Reading list to accompany Canazei2025 lecture

Assessing whether income inequality trends are significant

Stephen P. Jenkins
(LSE, EUI, and IZA)
Email: s.jenkins@lse.ac.uk

National and international statistical agencies commonly provide annual reports about income inequality, and these are basis of discussions about inequality trends – whether inequality is rising, falling, or stayed the same. Other researchers also derive time series of income inequality estimates. Whether estimates of inequality trends are statistically significant is rarely discussed, however. This lecture considers how one might make such assessments in the case when one has access to unit record data from the household surveys used to derive the inequality series.

Headings

1. Overview of conventional approaches to inference, and the problems with them.
2. Approaches to inference that address the problems.
3. Application to yearly data, 1977–2018, on income inequality for the UK – the same data as used in the Office for National Statistics to produce its official series.

Abstract of lecture *(based on joint work with Nicolas Hérault, Université de Bordeaux)*

It has long been known that conventional approaches to finite sample inference for inequality measures based on asymptotic methods or the standard bootstrap do not perform well, even in large samples. See, e.g., Davidson and Flachaire (*Journal of Econometrics*, 2007) and Cowell and Flachaire (*Journal of Econometrics*, 2007). The source of the problem is the heavy-tailed nature of income distributions, with a relatively high prevalence of influential high-income outliers in datasets. Can researchers do better? The answer is yes, in principle. Davidson and Flachaire (DF, 2007) and Cowell and Flachaire (CF, 2007) propose inference based on a semi-parametric percentile-t bootstrap in which the upper tail is modelled parametrically by a Pareto distribution, and show using simulated data that accurate inference is achievable with moderately large samples. (For a related approach, see also Alfons et al., *JRSS(A)*, 2011.) In this paper, we provide the first systematic application of the DF-CF inferential approach to real-world income data for multiple years (yearly, 1977–2018), i.e., the same survey data as used by the UK Office for National Statistics for official statistics on inequality. As part of this work, we extend the DF-CF approach to deal with weighted data and a range of inequality indices (Gini, Generalised Entropy family, percentile ratios, and top income shares). We confirm that inequality indices are more precisely estimated using the DF-CF approach rather than the conventional asymptotic approach. But according to our DF-CF estimates we cannot reject the null hypothesis of no inequality difference between most pairs of years over the period 1977–2018 at conventional levels of statistical significance.

Selected reading

'Conventional' inference

Biewen, M., and Jenkins, S. P. (2006). Variance estimation for Generalized Entropy and Atkinson inequality indices: the complex survey data case. *Oxford Bulletin of Economics and Statistics*, 68 (3), 371–383.

Cowell, F. A., (1989). Sampling variance and decomposable inequality measures, *Journal of Econometrics*, 42 (1), 27–41.

Langel, M. and Tillé, Y. (2013). Variance estimation of the Gini index: revisiting a result several times published. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, 176 (2), 521–540.

Improved inference

Alfons, A., Templ, M., and Filzmoser, P. (2011). Robust estimation of economic indicators from survey samples based on Pareto tail modelling. *Journal of the Royal Statistical Society (C): Applied Statistics*, 62 (2), 271–286.

Cowell, F. A. and Flachaire, E. (2007). Income distribution and inequality measurement: The problem of extreme values. *Journal of Econometrics*, 141 (2), 1044–1072.

Davidson, R. and Flachaire, E. (2007). Asymptotic and bootstrap inference for inequality and poverty measures. *Journal of Econometrics*, 141 (1), 141–166.

Dufour, J.-M., Flachaire, E., and Khalaf, L. (2019). Permutation tests for comparing inequality measures, *Journal of Business & Economic Statistics*, 37 (3), 457-470.

Midões, C. and de Crombrugghe, D. (2023). Assumption-light and computationally cheap inference on inequality measures by sample splitting: the Student t approach, *Journal of Economic Inequality*, 21 (4), 899–924.

Software packages in Stata

For the conventional approach to inference:

All packages downloadable in Stata from SSC using *ssc describe <name>*, where <name> is ..

geivars (by Jenkins, 1998: Cowell 1989 approach; GE indices; allows weights)
svylorenz (by Jenkins, 2005: Langel & Tillé 2013; Gini coefficient; allows weights, clustering, stratification)
svygei_svyatk (by Biewen & Jenkins, 2003: Biewen & Jenkins 2006; GE and Atkinson indices; allows weights, clustering, stratification)
dstat (by Ben Jann, 2023; can do all of the above!)

To apply the CF/DF improved inference methods, Nicolas Hérault and I are using the following:

Built-in Stata commands, including: *bsample, forvalues* loops, *postfile* (to collect results)

dstat (op. cit.)
paretofit (Jenkins and Van Kerm, downloadable from SSC) and *paretofit_obre2* (Van Kerm)