 400 (i.e. the purple line in figure 6, panel B). Data sources: Federal Reserve Board, Survey of Consumer Finances (SCF); and Saez and Zucman (2016), Appendix Table B5b.

46

Figure 8. Heterogeneity in Capitalization Factors to Infer Fixed Income Assets
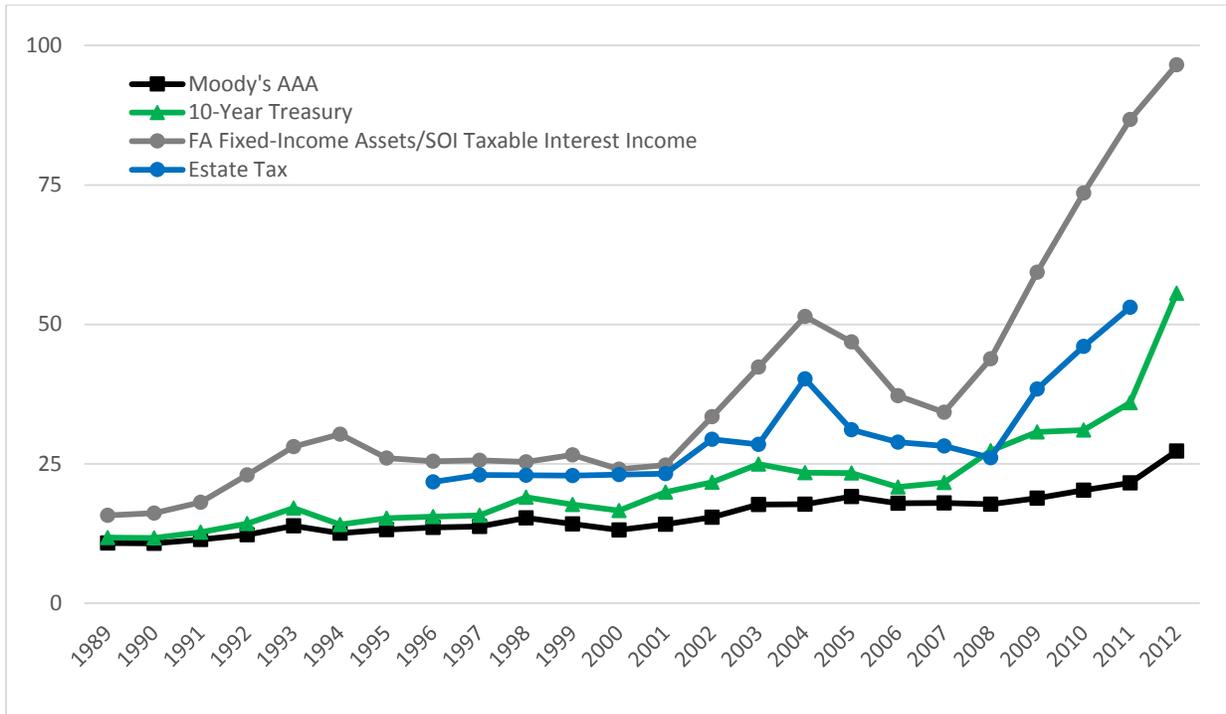
Notes: In a gross capitalization model, the capitalization factor for taxable interest income is the rate at which interest income will be grossed-up to infer fixed-income assets. The Moody's AAA line shows the inverse of the interest rate of the Moody's AAA corporate bond rate (seasoned issue, all industry, annualized) from the Federal Reserve H.15 data series. The 10-year Treasury line shows the inverse of the 10-year Treasury yield, annualized. The blue line shows the estimated rate of return on fixed-income assets among a set of matched estate tax-to-income tax filers (with more than $20 million in estate tax assets) from Saez and Zucman (2016) Appendix Table C6b. The grey line shows the ratio of the stock of fixed income assets in the Financial Accounts (table B.101) to SOI taxable interest income. Data sources: Saez and Zucman (2016), Appendix Table A11 and Appendix Table C6b; Moody's; United States Treasury.

