Journal of the Statistical and Social Inquiry Society of Ireland
Vol. L

Top Incomes in Ireland: Reconciling Evidence from Tax Records and Household Survey Data[1]

T. Callan
Geary Institute, UCD, Institute of Labor Economics and Callan Economics

K. Doorley
ESRI, Trinity College Dublin and Institute of Labour Economics

A. McTague
ESRI

(read before the Society, 22nd October 2020)

Abstract: There are substantial divergences between the pictures of income inequality provided by household survey data for the full population and analyses based on administrative data generated by the income tax system focusing on the share of income accruing to the top of the income distribution. Our analysis, similar to international work on this topic, finds that differences in concepts and measures play a very substantial role in explaining these differences in the Irish context. When analyses of household data and administrative data are aligned using the same concepts and measures, the contrasts are much reduced. Average incomes for the top 1% of the population do appear to be higher in tax return data than in SILC, again in line with international findings. There is a strong case for examining potential adjustments to survey data to ensure better representation of income levels at the very top of the income distribution. Such adjustments combine information from tax returns and household surveys to give better representation of top incomes, while retaining the advantages of household data's perspectives on low incomes and household context.

Keywords: inequality, income distribution
JELs: C80, D31

1. INTRODUCTION

Concern about the consequences and causes of economic inequality has become widespread. Academic research which has analysed the extent and impact of inequality has developed and expanded: what was previously seen as a "niche" area has become mainstream. Key publications which would formerly have had a limited audience have become best sellers and award winners (e.g., Piketty, 2014, Atkinson, 2018). International institutions which would have been regarded as favouring free market policies over distributional concerns are now taking a very different view. For example, in series of major reports, the OECD (2008, 2011, 2015) has documented extensively the widespread (though not universal) rise in income inequality in many OECD countries. It has also documented the consequences of rising inequality and explored policy options to attenuate or reverse this trend. Similarly, the International Monetary Fund now places great emphasis on inclusive growth, with inequality being seen as harmful to growth. For example, Aiyar and Ebeke (2019) argue that "the negative impact of income inequality on growth is higher the lower is intergenerational mobility". At national level too, inequality has become a key theme in public and political debate.

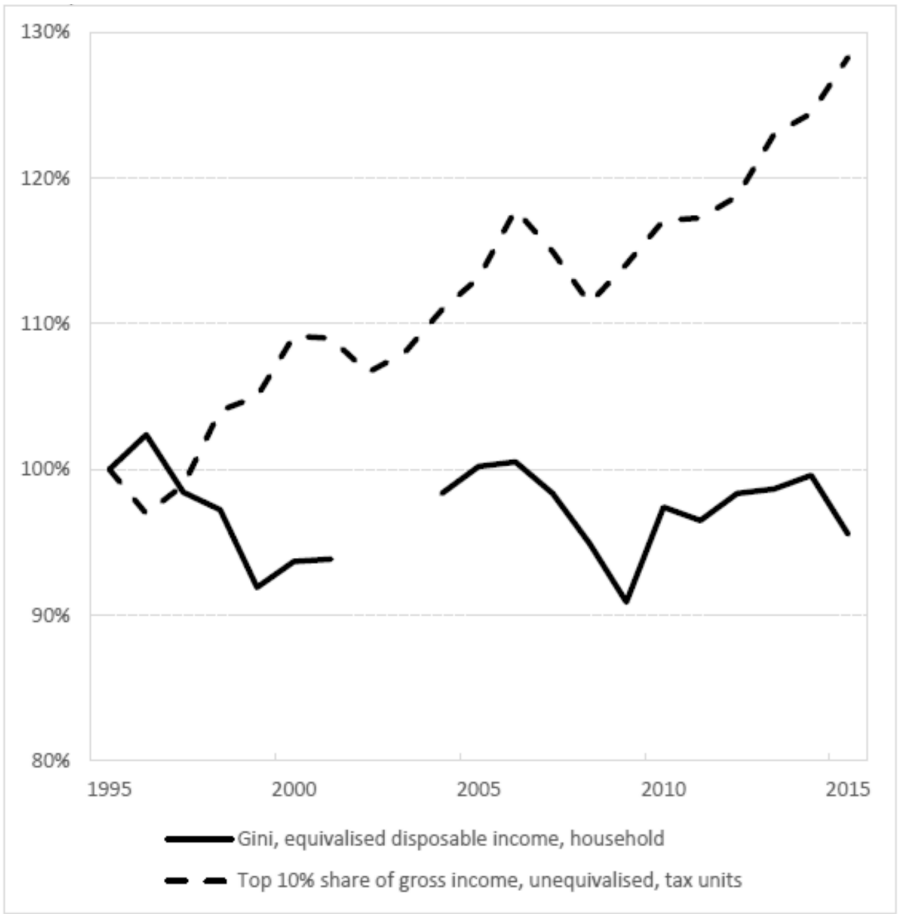
Underlying all of these developments are two quite distinct forms of data – household surveys and tax records– which sometimes seem to tell conflicting stories about trends in inequality.[1, 2] For recent overviews of the different strengths and challenges involved in these data sources, see Lustig (2018) and Kennedy (2019), on which the following discussion draws. Most studies and official statistics on income distribution (including the OECD's Income Distribution Database and the Luxembourg Income Study) are based on household surveys. These have a great number of strengths, as catalogued by Burkhauser et al. (2018): They focus on the household as an income-sharing unit; they incorporate adjustments using an equivalence scale to take account of how needs differ for households of different sizes and composition; and they take account of the impact of taxes and transfers on disposable income, which is therefore more closely linked with living standards. Household survey also allow a wide range of inequality measures, including many which take account of the distribution over the full range of income, such as the Gini coefficient; and they allow analysis at the level of persons (irrespective of age) rather than simply at household level.

[1] We thank Brian Nolan for helpful discussions and for clarifications of the methods used in the World Inequality Database.
[2] Burkhauser et al. (2017a) note that "the two main sources of information – household survey and personal tax return data – provide very different estimates of inequality trends".

There is, however, one major challenge for household surveys in measuring the income distribution. It is generally expected that household surveys will have difficulty in capturing accurately the very highest incomes – either because those with the highest incomes are less likely to respond to surveys – what Lustig (2018) terms "the missing rich" - or because the incomes they report in surveys do not match with the full income reported to the tax authorities. Tax return data has a comparative advantage in this area – while at low incomes a tax return may not be required, it is compulsory for others, meaning that tax returns provide full coverage of the relevant population for top incomes. Furthermore, the powers and penalties available to the tax authorities mean that the returns filed by taxpayers. Other advantages of tax return data noted by Kennedy (2019) include their large scale, allowing analysis of subgroups for which survey data would not have enough cases; and a strong longitudinal dimension, with taxpayers reporting incomes year after year.

