Wilkins, Burkhauser, Herault, and Jenkins (WBHJ), "What has Been Happening to UK Income Inequality Since the Mid-1990s? Answers from Reconciled and Combined Household Survey and Tax Return Data"

Jenkins, "Pareto models, top incomes, and recent trends in UK income inequality"

Discussion by

Michael Wolfson,
Professor and Canada Research Chair, uOttawa

IARIW Conference, August 2016, Dresden

Background

- major *increase* in interest in income inequality trends (from MCW's ancient perspective ☺)
  - growing use of income tax data at top end (top 1%)
  - increasing availability of unit record / microdata files
  - much work on international comparability
- longstanding and well-known issues of under-reporting of incomes, especially for the rich, in household surveys
  - n.b. possibly also under-reporting at low end of income distribution – not considered

The Challenge (Jenkins' opening para)

- "There is a bifurcation in the literature on income inequality levels and trends. On the one hand, most official statistics and academic analysis utilise data from household surveys and report estimates of the inequality of family or household disposable income summarised using Gini coefficients and other inequality indices calculated using all incomes from poor to richest.
- "On the other hand, there is the 'top incomes' literature that uses administrative record data on personal income tax returns, reporting estimates of top income shares – the share of total income received by the richest 1% or richest 10%, and so on."
- Challenge: bridging this "bifurcation"

Comparative Foci

- both papers study UK income inequality trends from mid-1990s to 2010/11 – using same data
- raison d'etre for both: serious under-reporting of top incomes in household surveys
- WBHJ – focus on impacts of
  - income definition
  - income recipient unit
  - inequality measure chosen
  - some comparisons to the US
- Jenkins – focus on methods of combining household survey and income tax data
  - individuals / gross income mainly

Challenges Restated & Distinguished

- there are major empirical / statistical issues with the quality of the source data for characterizing the distribution of income
  - household survey data and income tax admin data each used on their own yield very inconsistent results
- there are major choices in "concepts" underlying the distribution of income and assessments of inequality trends for which there is no "right" answer
  - income recipient unit
  - definition of income
  - equivalence scales
  - choice of inequality indicators / measures
  - (also accounting period, "ordering principle" (Love and Wolfson, 1976))

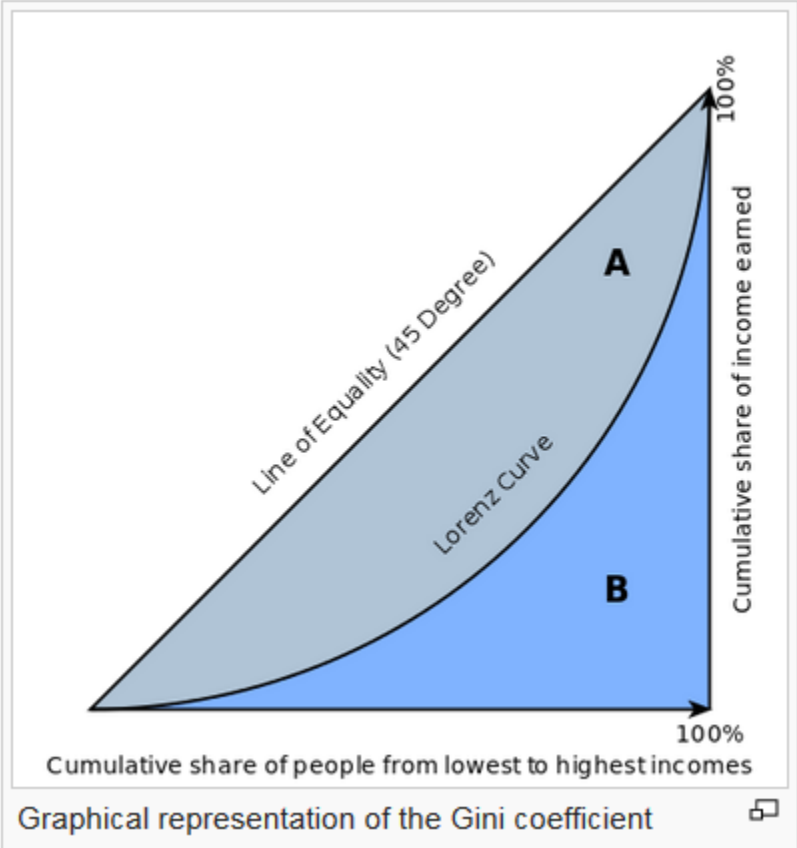
And Another Challenge

- "inequality" is not a scalar (duh!)
- summary inequality measures like the Gini and 90/10 ratio each capture only one feature of an infinity of points (or at least a large sample representation of millions of income units)
- and – using one representation of such an infinity of points – Lorenz curves can and do cross
- so in general we cannot make unequivocal statements that "inequality increased (or decreased)" based on one or even on only a few (poorly chosen) summary inequality measures
  - => ideally we need to examine the entire distribution