**Figure 9. Top Wealth Shares Using Alternative Capitalization Factors**
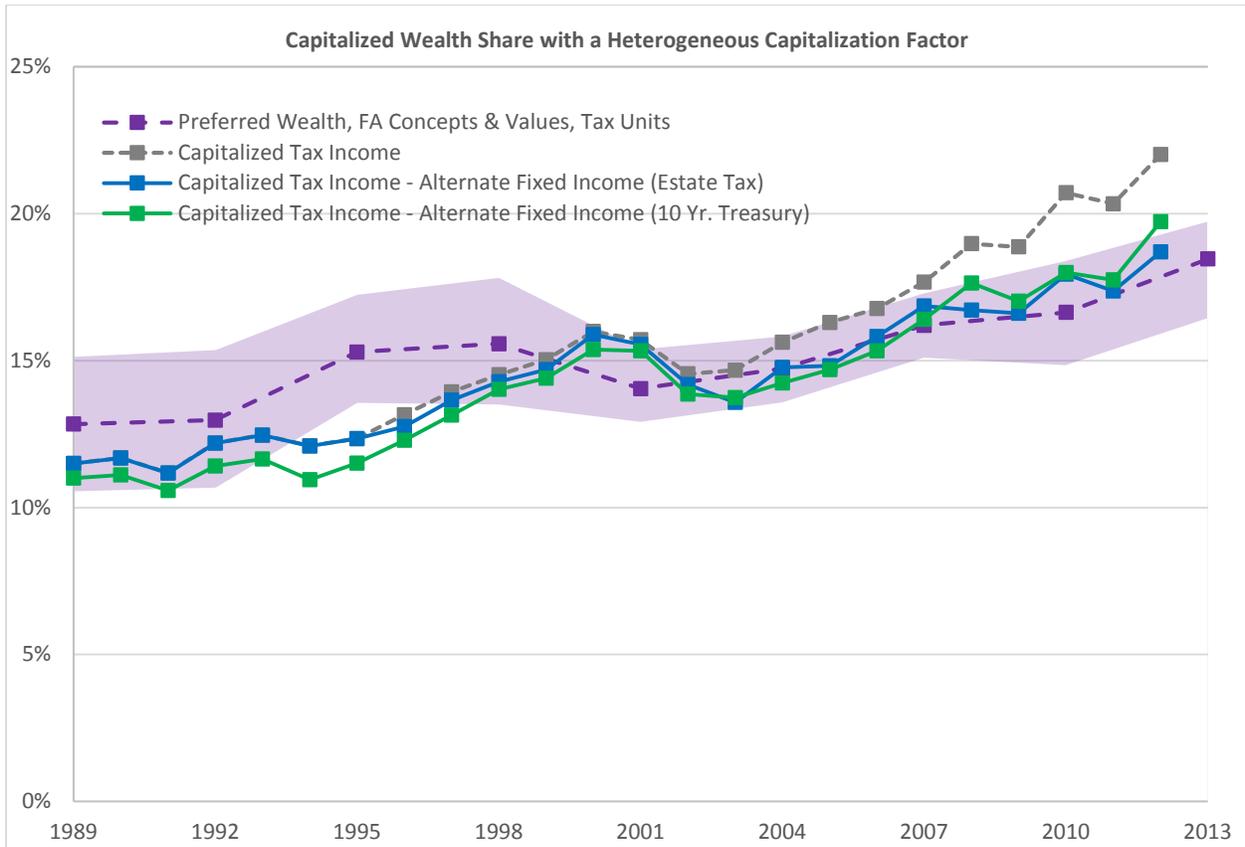
Notes: the purple line shows the share of wealth held by the top 0.1 percent in the preferred wealth measure (adjusted to match FA assets, adjusting away from families and toward tax-units) and is identical to the purple line in figure 6, panel B. The dashed grey line shows the share of wealth held by the top 0.1 percent in the capitalized income data (Saez and Zucman, 2016) and is identical to the dashed grey line from figure 6, panel B. The green line shows a version of the grey line where fixed income assets for the top 1 percent of income earners are generated by the inverse of the rate of return on the 10-year Treasury (the green line in figure 8). In the green line, the fixed income assets for the bottom 99 percent of income earners are still generated, as in the dashed grey line, by the ratio of fixed-income assets in the FA to SOI taxable income (the grey line in figure 8). Underlying data for the green line can be found in Saez and Zucman's (2016) Appendix Table B40. The solid blue line shows the estimated top 0.1 percent wealth share when fixed income is capitalized based on the rate of return on fixed income assets among estate tax filers with more than $20,000,000 in assets (see Saez and Zucman (2016) Appendix Table B36b). Data sources: Federal Reserve Board, Survey of Consumer Finances (SCF); Saez and Zucman (2016), Appendix Tables B1, B36b, and B40.

Figure 10. Top Income Shares, 1988-2012

A. Top 1% Income Shares

Preferred Income Measure
SCF Bulletin Income
Income Tax Data

B. Top 0.1% Income Shares

Preferred Income Measure
SCF Bulletin Income
Income Tax Data

Sources: Survey of Consumer Finances (SCF) and Piketty and Saez (2003 + updates). SCF incomes are collected for the calendar year prior to each triennial survey. See Sections I, III and on-line apendix for details on Income Tax, SCF Bulletin, and Preferred Income concepts.

Figure 11. Reconciling Top Income Shares

A. Top 1% Total Income Shares

- Preferred Income Measure
- Market Income, Families
- Market Income, Tax Units
- Income Tax Data

B. Top 0.1% Total Income Shares

- Preferred Income Measure
- Market Income, Families
- Market Income, Tax Units
- Income Tax Data

Sources: Survey of Consumer Finances (SCF) and Piketty and Saez (2003 + updates). SCF incomes are collected for the calendar year prior to each triennial survey. See on-line appendix for details on income tax, SCF Market income, and Preferred Income concepts.

Figure 12. Aggregate Income Concepts and Estimates

Sources: Bureau of Economic Analysis; Piketty and Saez (2003 + updates); Federal Reserve Board, Survey of Consumer Finances. See on-line appendix for more details.