There are, however, other trade-offs here. Tax return data does not provide the same breadth of demographic information as a household survey, tend to underrepresent lower income units, and may not include untaxed transfers. (Kennedy, 2019, Table 1.1). The income-sharing unit is effectively the tax unit, and this is often also the unit of analysis.2[3] Most analysis of tax return data is, of necessity, based on gross incomes rather than disposable incomes. Burkhauser et al. also note that inequality measures based on tax return data is restricted to "summary measures (top income shares) that do not incorporate differences across the full income range". This is because the tax return data do not provide a full picture for those on lower incomes. Instead, the top income methodology uses tax return data in combination with aggregate data on population and national income accounts to generate estimates of the shares of the top 1%, top 10% and so on.

Figure 1: Contrasting Perspectives on Income Inequality, Ireland 1995-2015: Gini coefficient from Household Surveys vs Top Decile Share from Tax Return Data)



Note: Index: 1995=100 for both Gini and top income share.
Sources: Top 10% share: World Inequality Database (wid.world), analysing "fiscal income" based on tax returns, after Nolan (2018) Gini coefficient: Equivalised disposable income at household level 1995-2001, Living in Ireland Survey analysed by Callan and Savage (2017) 2004-2015: CSO Survey on Income and Living Conditions

[3] The World Inequality Database now produces some statistics based on an equal split of income between adults in the tax unit, but a household level analysis is not possible from tax record data.

While these considerations may appear rather technical, they have important implications for our understanding of trends in inequality, as illustrated in Figure 1. This shows two of the most commonly quoted statistics on inequality from each of the distinct sources, household surveys and tax returns. The Gini coefficient is perhaps the most frequently quoted statistic on income distribution based on household surveys. On the tax return side, a Gini coefficient for the full population is not possible: tax return data do not cover the full population, and microdata for the full population are needed to calculate the Gini. However, tax return data, in combination with information from national accounts on aggregate income, allow the estimation of top income shares, as tax return data have comprehensive information on the upper reaches of the income distribution. Thus, top income shares, such as the share of the top 10% (or top 1%) are the statistics most usually provided. Figure 1 shows that the trends identified by these statistics are quite different. The Gini coefficient varies within a narrow range and is broadly stable. The top decile income share, by contrast, rises strongly over this period, with only a temporary recession-related fall in the 2008-2010 period. By 2015 the top decile share is almost 30 per cent higher than in 1995, while the Gini coefficient is close to its 1995 level. The "top incomes" statistics, based on gross incomes, unequivalised, at tax unit level suggest a strong rise in inequality over the period. By contrast, the household survey statistics – based on disposable incomes, equivalised, and at household level – suggest a broad stability or slight fall in inequality over the full period.

What gives rise to this sharp contrast? Is it the fact that top incomes analysis has better coverage of incomes at the very top of the income distribution? Or does the contrast arise more from technical factors – differences in income concept (gross versus disposable) or income unit (household versus tax unit) or from the contrast between equivalised and unequivalised incomes? There is a complex relationship between commonly reported measures of inequality based on household surveys and the usual results reported in the top incomes literature. In this paper we first review what international research has found in this area (Section 2). We then turn to a closer examination of the Irish situation (Section 3). We re-analyse household survey data in a way which aligns the concepts and measures used with those in the top incomes literature. Our analysis finds that the apparent gap between the two methods is much less than what commonly used headline measures would suggest. We also indicate the scale of the impact of the different technical factors involved in driving apart the survey and tax return measures of inequality. Section 4 provides evidence on the profile of top decile income units and shows the extent to which differences in income definitions and measurement choices affect the composition of the top 10 per cent group. The final section draws together the main conclusions and some issues for further research.

2. INTERNATIONAL EVIDENCE

Household surveys have for many years been the first port of call for broadly-based analysis of inequality levels, trends and international comparisons. Within the European Union, the standard source (since 2003/4) is the EU Survey on Income and Living Conditions, which was designed to provide comprehensive and comparable statistics on income distribution and risks of poverty across the EU. This provides the basis for a wide range of indicators of poverty and social exclusion which are used in a policy setting at national and EU level. Broader international comparisons have been made possible by the Luxembourg Income Study, which harmonises household survey data from a wide range of countries. The OECD's Income Distribution Database, on which its substantial analyses of inequality rest, is also based on household survey data.

There have been ongoing concerns about whether household surveys provide good representation of incomes at the top of the distribution. Partly for this reason, and partly because tax records are available over a much longer period of time, there has been a revival of the methods used by Kuznets (1953) to examine shares of income accruing to those at the top of the income distribution – often the top 1%, top 5% or top 10%. The recent revival of interest led to a succession of studies has constructed top income share series for a large number of countries (starting with Piketty (2001) for France, Piketty and Saez (2003) for the US, Atkinson (2002 and 2005) for the UK, and the two multi-country volumes edited by Atkinson and Piketty 2007 and 2010). The World Inequality Database (Alvaredo et al., 2018) now provides both a series of publications and an online platform which allows users to examine top incomes in a historical and international context. Ireland is included in this database thanks to contributions by Nolan (2007, 2018). For reviews of this top income literature see Atkinson, Piketty and Saez (2011), Alvaredo et al. (2013) and Alvaredo and García-Peñalosa, (2018).

One thread which emerged in the debate surrounding this work was whether this new perspective from tax return data cast doubt on the picture of inequality which had emerged from previous work using household survey data. For the US and the UK, a small number of studies have examined this issue in depth. Burkhauser et al. (2012) use income data from the US Current Population Survey – a household-based survey - to derive estimates of top income shares which are comparable to those based on tax return data. The methods and definitions developed by Piketty and Saez (2003) and others for use with tax return data were applied by Burkhauser et al. to the household survey data. Resultant estimates for top income shares, in terms of levels and trends, were found by Burkhauser et

al. to be "nearly identical" to results from the tax return data for groups in the richest decile, with the exception of the richest 1 per cent. "Even for estimates of the share held by the top 1 per cent the two data sources are broadly in agreement about trends over much of the past 40 years. It is only during a six-year period in the late 1990s that the trends diverge for reasons that are not easily explained by changes in the nature of the two data sources". Yonzan et al. (2018) perform a similar comparison based on the Current Population Survey for 2013 and also conclude that there is substantial difference of income only within the very top percentile, and almost all of this is driven by the non-labour portion of income – business income and capital income.

In the UK, Burkhauser et al (2018b) note that household surveys do not capture income at the extreme top of the income distribution very well. They argue that using tax return data in combination with survey data is a potential approach to address this problem because tax data are likely to have much better coverage of top incomes. They examine "a pioneering variant of this approach", the SPI3 adjustment used in Britain's official income distribution statistics since 1992.[4] However, they argue that a new approach, in the same tradition, can improve data quality at the top of the income distribution. Burkhauser et al. (2018a) provides the first systematic comparison of top income shares derived from household survey (Households Below Average Income, based on the Family Resources Survey) and tax return data (the Survey of Personal Incomes, SPI). Having reconciled the definitions, they find evidence that very high incomes are not well covered in the survey. The nature of the adjustment they propose is built on careful investigation of the nature of the gaps between top incomes as recorded in the household survey and in tax returns. Ultimately they find evidence favouring an approach which defines 50 income groups, each containing 0.1% of the total population, for the top 5% of income units, and replaces survey-measured incomes with average tax return incomes for each of these high-income groups. It is noteworthy that in the UK, the Office for National Statistics intends to finalise a top income adjustment and include it in regular Office for National Statistics (ONS) releases from 2020 onwards (Shine et al., 2019). Top income adjustments have also been made to SILC data. Bartels and Metzing (2019) estimate a Pareto distribution for the top of the income distribution, based on tax return data. They then use these Pareto-estimated top incomes to replace the incomes of the top 1% of tax units in SILC.