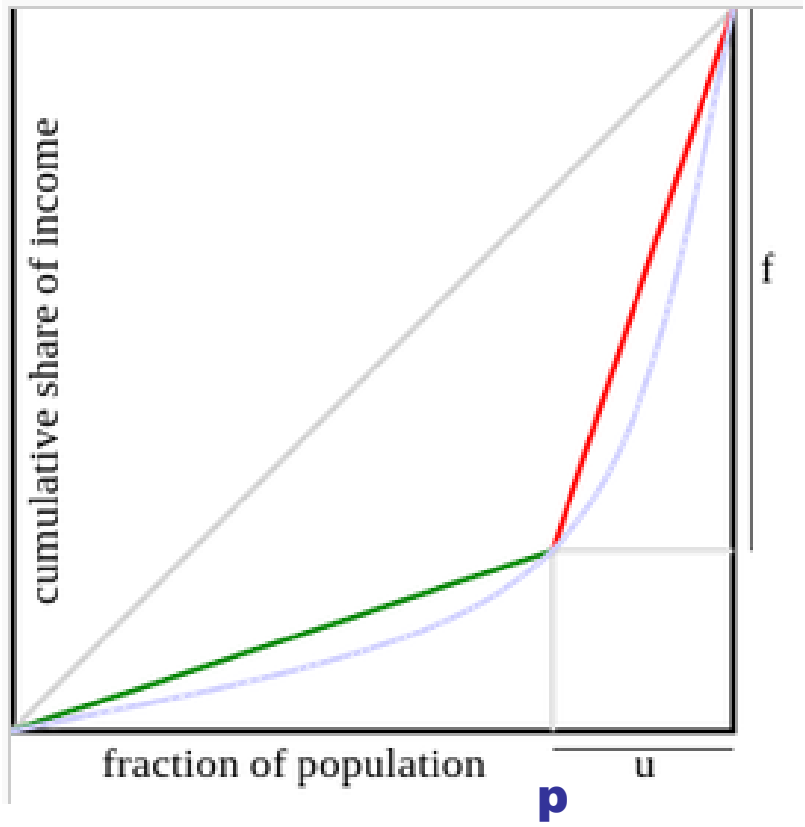
Addressing the Empirical Challenge

- both papers build on the same idea:
  - household survey data are of better quality than income tax admin data for the bottom p% of the population, where e.g. p = 90 or 95 or 99%
  - income tax data are better quality for the top (1-p)%
- great news: public use unit record income tax data are available in the UK!
- methodological question: how to combine these two data sources in order to benefit most from their comparative areas of strength

Graphical Representation of a Simple Combination Process



standard Lorenz curve



partitioned into two (easily additively separable) Lorenz curves at p

The Data in a Bit More Detail (I)

- income tax – three stratified samples combined
  - employees and occupational pension plan recipients
  - those with self-employment, rental, and “complexity”
  - those with too much tax deducted at source
- but these samples do not cover the entire population, so both population and income control totals are used to adjust for under-coverage
- n = 57,000 in 1995/6; 677,500 in 2010/11 (!)
- after-tax income available, but focus is on before tax total income for “comparability” with broader published literature

The Data in a Bit More Detail (II)

- everything else: Family Resources Survey
- n.b. by Department for Work and Pensions, not ONS
- (poor quality?) adjustment is made for under-coverage of top incomes, but WBHJ ignore it
- survey has insufficient detail directly to match income tax concepts – total income, individuals rather than families – but approximations were made
  - a kind of “reverse engineering” of suppressions made to construct non-identifiable public use file
  - and under-coverage of top incomes adjustment was undone

Instead: WBHJ Top Income Adjustment

- sort both surveyed individuals and income tax returns (each appropriately weighted) in ascending order of total income
- divide each sample into 1,000 equal sized population groups (“mtiles”, rather than deciles or percentiles)
- assign each mtile the average income of its members
- take the first 900 mtiles from the survey (i.e. first 90%), then concatenate last 100 mtiles (i.e. top 10%) from the income tax data
- voila – 1,000 mtiles with improved top end = totally consistent with published income tax data, by construction

Figure 2. Top 1% income shares: estimates from survey and tax return data

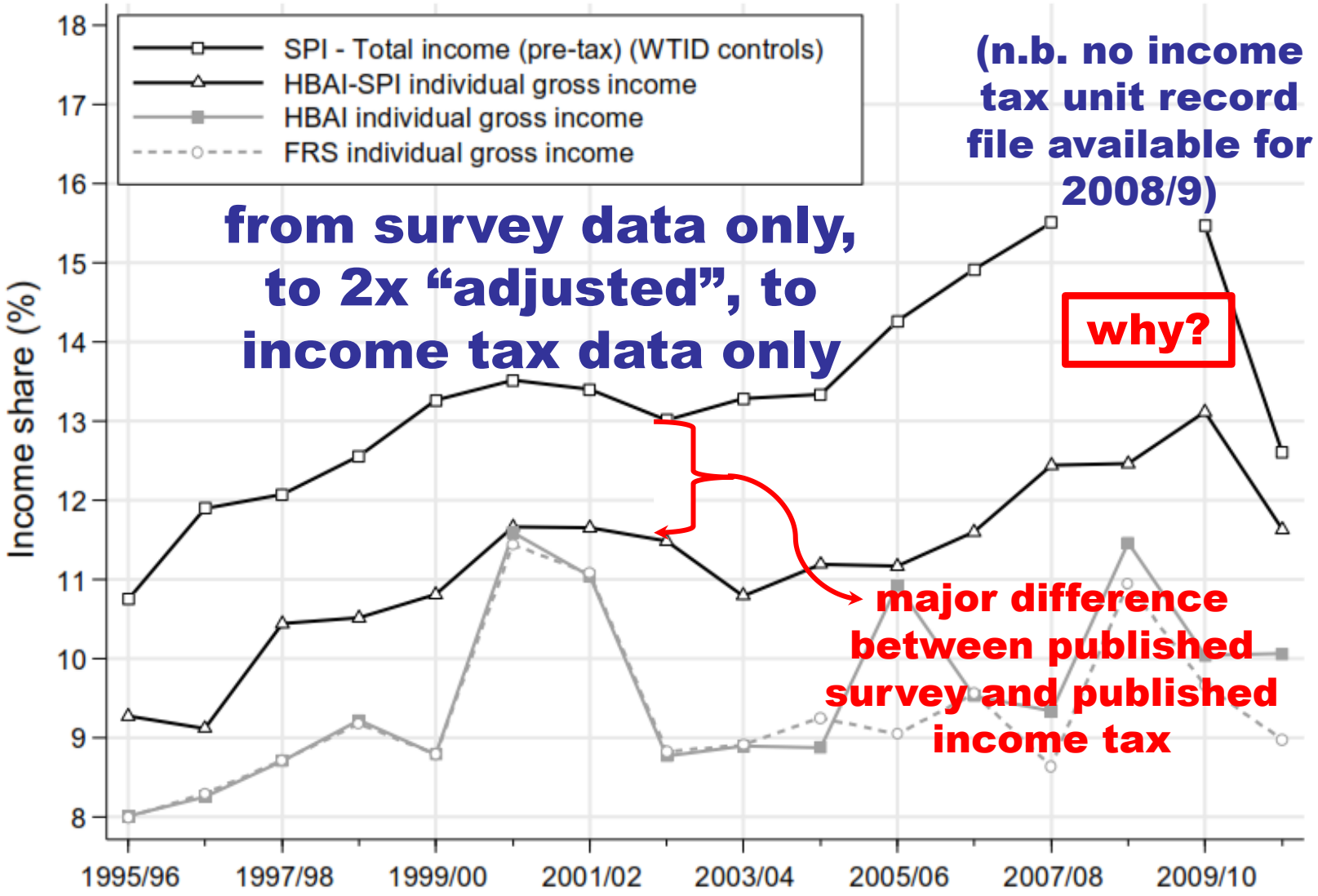
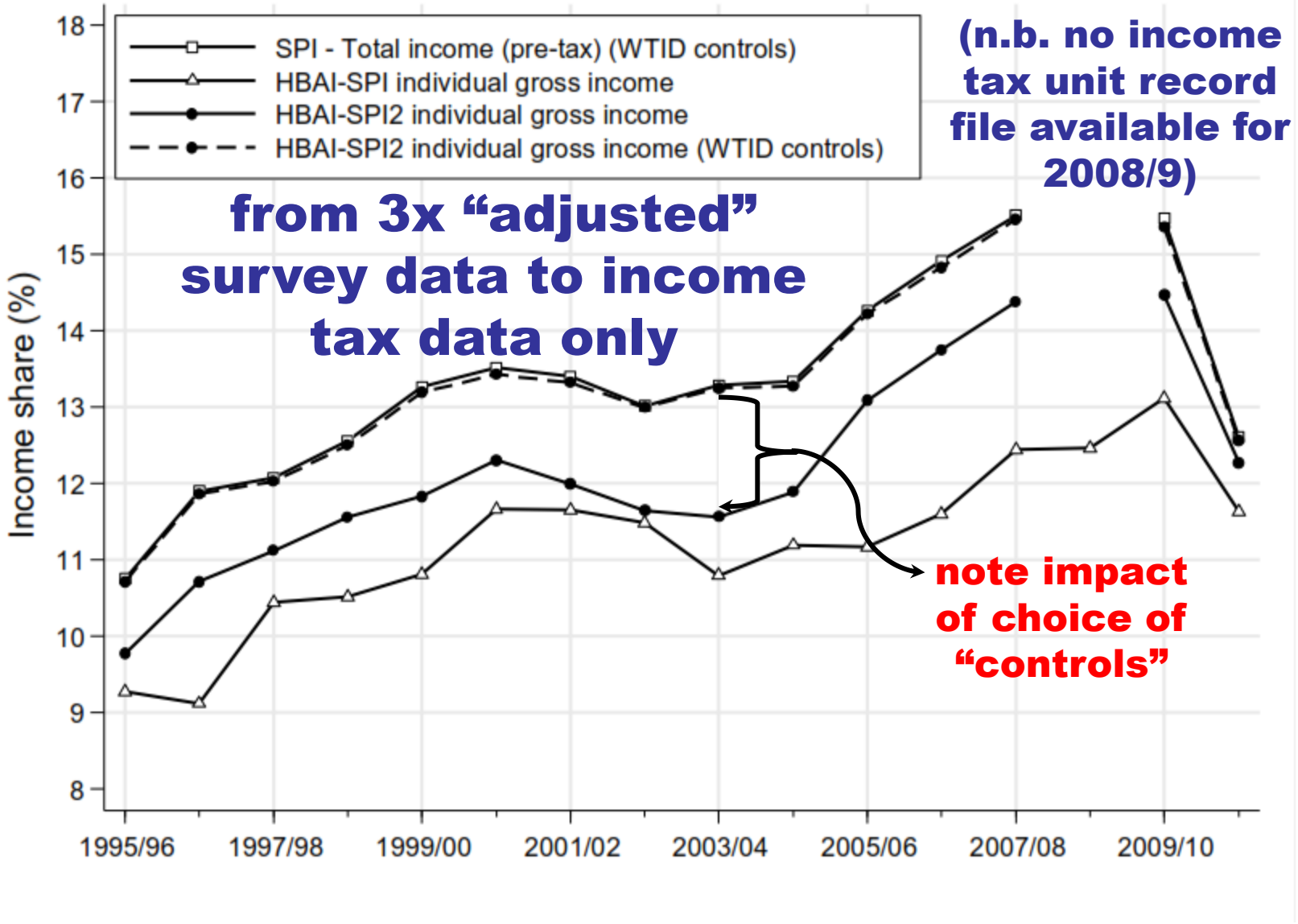


Figure 3. Top 1% income shares: estimates from SPI/WTID data and combined survey data



WBHJ Main Conclusions

- “the more that the survey data are adjusted to better capture top incomes, the greater is the effect on inequality that is observed.”
- “have shown with our … adjustment that it is possible to take even greater account of top incomes … than the official income distribution statistics in the UK”
- “the magnitude of the estimated increase in inequality, especially between 2004/05 and 2007/08, depends on whether the index is middle- or top-sensitive.”
- “The flexibility of unit record survey data also allows us to employ different definitions of income and the income-sharing unit.”