In this paper, we follow the first half of the Burkhauser et al (2018a) strategy.[5] We perform a detailed comparison of top income shares for Ireland, comparing the results from tax returns in the World Inequality Database (Nolan, 2018) with appropriately constructed measures based on EU SILC. Only such a comparison can find the extent and location of any potential income undercoverage.[6] Is it restricted to the top 1% or does it extend more broadly? Does it affect the broad trend in inequality? These are the issues to which we now turn.

3. TOP INCOMES IN IRELAND: RECONCILING SURVEY ESTIMATES AND TAX RECORDS

As seen in Figure 1, top income analysis for Ireland indicates a sharp rise in inequality, while the Gini coefficient for household disposable income (adjusted for household size and composition using an equivalence scale) is broadly stable. This is not, however, a "like for like" comparison. There are major differences in the concepts and measures used, as well as differences in the data sources. To what extent is this observed difference due simply to the combined effects of these methodological differences, and to what extent does it reflect differences in the coverage of top-level incomes?

In this section, we are able to answer this question by re-analysing the household survey data, taking one step at a time from the concepts and measures used in top income analysis to the standard household survey measures.[7] We focus initially on the income share of the top 10% of tax units – one of the key measures used in top income analysis. We also note the impact on the Gini coefficient, where this is possible. Our analysis uses SWITCH, the ESRI's tax-benefit model. The flexibility provided by the SWITCH modelling process allows us to group individuals within households into tax units. A tax unit is defined as a single person, or a married couple, together with their dependent children (aged under 18, and not in employment, in which case they would be counted as independent tax units). At present, we must use simulated tax liabilities and simulated welfare payments in our analysis. While this has some drawbacks – chiefly the fact that actual tax liabilities at high incomes tend to be lower than those simulated by SWITCH[8] – this has little impact on the key comparisons of top income shares for pre-tax income which follow.

[4] SPI stand for Survey of Personal Incomes, which is a very large-scale sample of tax returns constructed by the UK revenue authorities.

[5] The second part of the Burkhauser et al. strategy is to adjust household survey data based on information from tax return data. We discuss this aspect in our conclusions, in the context of future research possibilities.

[6] The term "undercoverage" includes any mixture of underrepresentation of top incomes in the sample (e.g., through increased non-response at very high income levels) and/or underreporting of incomes by survey participants.

[7] Burkhauser et al. (2018a) refer to this as "cross walking" from survey-based definitions to the taxa data definitions.

[8] SWITCH models the main personal credits, and reliefs in respect of pension contributions, mortgage interest relief and health insurance. It does not model the full range of reliefs, some of which are highly concentrated on top income earners. (See

The income concept used in the World Inequality Database estimates for Ireland and many other countries is "fiscal income". A clear understanding of this concept, and how it relates to the disposable income measure used in much of the analysis of household income distributions is essential in reconciling apparently conflicting results on inequality levels and trends.

"Fiscal income" refers to income as measured by the fiscal authorities. Irish revenue authorities state that in their Income Distribution Statistics, which are used by WID

> "Total" income is the total income of taxpayers from all sources as estimated in accordance with the provisions of the Income Tax Acts. It is net of such items as capital allowances, allowable interest which is not subject to relief at the standard rate, losses, allowable expenses, retirement annuities and superannuation contributions.
>
> Revenue (2012)

Thus, non-taxable sources of income – including many welfare benefits – are excluded from fiscal income. Moreover, fiscal income excludes retirement annuities (largely paid by the self-employed) and superannuation contributions. The differences between this fiscal income concept, and the concept of disposable income as used by SILC in the construction of Ireland's national poverty indicators, are summarised in Table 1.

Table 1: Differences between Disposable Income and Fiscal Income

Fiscal income
less
Income tax
Employee PRSI
Universal Social Charge
plus
Non-taxable social welfare payments
Employee superannuation contributions
Retirement annuity contributions
Equals
Disposable income

These are major differences in the basic income concept, which can be expected to have a substantial impact on measures of inequality. Income taxes (including USC and employee PRSI) are strongly progressive in most countries, and particularly so in Ireland (Roantree, 2020). Exclusion of non-taxable social welfare payments also omits a further progressive impact of the transfer system. One countervailing factor is the exclusion from fiscal income of superannuation and retirement annuity contributions. This is likely to reduce measured inequality in fiscal income, as studies of the distribution of tax relief on superannuation and retirement annuity contributions (Doorley et al, 2017) suggest that these are concentrated towards higher earners.

We implement this definition of fiscal income on SILC data for 2013, 2014 and 2015 using SWITCH, the ESRI tax benefit model. Before examining the implications in terms of top income shares or other distributional measures, it is useful to compare the aggregate level of fiscal income estimated using SILC and SWITCH. We find that for 2015, this estimate is close to €75 billion. This compares with a figure of €87 billion based on tax returns to the Revenue authorities (and published by CSO).[9] The sources of this sizable apparent discrepancy deserve further investigation but are outside the scope of the current paper. Our focus is on top income shares: If the gap in aggregate income is evenly spread across the distribution, then estimates of top income shares may be similar – our analysis helps to establish how well SILC covers the top income echelons compared with others.

Revenue, 2019; Collins and Walsh, 2010; Kennedy et al., 2016) . This affects analyses of disposable income rather than the comparisons of fiscal income which are the core element of this study.

[9] We average the Revenue statistics for 2014 and 2015, to take account of the time period covered by SILC incomes.

.
The differences between disposable income and fiscal income are very large and, as will be seen, contribute substantially to the apparent differences in inequality (see Figure 1) as measured by WID, based on tax return data, and by CSO, ESRI and others based on analysis of SILC. It is, however, only one of many such differences in concepts and methods, as Table 2 summarises.

Table 2: Differences in Concepts and Methods

Item	*World Inequality Database*	*Household Income Distribution Analyses (e.g., SILC, Eurostat, OECD)*
Data source(s)	Tax returns, National Accounts	Household Survey (SILC)
Income concept	Fiscal income	Disposable income
Income sharing unit	Tax unit	Household
Adjustment for needs ofincome sharing unit	None	Equivalisation using anequivalence scale (1 for first adult, 0.66 for otheradults, 0.33 for children)
Unit of analysis	Tax unit	Individual
Main inequality measure(s)	Share of top 10%	Gini coefficient

Table 2 spells out the other important differences which can contribute to differing picture of inequality.