A few points for note / discussion

- KUDOS! – authors include underlying data and analysis details + standard errors
- but why aggregate into mtiles at all?
  - recall Gastwirth and grouping bounds => unnecessary grouping errors reduce inequality (though probably small, depends on measure)
  - why not use microdata directly?

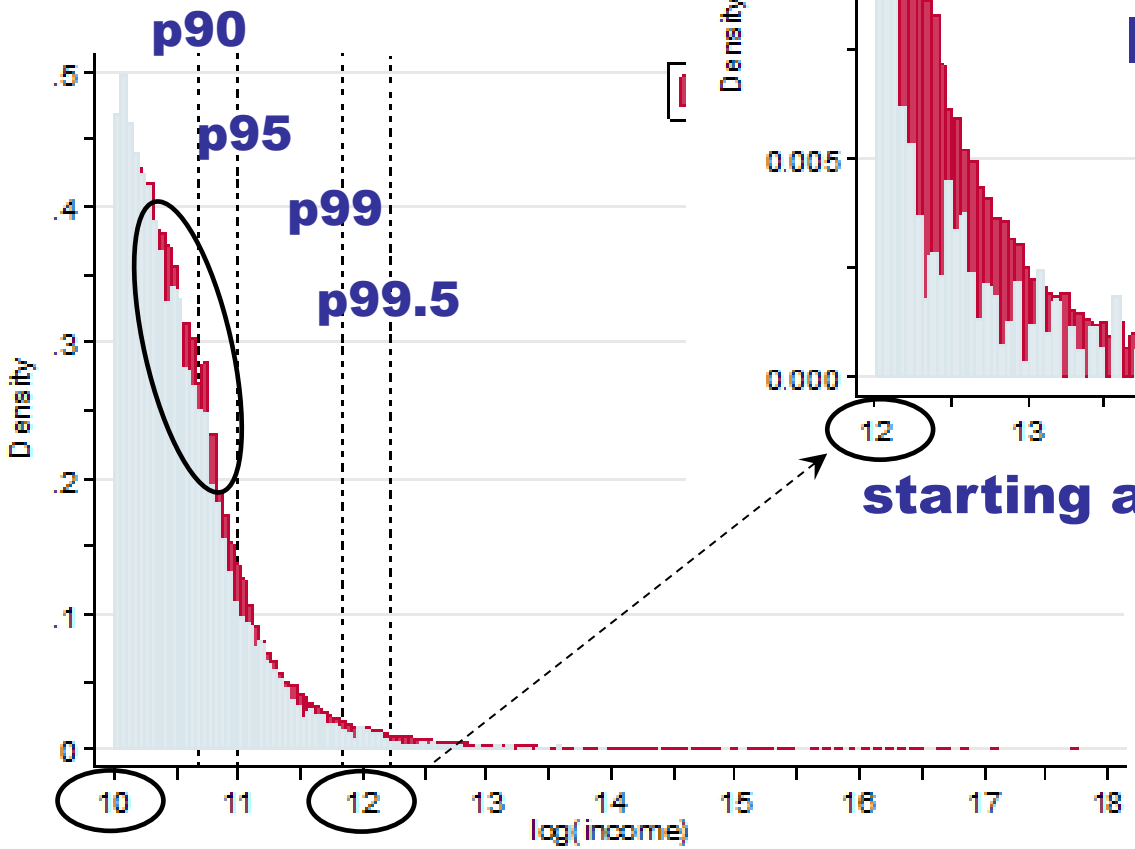
Now onto Jenkins paper

- also addressing same question as WBHJ – how best to improve estimates of UK income inequality, and trends therein, by using both survey and income tax data
- but focus is on other “semi-parametric” methods
- specifically, fitting some sort of Pareto distribution to the upper tail of the income distribution based on income tax unit record data

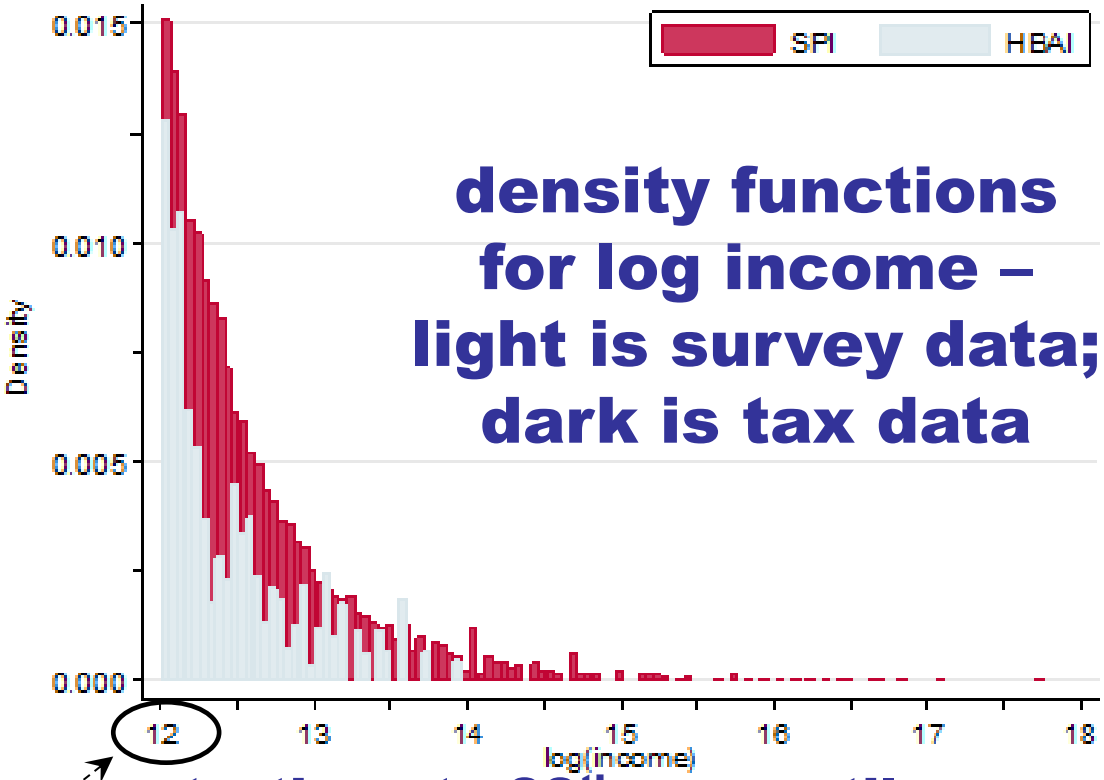
"Main Contributions of the Paper"

- new assessment of UK income inequality trends
- comparison of WBHJ approach and semi-parametric / Pareto tail approach
- new evidence on the extent of under-coverage of high incomes
- new analysis for the choice of functional form for fitting an equation to the upper tail
  - comparison of two versions of Pareto distribution
  - joint analysis with choice of "joining" threshold – by looking at p90, p95, and p99
    - n.b. assumes nil under-coverage below threshold?

Jenkins Figure 2 (2007)



starting well below 90th percentile



density functions for log income – light is survey data; dark is tax data

starting at ~99th percentile

n.b. no obvious threshold for starting to use income tax data

Three methods to adjust for under-reporting of upper incomes in UK income distribution survey

A. fit standard Pareto function to upper tail of income tax distribution data
   - widely used empirically, hence useful for comparisons

B. as in WBHJ just described – "non-parametric" (though grouped into mtiles – why?)

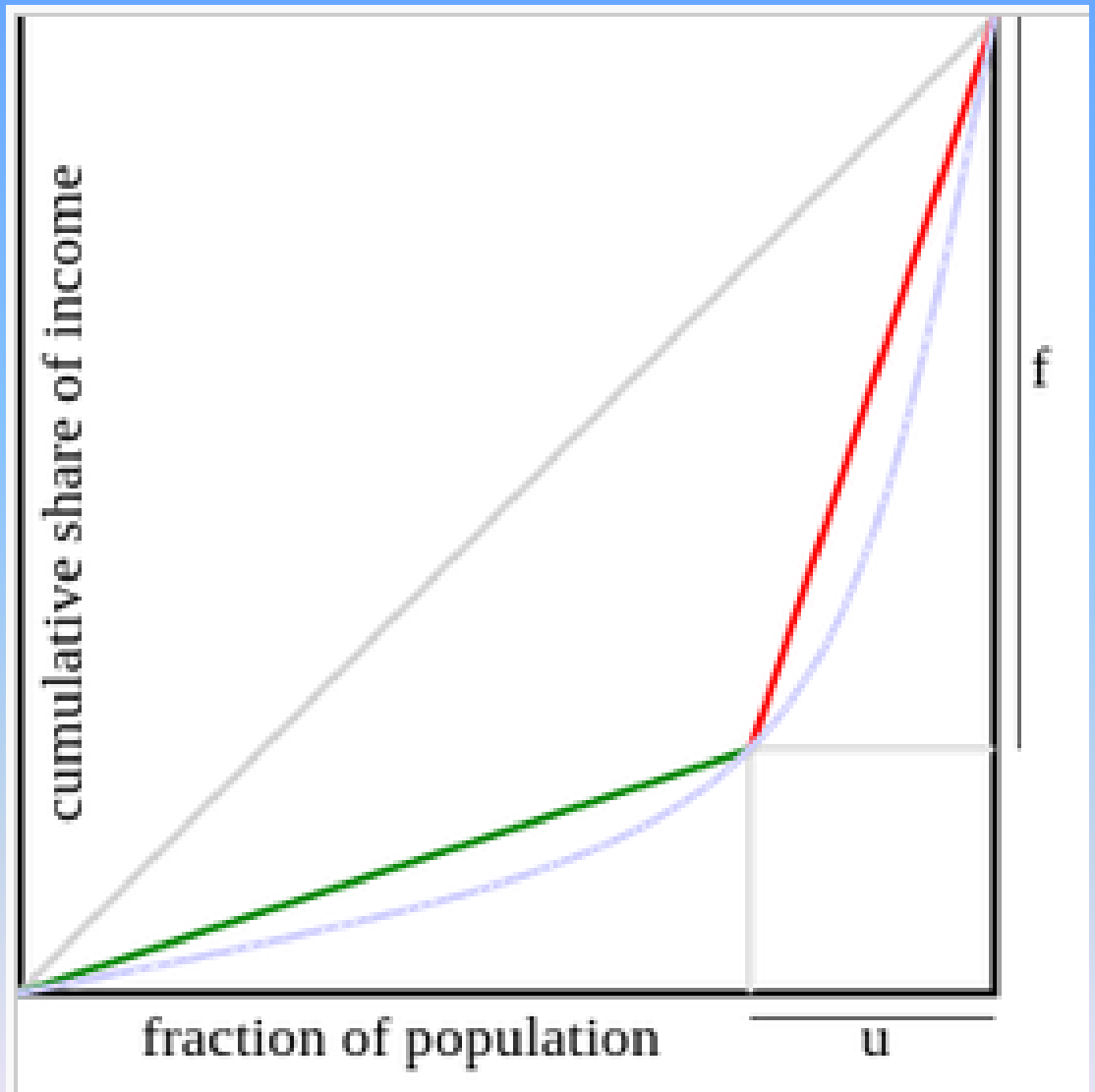
C. try fitting fancier Pareto function to upper tail of income tax distribution data

n.b. two kinds of "sensitivity problems" mentioned

- of inequality to outlier observations in upper tail
- bottom- versus top-sensitivity of measure itself

Interesting Claim – Meaning?