- WID analysis focuses for the most part on the tax unit as the income unit, and the unit of analysis. SILC-based analyses tend to focus instead on the household as the income sharing unit, and the individual as the unit of analysis.
- WID analyses mainly focus on aggregate income within the tax unit, with no adjustment for the numbers of adults and children relying on that income. SILC based analyses, by contrast, mainly focus on income distribution adjusted, using an equivalence scale, for the size and composition of the household.
- WID analyses tends to focus on the share of the top 10% (or top 1%) while SILC analyses tend to focus on broader measures of inequality such as the Gini coefficient or the full set of decile shares.

Table 3 outlines a path from the usual concepts used in analysis of SILC and other household survey data to analysis of household survey data using the same concepts as in the WID's top income analysis. We have implemented this using SWITCH, based on SILC 2015, and can now see the extent to which the changes in income concept and methods help to narrow the apparent gap between household survey-based estimates (SWITCH/SILC) and those based on tax return data (WID).[10]

Table 3: Top income share and Gini coefficient for Alternative Income Concepts: Leading to Comparable Top Income Shares for Tax Return Data and SILC 2015

Income concept	*Unit of analysis*	*Equivalised /unequivalised*	*Source*	*Share of top 10%*	*Gini*
Household income distribution analysis:				(%)	
Disposable income	Household	Equivalised	SILC	23.9	30.8
Intermediate steps: Disposable income	Household	Equivalised	***SWITCH***	22.6	29.4
Disposable income	Household	***Unequivalised***	SWITCH	22.9	32.0
Disposable income	***Tax unit***	Unequivalised	SWITCH	29.1	43.5
Fiscal income	Tax unit	Unequivalised	SWITCH	38.2	56.6
Top income analysis:					
Fiscal income	Tax unit	Unequivalised	WID	37.2	

Notes: On each row the item in bold represents the change in concept or data source from the previous line.

[10] While it is possible to replicate the WID concepts using household survey data, the converse is not true.

There are some conceptual differences between the SWITCH data and the standard SILC data, as SWITCH works with simulated welfare entitlements and simulated tax liabilities – assuming 100% take up of welfare entitlements and 100% compliance on tax. Furthermore, SWITCH works with annualised current income, which is relevant to modelling of current welfare and medical care entitlements, rather than the annual measure used in CSO's published SILC statistics on income distribution and risk of poverty. Nevertheless, the Gini coefficients and shares of the top decile are similar for SILC and SWITCH.

SILC 2015 estimates the top income share to be just under 24 per cent, when based on equivalized disposable income at household level. This compares with a share based on "top income methodology" (gross, unequivalised tax unit income from tax return data) of just over 37 per cent. The "like for like" comparison is a top decile share of just 38.2 per cent from the household survey as against 37.2 per cent from the tax return data. Alignment of the concepts used in the analysis brings the gap between the estimates from 13 percentage points down to one percentage point. The intervening rows show that the major contributory factors are

- moving from household to tax unit as the unit of analysis and
- moving from disposable income (post tax, post transfer) to fiscal income (as defined earlier – this is before all income-related taxes and includes only taxable social welfare payments).

While movement between equivalized and unequivalised income has little impact on the top income share, it does have a substantial effect on inequality as measured by the Gini coefficient, across the full distribution.

Table 3 illustrated this "crosswalk" using a particular path (from equivalized to unequivalised, then household to tax unit, then disposable to gross income). The impact of each transition may be dependent on this path, and on the inequality measure examined. We examine this issue in Table 4 (for the Gini coefficient) and Table 5 (for the top decile share).

Table 4: Impact of Changes in Income Concept on Gini Coefficient

Baseline definition					
Change in income concept	Changes in Gini in percentage points				Average change
Unequivalised to equivalised	Fiscal, Tax Unit	Disposable, Tax Unit	Fiscal, Household	Disposable, Household	
	-4.6	-7.0	-1.4	-2.6	-3.9
Fiscal to Disposable	Tax Unit, Unequivalised	Tax unit, Equivalised	Household, Unequivalised	Household, Equivalised	
	-13.2	-15.5	-14.3	-15.5	-14.6
Tax unit to household	Fiscal, Unequivalised	Fiscal, Equivalised	Disposable, Unequivalised	Disposable, Unequivalised	
	-10.2	-7.0	-11.4	-7.1	-8.9

For the Gini coefficient, we find that there is a clear ranking in terms of impact. A move from fiscal to disposable income sees the Gini fall by close to 14 or 15 points, about half of the total impact of these changes. The move from household to tax unit level has the next greatest impact, between 7 and 10 percentage points. The impact of moving from unequivalised to equivalised income is to reduce the Gini coefficient by an average of 4 percentage points. There is some path dependence here: the impact of equivalisation is greater at tax unit level and is sharply attenuated if the income concept has already moved to household level.

Table 5: Impact of Changes in Income Concept on Top Decile Share

Baseline income concept					
Change in income concept	*Changes in Gini in percentage points*				*Average change*
Unequivalised to equivalised	Fiscal, Tax Unit -3.7	Disposable, Tax Unit -3.6	Fiscal, Household -0.1	Disposable, Household -0.3	-1.9
Fiscal to disposable	Tax Unit, Unequivalised -9.1	Tax unit, Equivalised -9.0	Household, Unequivalised -7.7	Household, Equivalised -7.9	-8.4
Tax unit to household	Fiscal, Unequivalised -7.6	Fiscal, Equivalised -4.0	Disposable, Unequivalised -6.2	Disposable, Unequivalised -2.9	-5.2

For the top decile share, the same overall ranking holds. The greatest impact comes from a move from fiscal income to disposable income– a reduction in the top decile share of about 8 percentage points. The next largest impact is for a move between household and tax unit level – an average fall of some 5 percentage points. Again, the impact of equivalisation depends on whether this is at household level, when the impact is very small (less than half a percentage point) or at tax unit level, when the impact is between 3 and 4 percentage points. The impact of moves between unequivalised and equivalised incomes, and between tax unit and household levels of analysis does, however, depend on which of these is undertaken first. The impact of equivalisation is lessened if a move to household level has taken place first; and vice versa.

Table 6 documents the estimation of top income shares using WID concepts (fiscal income, unequivalised, at tax unit level) to SILC data from 2013, 2014 and 2015. We construct a household survey-based estimate of the top decile share, comparable to that used in top income studies, using fiscal income, unadjusted by an equivalence scale, with the tax unit serving as income sharing unit and unit of analysis.[11] We report the top decile share, and also break this into the component attributable to the top 1 per cent of tax units, and the next 9 per cent.

Table 4: Estimates of top income shares of fiscal income from World Inequality Database (WID) and from SILC using WID concepts

	Top 1%		Next 9%		Top 10%	
	WID*	SILC, authors' estimates	WID*	SILC, authors' estimates	WID*	SILC, authors' estimates
	%	%	%	%	%	%
2013	9.8	7.7	25.9	31.5	35.7	39.3
2014	10.0	8.3	26.1	30.7	36.1	39.0
2015	11.5	8.2	25.7	30.0	37.2	38.2

Notes: *World Inequality Database (wid.world), "fiscal income", after Nolan (2018).

These results show that survey-based estimates of the share of the top 1 per cent of tax units tend to be well below the corresponding estimates based on tax returns. Nor does the SILC based series capture the rise in the top income share for 2015. However, the share of the next 9 per cent of tax units based on survey data is above that estimated from tax returns, so that the share of the top 10 per cent of tax units is quite close for the two sources.

[11] Kennedy et al. (2019) report top decile shares based on tax records. While this is not strictly comparable with the WID concept, which takes into account incomes and persons not recorded by the tax authorities, the orders of magnitude for the top decile share are similar.

4. PROFILING TOP INCOME GROUPS

Given that SILC data is capturing much of the share of the top 10 per cent of the income distribution – with the partial exception of the top 1% - it is of interest to examine the profile of top income cases (as has been done, for example, by Lemieux and Ridell, 2015, for Canada, and Peichl et al., 2010, in Germany). A word of caution is needed, however. The picture of the top 10%, for example, may vary significantly depending on the income concept and unit of analysis used.[12] Table 5 illustrates this by examining those individuals (adults and children) who are in the top 10% of tax units based on fiscal income. We examine where these individuals, are located in the distribution of income based on other concepts. Thus, moving to a household unit of analysis, about three-quarters of those in the top 10% on a tax unit basis remain in the top decile, with the remainder in the 9th decile. When needs are taken into account by using an equivalence scale, the proportion remaining in the top decile falls a little further, and some cases are found in decile 8. A further shift to disposable income (post-tax, post-transfer) means that more than 1 in 3 of the individuals in the top decile on tax unit basis are now found in deciles 6 to 9. Most of these are in decile 9 or decile 8. These substantial transitions caution against a simple view that one can take a single income concept and identify "the rich".

Table 5: Proportion of All Persons (Adults and Children) in Top 10% of Tax Unit Fiscal Income by Deciles of Other Income Concepts

Income sharing unit	*Tax Unit*	***Household***	*Household*	*Household*
Income concept	*Fiscal income*	*Fiscal income*	*Fiscal income*	***Disposable income***
Equivalisation	*None*	*None*	***Equivalised***	*Equivalised*
Top 10%	100.0	76.1.	71.4	65.0
Decile 9		23.9	24.4	24.3
Decile 8			4.2	8.3
Decile 7				1.9
Decile 6				0.5
Deciles 1-5				
All	100.0	100.0	100.0	100.0

Notes: **Bold** indicates the change in income concept from the preceding measure, moving from left to right. A bracket symbol indicates that two cells have been grouped to respect rules on statistical confidentiality.

Table 6 analyses the adults living in top income tax units, ranked by pre-tax, post-transfer income per tax unit. Given that a very high proportion of these tax units (more than 8 out of 10) are married, it is unsurprising to find that numbers of men and women are similar. Analysis by age group shows that there is a strong concentration of adults aged 30-49 in the top income group: 62 per cent of adults in top income units are in this age group, compared with 39 per cent in the general population. There are relatively few top income cases in the young (under 30) or older (over 65) age groups.

Table 6: Profile of Adults in Top Income Tax Units, Ranked by Fiscal Income of Tax Unit

Characteristic		*Top 10%*	*All tax units*
Gender	Male	51.0	48.7
	Female	49.0	51.3
	Total	100.0	100.0
Age	<30	3.1	17.6
	30-49	62.6	39.3
	50-64	29.9	25.1
	65+	4.3	18.1
	Total	100.0	100.0
Marital Status	Married	87.5	55.8
	Single	12.5	44.2
	Total	100.0	100.0

Source: Authors' estimates, using SWITCH and SILC

[12] The sample size is too small to permit a similar analysis for the top 1%.

The World Inequality Database also analyses top incomes per adult, on the basis of an equal split of income between adults in a tax unit. We find that on this basis, married individuals would form close to three quarters of the top income population – rather less than the 88 per cent found when analysing on the basis of aggregate tax unit income.

How does the composition of income for top income units compare with that of all units? We examine this issue focusing on direct income (earnings from employment, self-employment, investment income and occupational pensions). Table 7 shows that wage income is the dominant source for all units, and even more dominant for the top decile. Because of the small number of cases in the top 1%, we analyse the top 2% in order to respect rules on statistical disclosure control. For the top 2% of incomes, self-employment becomes much more important than for those lower down the scale, even for the rest of the top 10%. The importance of self-employment incomes at the highest income levels is confirmed by analyses of the Revenue income distribution statistics (see Kennedy et al, 2019, Table 2).

Table 7: Composition of incomes for top income cases and for all tax units

	Top 2%	Next 8%	Top 10%	All units
Employee Income	75.2	83.6	82.0	71.9
Self-Employed Income	18.8	10.0	11.8	11.4
Investment Income	2.2	1.8	1.9	2.7
Pension Income	3.8	4.6	4.4	14.0
Total	100.0	100.0	100.0	100.0

Source: Authors' estimates, using SWITCH and SILC

5. CONCLUSIONS

There are now two main sources of data on income distribution. Household based surveys have tended to report mainly on inequality in household level disposable income, equivalised to take account of how household needs differ by family size and composition. Top income shares, on the other hand, tend to focus on the tax unit as the unit of analysis, because administrative records are obtained from such units, and are unable to be combined into households. Tax return data is also typically analysed in terms of fiscal income, and without adjustment to take account of the number of persons supported by that income. However, tax return data is likely to obtain better coverage of those at the very top of the income distribution. Thus, when differences emerge between the pictures of inequality arising from household and tax return data, it is unclear the extent to which they reflect differences in the data and coverage of these sources, and to what extent it reflects differences in the concepts and measures applied to the data.

Our analysis, following methods similar to Burkhauser et al, (2018a), finds that differences in concepts and measures play a very substantial role in accounting for the divergence in results. Estimates of the share of the top 10% of tax units in fiscal income from the two sources – SILC and top income analysis based on tax returns and national accounts – are quite close. Average incomes for the top 1% of the population appear to be substantially higher in tax return data than in SILC – a pattern that has often been observed internationally.

At present, we are faced with a number of trade-offs in the choice of data source for distributional analysis. For example, tax returns give more precise information on top incomes, but cannot inform us as to the household context or the incomes of those low-income individuals who are not included in tax returns. These trade-offs cannot be resolved by a simple either/or choice – both have contributions to make to our knowledge, with the balance between the two depending on the question at hand. We concur with Kennedy (2019) who states that "The future of best-practice tax policy analysis is likely to combine the unique advantages of tax, survey and national accounts data". In this spirit, we make the following suggestions, based on our findings to date.

One key difference between the "top incomes methodology" and the household survey based analysis relates to the income sharing unit. For the former, data is collected and analysed at tax unit level – this can be a single individual, a married couple or civil partners. For the latter, it is most often the case that analysis is undertaken at household level. There is, however, potential for moving to a common unit, under the assumption of equal sharing between couples. WID provides some analysis at this level, and household based surveys can do likewise. The substantive value of this is that it takes a step towards recognising variation in needs across tax units. A further step in this direction would be to adjust for the numbers of child dependants in the tax unit.