"Approach C, used in this paper, *combines estimates from the two types of data source* rather than *combining data per se as Approach B does. It is thus identical to Approach A."*



n.b. even the Gini is additively decomposable in this case

Choosing a Functional Form for the Upper Tail of the Income Distribution

- "there are implementation issues concerning the choice of model and the top income range over which they are fitted. There is also a prior question of whether top incomes are described better by a model other than a Pareto one. …
- "Addressing all these issues is complicated by a chicken and egg problem: most methods for choosing the appropriate model are conditional on a given threshold; and most methods for choosing the threshold have been applied to a single model.
- "but there can be an information overload, potentially worsened by having 15 years of data covering a period when the income distribution changed." (p12)

Statistical Estimation Aspects

- first check whether upper tail is Paretian
  - used interesting set of diagnostics: mean excess and Zenga plots,
  - notwithstanding relatively large sample size, results are equivocal
- not OLS; rather used (variant of) MLE for simple versus extended (type I vs II) Pareto model
- also estimated for upper tail starting at four different thresholds: p90, p95, p99, and p99.5
- 2 Pareto's x 4 thresholds x each of many years => lots of curves to compare
  - used Kolmogorov-Smirnov test = sup of absolute differences in CDFs

Pareto Functional Forms

A random variable X is said to follow a Pareto distribution if its density function $f(x)$ is such that

$$f(x) = \frac{\alpha x_0^{\alpha}}{x^{\alpha+1}}, \qquad 0 < x_0 \leq x, \tag{1}$$

where α is the so-called shape parameter, which measures the heaviness of the right tail, and x_0 is a scale parameter. The corresponding cumulative distribution function (cdf) is thus

$$F(x) = 1 - \left(\frac{x}{x_0}\right)^{-\alpha}, \qquad 0 < x_0 \leq x. \tag{2}$$

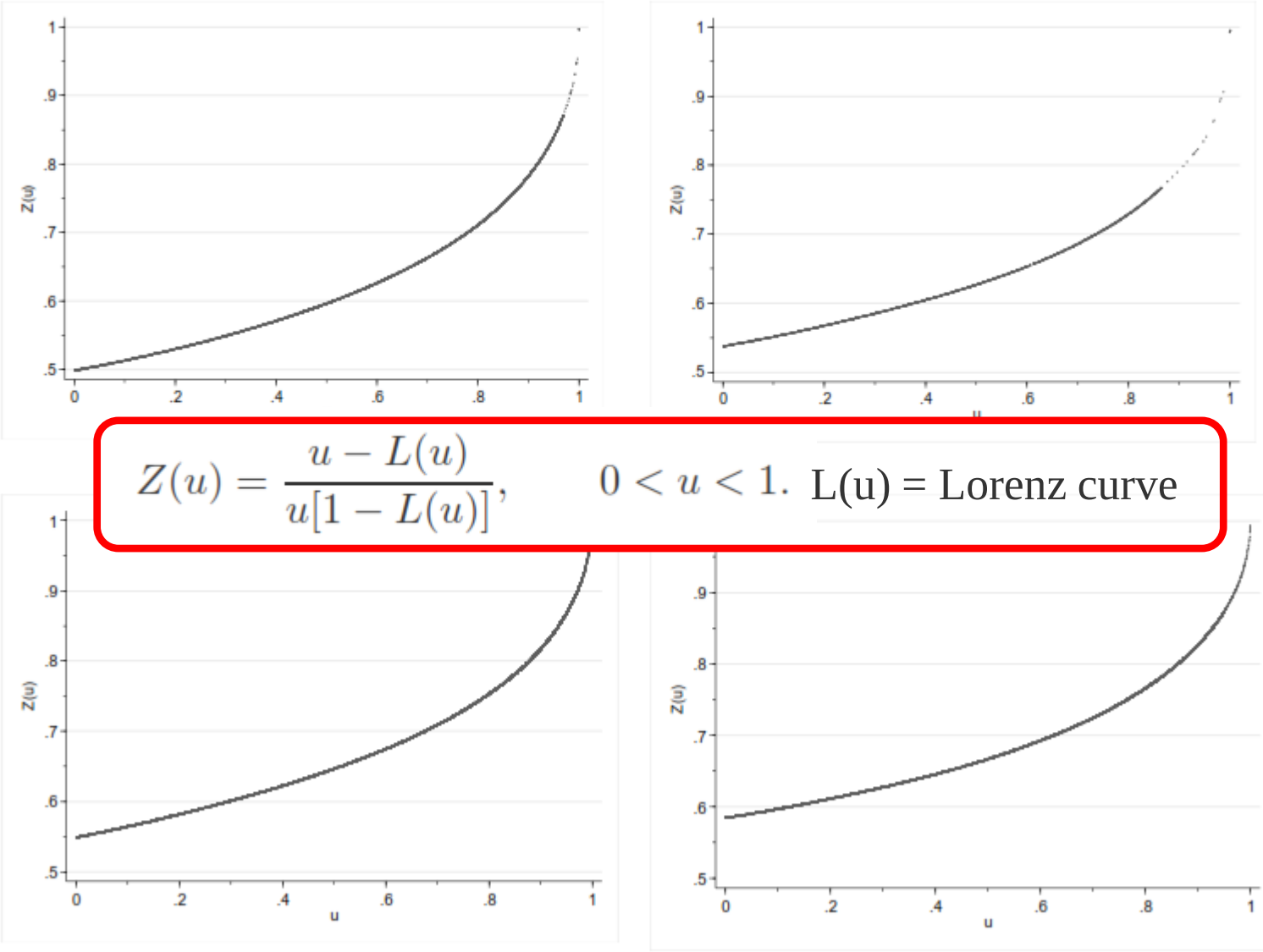
Pareto II, also known as Lomax distribution, where

$$F(x) = 1 - \left[1 + \frac{x}{b}\right]^{-\alpha}, \qquad x > 0. \tag{4}$$

Figure 4. Zenga plots for top incomes, tax return data, by threshold and year

Threshold = £60,000 p.a. Threshold = £120,000 p.a.

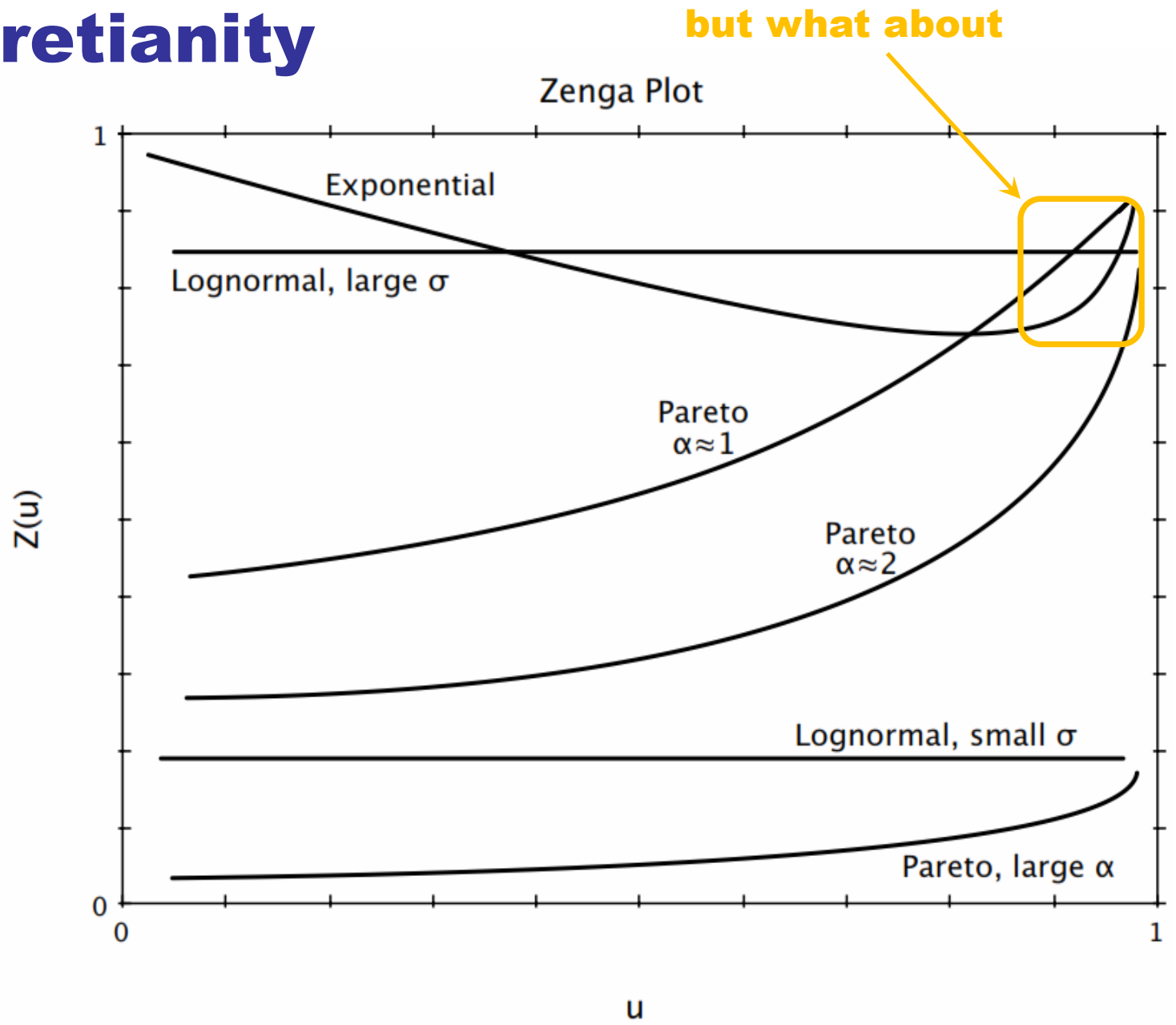
1996



$$Z(u) = \frac{u - L(u)}{u[1 - L(u)]}, \qquad 0 < u < 1.$$ L(u) = Lorenz curve

Notes. Author's estimates from SPI data. For plots for other years and thresholds, see Appendix D. On the Zenga plot, see the main text and Cirillo (2013).