There is a strong case for examining potential adjustments to survey data to ensure better representation of income levels at the very top of the income distribution. Such adjustments combine information from tax returns and household surveys to give better representation of top incomes, while retaining the advantages of household data's perspectives on low incomes and household context. Burkhauser et al. (2018a, b) suggest an improvement on the pioneering adjustment of UK household data using tax returns from the Survey of Personal Incomes can now be implemented. Shine et al. (2019) confirm the commitment of the Office of National Statistics to such an approach. Bartels and Metzing (2019) apply adjustments to the incomes of the top 1% in SILC, based on estimates of the Pareto distribution of top incomes using tax return data. The analysis undertaken here represents a building block which can be used in investigating appropriate top income adjustments for the Irish case. This will allow distributional analysis to move beyond balancing the pros and cons of household survey and tax return data, to investigations which combine some of the best aspects of both approaches.

References

Aiyar, S. and C. Ebeke (2019) "Inequality of Opportunity, Inequality of Income and Economic Growth" IMF Working Paper, February.

Atkinson, A.B. (2002) "Top Incomes in the United Kingdom over the 20th Century", University of Oxford, Discussion Papers in Economic and Social History, No. 43.

Atkinson, A. B. (2015). *Inequality. What can be done?* Cambridge, MA: Harvard University Press

Alvaredo, F. & García-Peñalosa, C. (2018) Top incomes, wealth and inheritance: special issue in homage to Tony Atkinson, *Journal of Economic Inequality* Vol. 16, No. 131

Alvaredo, F., A.B. Atkinson, T. Piketty and E. Saez (2013) "The Top 1 Percent in International and Historical Perspective", *The Journal of Economic Perspectives*, Vol. 27, No. 3/

Alvaredo, F., L. Chancel, T. Piketty, E. Saez, G. Zucman (2018) *World Inequality Report*, World Inequality Lab, https://wir2018.wid.world/

Alvaredo, F., A.B. Atkinson, L. Chancel, T. Piketty, E. Saez, G. Zucman "Distributional National Accounts Guidelines: Methods and Concepts Used in WID.world" WID.worl Working Paper Series N° 2016/2

Bartels, C. and M. Metzing (2019) "An integrated approach for a top-corrected income distribution", *Journal of Economic Inequality*, Vol. 17, pp. 125-143.

Bee, A. and J. Rothbaum (2019) "The Administrative Income Statistics (AIS) Project: Research on the Use of Administrative Records to Improve Income and Resource Estimates", SEHSD Working Paper No. 2019-36.

Brewer, M. (2019) "Top Incomes in the UK: Analysis of the 2015-16 Survey of Personal Incomes", Institute for Social & Economic Research, Working Paper No. 2019-06.

Burkhauser, R., N. Herault, S. Jenkins, R. Wilkins (2018a) "Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data", *Oxford Economic Papers,* Vol. 70, No. 2, 301-326.

Burkhauser, R., N. Herault, S. Jenkins, R. Wilkins (2018b) "Survey Under-Coverage of Top Incomes and Estimation of Inequality: What is the Role of the UK's SPI Adjustment", *Fiscal Studies*, vol. 39, no. 2, pp. 213-240.

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*, Vol. 94, No. 2, pp. 371-388.

Collins, M. and M. Walsh (2010) "Ireland's Tax Expenditure System: International Comparisons and a Reform Agenda" *Studies in Public Policy No. 25*, Trinity College Dublin

CSO (2017) Estimates of Irish Pension Liabilities 2015, https://www.cso.ie/en/releasesandpublications/ep/p-eipl/eipl2015/eipl/

CSO (2019) Standard Report on Methods and Quality for the 2018 Survey on Income and Living Conditions (EU-SILC)

Doorley, K., T, Callan, M. Regan and J. Walsh, 2018. "The tax treatment of pension contributions in Ireland," Quarterly Economic Commentary: Special Articles, Economic and Social Research Institute (ESRI).

Finnie, R., and I. Irvine (2006) "Mobility and Gender at the Top Tail of the Earnings Distribution" *The Economic and Social Review, Economic and Social Studies*, vol. 37(2).

Jenkins, S. (2016) "Pareto Models, Top Incomes and Recent Trends in UK Income Inequality" *Economica*, doi: 10.1111/ecca12217.

Kennedy, S., Y. Jin, D. Haugh, P. Lenain (2016) "Taxes, Income and Economic Mobility in Ireland: New Evidence from Tax Records Data", *The Economic and Social Review,* Vol 47 No 1, Spring/

Kennedy, S., D. Haugh and B. Stanley (2019), "Yours inclusively? Income mobility in Ireland, 10 years of tax record microdata", OECD Economics Department Working Papers, No. 1578, OECD Publishing, Paris. https://doi.org/10.1787/f3df1668-en.

Kennedy, S. (2019), "The potential of tax microdata for tax policy", OECD Taxation Working Papers, No. 45, OECD Publishing, Paris, https://doi.org/10.1787/d2283b8e-en.

Lemieux, T. and W. Craig Ridell (2015) "Top Incomes in Canada: Evidence from the Census", NBER Working Paper 21347, July.

Lustig, N (2018), 'Measuring the distribution of household income, consumption and wealth', in J.E. Stiglitz, J.-P. Fitoussi and M. Durand M (eds), *For Good Measure: Advancing Research on Well-being Metrics Beyond GDP*, Paris: Organisation for Economic Co-operation and Development Publishing.

Murphy, B., P. Roberts and M. Wolfson (2007). *A Profile of High Income Canadians*. Statistics Canada, Income Statistics Division, Income Research Paper Series

Nolan, B. (2007) "Long-Term Trends in Top Income Shares in Ireland" in A.B. Atkinson and T. Piketty (eds.) *Top Incomes over the 20th Century, A Contrast Between Continental European and English-Speaking Countries*, Oxford: OUP.

Nolan, B. (2018) "Updating Top Incomes for Ireland to 2015", WID.world Technical Note 2018/02.

OECD (2008) *Growing Unequal? Income Distribution and Poverty in OECD Countries*, Paris: OECD Publishing.

OECD (2011), Divided We Stand: Why Inequality Keeps Rising. Paris: OECD Publishing. https://doi.org/10.1787/9789264119536-en.Google Scholar

OECD. (2015). *In it together: Why less inequality benefits all*. Paris: OECD Publishing. https://doi.org/10.1787/9789264235120-en

Peichl, A., Schaefer, T. and Scheicher, C. (2010), "Measuring Richness and Poverty: A Micro Data Application to Europe and Germany" *Review of Income and Wealth*, 56: 597-619. doi:10.1111/j.1475-4991.2010.00404.x.

Piketty, T. *Capital in the Twenty-First Century*. Cambridge Massachusetts: The Belknap Press of Harvard University Press, 2014.