Cirillo (2013) on Paretianity



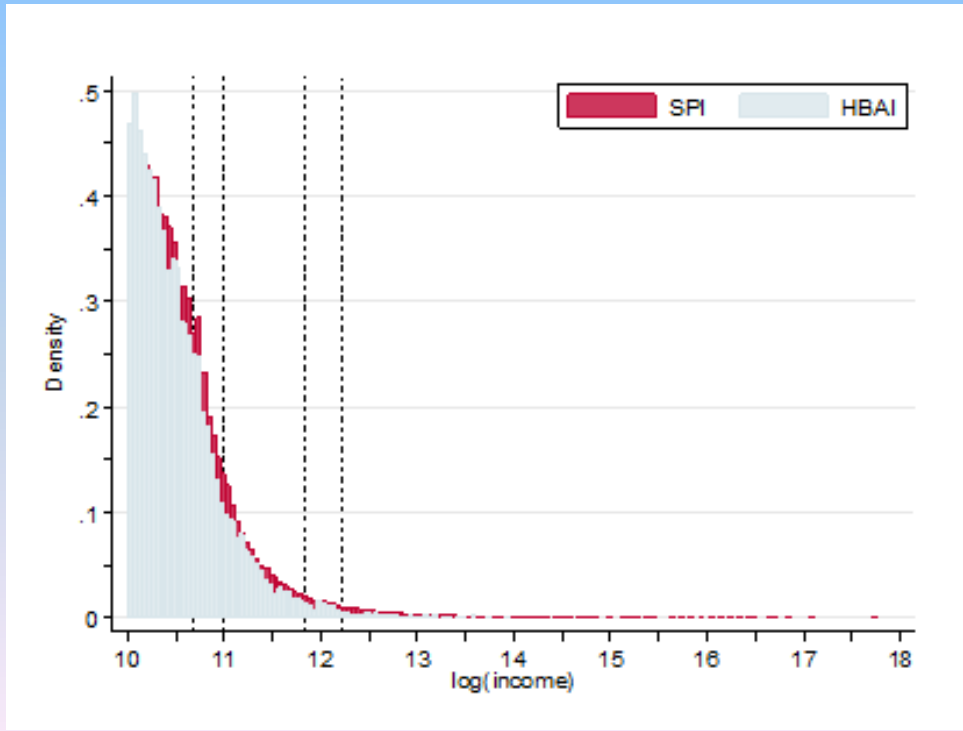
Statistical Estimation Conclusions

- for Pareto I, parameter is sensitive to choice of threshold
- for Pareto II, parameters more robust to changes in threshold

- but are we really sure the tail is Paretian?

Discussion – Both Papers

- good – used a range of inequality measures
- but recall – sensitivity to control totals
- what about under-reporting at the low end?
- why not take the sup of the two distributions overlaid, as in Figure 2?



Choice of Inequality Measures

- both studies looks at two kinds:
  - shares of top income quantiles, e.g. top 1%
  - “summary inequality measures that take into account income differences over the full income range (Gini coefficients and generalized entropy indices) rather than top income shares. “ (Jenkins p3)
- but what about
  - summary measure defined for non-positive incomes => cannot use logs; i.e. no Theils
  - polarization / disappearing middle indicators
  - definitions of “top-sensitivity”

the review of income and wealth

Explore this journal >

STASIS AMID CHANGE INCOME INEQUALITY IN CANADA 1965-1983

Michael Wolfson

First published: December 1986 Full publication history

we can avoid logs and still have Lorenz consistent inequality measures that span the range from bottom- to top-sensitivity (MCW, 30 years ago)

[4]The exponential measure is simply:

$$\frac{1}{n}\sum_{i=1}^{n} e^{-y_i}$$

for n incomes y_i with mean 1. This measure has the advantage over other "bottom sensitive" measures like the Theil-Entropy or Theil–Bernoulli measures that it is well defined for zero and negative incomes. Such incomes are fairly common and legitimate in actual microdata. The CV is the squared coefficient of variation, again for n incomes y_i with mean 1 this is:

$$\left[\frac{1}{n}\sum_{i=1}^{n}(y_i-1)^2\right]-1$$

Note that the CV is the most "top sensitive" measure used by Cowell (1984).

we can use income tax return data to adjust the upper tail of household survey data by sophisticated statistical matching (MCW, 28 years ago)

Review of Income and Wealth
Series 34, No. 4, December 1988

A PROTOTYPE MICRO-MACRO LINK FOR THE CANADIAN HOUSEHOLD SECTOR

BY HANS J. ADLER AND MICHAEL WOLFSON

Statistics Canada

This paper seeks to forge a link between Canadian macro and micro data relating to the household sector. The analysis is in three parts. The first part begins with National Accounts data on the personal sector. These data are adjusted to remove transactions relating to non-biological persons, so that the result is income and expenditure for the household sub-sector. The second part starts with the annual household survey used to collect income distribution data. These survey data are augmented in various ways to account for under-reporting and to add information from other micro data sets—particularly the periodic survey of household expenditure patterns and a sample of individual income tax returns. The result is a comprehensive, albeit partially synthetic, household micro data set. In the final part of the paper these two largely independent data sets are compared, and the general quality of the results is discussed.

Key Conclusions (MCW's take):

- official statisticians need to clean up their acts!
  - methodologies have been known for decades
  - underlying reason for inaction: survey statisticians (unlike National Accountants – think "macro-editing") are resistant to tampering with their surveys; willing to spend gobs of time on complex sample design and sampling error, but little on non-sampling errors – a seriously dysfunctional statistical culture
- academics need more non-parametrics
  - why bother with Pareto I vs II; go for the best possible household income distribution microdata set, even if substantially synthetic – then compute inequality indices (plus all sorts of other stats) directly

Conclusion – Topics for Discussion (I)

- should we (finally) recognize that non-positive incomes (and certainly net worth) are real, and therefore abandon the use of logs in inequality measures (and in econometric analyses more generally)?
  - e.g. no more Theil or Theil-Bernouilli = MLD
  - e.g. exponential for bottom-sensitive
- what about polarization versus {Lorenz dominance / Pigou-Dalton condition of transfers inequality measures} ?

Conclusion – Topics for Discussion (II)

- with currently available data, how best to improve upper tail of income distribution – especially (semi) parametric or not?
  - for countries with adequate microdata, why "dumb down" our methods?
- why not build on the synergy between these kinds of "corrections" for the income upper tail -and- the routine construction of proper micro foundations for the household sector – especially given Stiglitz / Sen / Fitoussi?