Revenue (2019a) *Analysis of High Income Individuals' Restriction 2017* available at: https://www.revenue.ie/en/corporate/documents/research/ror-2017-report.pdf

Revenue (2019b) *Costs of Tax Expenditures* https://revenue.ie/en/corporate/documents/statistics/tax-expenditures/costs-tax- expenditures.pdf

Roantree, Barra, Understanding Income Inequality in Ireland. Journal of the Statistical and Social Inquiry Society of Ireland, 2020.

Ruiz N and Woloszko N (2016), 'What do household surveys suggest about the top 1% incomes and inequality in OECD countries?', Economics Department Working Paper, No. 1265. Paris: Organisation for Economic Cooperation and Development

Shine, M., C. Clark, R. Tonkin and D. Webber (2019) "Using tax data to better capture top earners in household income inequality statistics" ONS article, available at https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/usingtaxdatatobettercapturetopearnersinhouseholdincomeinequalitystatistics/2019-02-26

Yonzan, N., B. Milanovic, S. Morelli and J. Gornick (2018) "Comparing top incomes between survey and tax data: US case study" Luxembourg Income Study newsletter Issue 2018-6, online at http://www.lisdatacenter.org/newsletter/nl-2018-6-h-4/

FIRST VOTE OF THANKS PROPOSED BY AEDÍN DORIS

I am delighted to propose a vote of thanks to Tim Callan, Karina Doorley and Alyvia McTague for their important paper on the evidence on top incomes in Ireland. As the authors note, there has been increasing interest in the topic of the share of earnings, income and wealth going to the very top of their respective distributions in recent years. However, until recently in many countries, the only data that we have had available to study these distributions are household survey data. As a frequent user of such surveys, I have always found it hard to shake off the suspicion that the typical one-percenter might be reluctant to answer them. Under-representation of the top one percent is a clear concern.

In addition, in a typical Irish household survey, the number of households surveyed is 4,000-6,000, with 10,000-15,000 individuals. If you then wish to focus on the top one percent of the distribution, you can be looking at one or two hundred individuals and can quickly run into problems of small cell sizes (and indeed confidentiality) if examining the characteristics of these individuals. Indeed, the authors faced this problem themselves in the paper; when breaking the top earners down by income sources, they could report the breakdown only for the top two percent. For both of these reasons, using tax data is attractive as they contain earnings data on the entire population of workers so sample size is not an issue. However, they exclude non-workers and typically include very few demographic variables or details on household membership.

Analyses of tax data typically find higher concentrations of income at the very top of the distribution than analyses of household survey data, and it can be hard to figure out whether these differences stem from differences in definitions or differences in coverage or both. With the care that is characteristic of these authors, they tackle this issue directly by analysing household data from EU-SILC in detail (using their SWITCH model) and comparing the results to those in the World Inequality Database, which is based on tax data. They find that once account is taken of the differences in income definitions and in the unit of measurement, survey data is pretty good at matching to tax data, at least for the top 10% of incomes. And they show that the biggest contributor to the differences between the EU-SILC and WID-based top decile shares are the differences in income definition, followed by the move from tax unit to household. These are very useful results for those of us trying to understand competing inequality figures.

For those in the top one percent, however, the findings are less reassuring – the top income proportions are further off, and so are the trends – although admittedly, the trends are only presented for three consecutive years, so comparisons over a longer time span when more data become available will be interesting. One response to this would be to say that the top decile covers a large enough portion of the population that it matters for overall inequality. Therefore, it is far more important to match the survey data to the tax data on the top decile than on the top one percent. But there are good reasons for the additional interest in the top one percent – or even, as in some recent studies, in the top 0.1 percent or the top 0.01 percent.

For one thing, in many countries, the WID database indicates that the share of taxable income (which they call fiscal income) going to the top one percent has been increasing over several decades, often doubling since the 1980s, including in Ireland. In Ireland, it rose from 5.2% in 1986 to 13.8% in 2006. The Great Recession then took its toll and it dropped back to 9.6% in 2009, but has since risen to 11.5% in 2015, the most recent year for which data are available. Although these numbers are very high, they are not particularly high by international standards; the 2014 figure for the US is 20.4%, which itself is an almost doubling since 1980.

As is well-known, this rise has coincided with a rise in overall earnings inequality in many countries. And of course earnings inequality matters because if it is rising, tax and transfer policy has to work harder and harder just to keep income inequality standing still. Could what's happening in the top percentile of the distribution be the result of forces that are also causing earnings at the bottom of the distribution to stagnate?

At least one explanation of the growth of the share of the top one percent does suggest so. Matching models of the labour market, which are now the standard way of analysing the labour market, are based on the idea that firms generate a surplus, whose division between the firm and workers is decided by their relative bargaining power. In these models, if the share going to senior management has been rising, then it will reduce the share going to workers. Many countries have seen a reduction in union power in recent decades. In addition, reductions in marginal tax rates for the very highly paid in the 1980s may have increased the incentive to negotiate increased gross pay. In this explanation, in which bargaining power is central, what is happening at the very top could be having a significant effect on the rest of the distribution.

If this is what's going on, the solution to increasing earnings inequality is increasing marginal tax rates for the very highly paid and strengthening workers' bargaining power. This argument is not valid if the reason for the growth of the share going to the very top of the earnings distribution is that market forces have strengthened, as illustrated by an increase in performance-related pay (including the awarding of share options), which has led to higher productivity. In this case, the pie is growing, and the fact that a few people at the very top of the distribution are benefitting disproportionately is of no consequence for overall inequality.

The problem is that there are reasons to suspect that the growth in executive pay – and pay in financial services – is due to rent capture rather than increases in productivity or the operation of efficient markets. First, there is little evidence that share options increase productivity. In addition, earnings in the financial services sector have grown most in countries in which deregulation has been strongest, with little evidence that this has increased the efficiency

of financial markets, and some strong evidence to the contrary. In this case, the solution to increasing earnings inequality is the re-regulation of some markets and reform of corporate governance to align executive pay better with relative firm performance. There are also broader problems that can arise if income of the top one percent is growing so much faster than the rest of the distribution's. For one thing, it may reduce the faith in the idea that our society is built on meritocratic lines. This could undermine the social contract and indeed democracy.

In summary, I am convinced of the importance of the one percent. And if our current household survey data is not quite capable of capturing them adequately, then the current efforts underway at the Central Statistics Office to merge administrative earnings data into other household survey datasets – particularly larger surveys such as the Labour Force Survey and the Household Budget Survey – should be encouraged and prioritized. Alternatively, the collection of some key survey variables (such as level of education) along with tax data could be considered. Finally, a great benefit of tax data that I haven't mentioned previously is that it allows the construction of panel data sets that are hugely expensive otherwise. We cannot say much about lifetime earnings without panel data. We cannot comment on the distinction between permanent and transitory earnings and income inequality without panel data. These are crucial to devising appropriate policy responses to inequality.

SECOND VOTE OF THANKS PROPOSED BY ROBERT SWEENEY

This is a very interesting paper. The issue of whether survey data, the official data used by most researchers and policymakers, is correctly capturing distributional trends is of obvious political and economic import. Given what appears to be the unusual performance of Ireland of falling, as opposed to rising, inequality over the last 25 or so years, the stakes are perhaps even higher here. Working for an organisation whose core mission is to research and advocate on issues surrounding inequality I, for one, have been unsure whether to accept the conventional narrative or not. I now have greater confidence that this is the case - that income inequality has fallen - though I would like to see the analysis extended so that there can be no room for doubt.

The paper attempts to reconcile evidence from survey data, SILC, with data based on tax records. As the authors note, each type of data has its own strengths and weaknesses. The main problem with survey data is that they fail to accurately capture incomes at the very top due to under-reporting and non-response by the rich. This is less of a problem for tax return data given the obligatory nature of paying taxes or filing returns. Problems with tax data include a lack of demographic information, underrepresentation of lower income groups, and exclusion of non-taxable income. Importantly, the unit of analysis is the tax unit, which may be an individual or may be a married couple.

Some further points the authors did not discuss should be borne in mind. For instance, the definition of taxable income may not necessarily be the same across years. This may or may not impinge on a reconciliation of different sources, but does bear on tax data trends (see Galbraith, 2018). Moreover, neither survey data nor the tax data capture income that escapes taxation. Again this would not affect a reconciliation of survey and tax data sources, but does point to the limitations of those sources. For instance, at 0.44% of GDP per annum individual tax evasion in Ireland is estimated to be are around EU norms, though when scaled by GNI* it would be above average (Vellutini et al., 2019: 87; 189).

As to the meat of the paper, most of the findings in the reconciliation are as expected. The effect of moving from unequivalised to equivalised income is comparatively modest. The move from fiscal to disposable income results in a sharp decline in inequality. The move from the tax unit to the household also exerts a powerful influence, resulting in a somewhat less sharp though still very significant decline in inequality. This, to me, was an interesting result, about which I had few priors.

The core result relates to the WID or tax data, and the SILC data adjusted to be comparable to the tax data. The reconciliation is performed using three years of data, 2013-15. On a like-for-like basis, the share of income going to the top ten percent according to SILC is similar to that found using tax data. Surprisingly, tax data understate the income share of the top decile compared to SILC. On the other hand, SILC understates the share of the top one percent compared to the tax data, and the gap between the SILC and the tax data grows for each year. In aggregate the two sources are quite similar.

The authors also profile the top ten percent. They look at the demographic make-up of the top decile using fiscal income and the tax unit, and see how it changes when the survey concept measure of equivalised household disposable income is gradually applied. Interestingly, only about two thirds of those in the top decile of equivalised disposable income are in the top ten percent of fiscal income tax units. The top tax payers are also mostly middle-aged and married.

A natural question arises: can these results be extended back in time such that we are sure inequality has indeed fallen? The results are consistent with a number of possibilities. One is that the survey and tax data create fundamentally different results, even after adjustments are made to reconcile them. In this case, the results obtained

in this paper are merely coincidental, an anomaly of the three years in question. It would be quite a coincidence if this were the case - that the three years chosen happened to be the ones in which the two sources arrived at similar results.

The other possibility is that they are consistent with one another. This is the most likely possibility given, as above, the coincidence needed for this not to be the case. Granted, the top one percent do grow further apart in each year. But if this were systematic then surely the levels of concentration - 11.5% as per WID vs 8.2% as per SILC - would be even further apart if the two series had been diverging before.

With only three years of data one cannot be definitive. There are, of course, middle grounds between these two poles - for instance, broad consistency in some periods, but not in others. The lower concentration of the top decile measured with SILC data is also surprising. These are some of the kinks that could be ironed out if the series were extended back further. Finally, I would also point out that the SWITCH model plays an important role in the paper, but remains unspecified. Nevertheless, the results and the paper do provide greater assurance that survey trends, despite their flaws, can be trusted.

References

Galbraith, J. K. (2019). Sparse, inconsistent and unreliable: Tax records and the World Inequality Report 2018. *Development and Change*, 50(2), 329-346.

Vellutini, C., Casamatta, G., Bousquet, L. and Poniatowski, G. (2019). Estimating international tax evasion by individuals, *European Commission*, Working paper 76.

DISCUSSION

Brendan P. Ryan: Thank you for the presentation. I would just like to ask Tim a question about the treatment of mortgage payments in assessing the distribution of disposable income. Clearly this has major implications as one moves through the age categories - as older people tend to have paid off their mortgages. In addition is there any treatment of imputed rent? Again, people who own their own homes with no mortgage payments may have a higher disposable income than people renting similar homes.

Brendan O'Dowd: My question is whether there was a divergence in family sizes between top and middle income households since 1995, and whether this could counteract or mask growing inequality as described by equivalised measures such as equivalised GINI. I wonder if low and middle income households are now "priced out" of having larger families due to rising costs of childcare and large houses (4+ bedrooms, say), and whether this is more acute than in 1995. My understanding is that household size or family size is the denominator in equivalistion (reduced factor of 0.5 or 0.3 for children notwithstanding). So a reduction in both family size and income for lower income households since 1995 could mean that equivalised income for this cohort is static. Therefore equivalised measures may show that inequality has not risen, but perhaps the real effect of growing inequality could be a restriction on family size specifically for low or middle income households.

Barra Roantree: Thanks to Tim, Karina and Aly for a really important paper. A question relating to capital gains: am I right in saying these are not fully captured in either the WID style measures or SILC? If so, given evidence from elsewhere (notably the UK e.g. Advani and Summers, 2020) of the growing importance of such income for the top 1%, what can we do to improve on the inequality statistics we have in that area?

Patrick Honohan: How is it that so many in the top decile by tax unit end up not in the top decile for household-level equivalised (post-tax)?

Jean Acheson: On method - These are interesting results for making the survey data more like tax data – with your specific WID source, though, can you make the tax data more like survey data i.e. create disposable income by subtracting direct taxes and employee PRSI from the gross/fiscal income in the tax data, and rerun the analysis? (Yes to Tim's query – the tax data held by the tax authority or CSO can do this – for example, net pay was a feature of TWSS policy design). On interpretation - The disposable-to-fiscal-income effect is stronger than the tax-unit-to-household income effect when looking at the Gini – can we conclude that the progressivity of the tax/welfare system is 'more important' for reducing inequality than household formation choices? On Aedin's discussion on explaining top 1% trends – an additional feature of tax return data (aside from panel and other aspects mentioned) is observation of employer type e.g. SME/corporate/foreign multinational. The latter category in particular conceivably has an effect in the Irish context?

Tim Callan: We are grateful for these comments and suggestions. Two brief points in response. First, on Patrick Honohan's query as to how so many in the top decile by tax unit end up not in the top decile for household-level equivalised income. A full answer to this would require a "step by step" analysis similar to that in Table 4. Absent this, our expectation is that equivalisation – taking account of the needs of larger households – plays a large role, as many of those in the top decile by tax unit have children. Second, a number of comments point the potential for a greater use of administrative data, either on its own, or in combination with household survey data, to provide a more accurate and comprehensive picture of income distribution trends. We agree that this should be a priority in further development of the national statistical system.