

Progressivity of US Tax & Transfer System

Motivation

The analysis performed in this exercise is relevant to policy aimed at achieving some degree of redistribution. The tax code is an instrument used to raise revenue for public goods, but also redistribute income. How successful are the different components of the tax & transfer system, according to different metrics income equality?

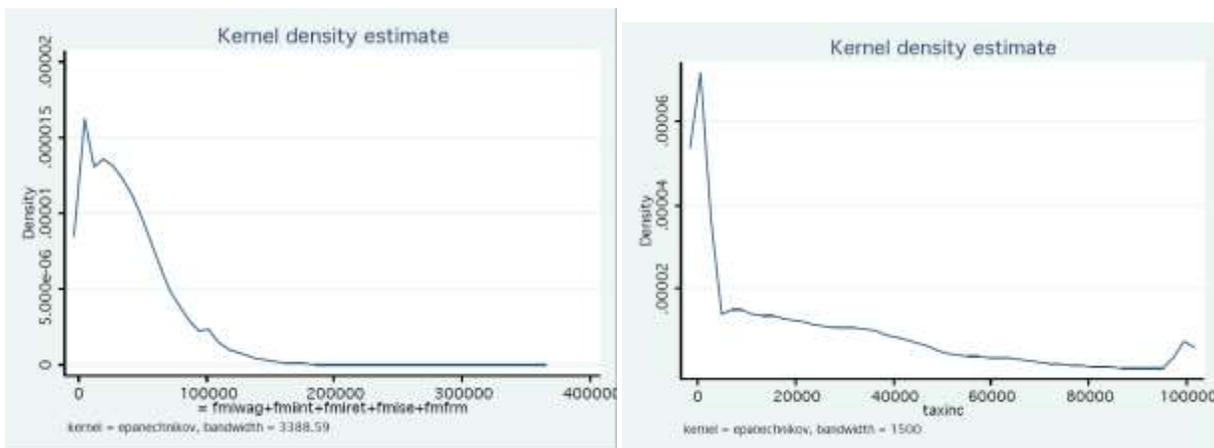
Positive Question

What does the income distribution look like for different definitions of income; pre-tax pre-transfer income, total taxable income, and after-tax after-transfer income? What are different measures of progressivity? How successful are the different components of the US Tax and Transfer system at redistributing income according to different progressivity measures, and starting with different before-tax income distributions?

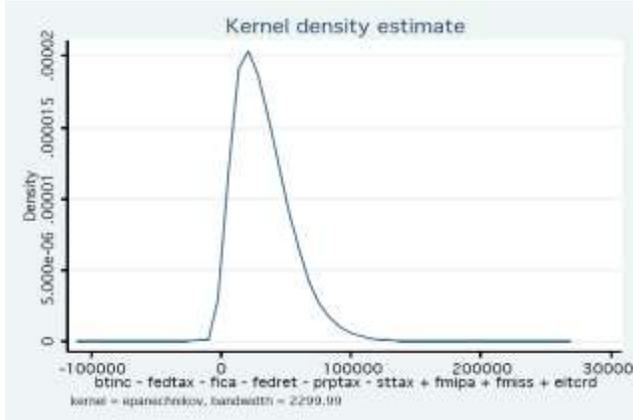
Methodology

To analyze income inequality in the US I generated multiple kernel density estimates. The Stata default bandwidth, Epanechnikov, generally yielded the best interpretation of the income distribution as it did not smooth out too many features of the distribution, however, also did not over-respond to all data points and become confusingly jagged.

Directly below are the Epanechnikov Density Estimates of Before T&T Inc, & Taxable Income:



Also, I created an Epanechnikov Density Estimate of after T&T Income:



I chose to use the following measures of progressivity to analyze the T&T (tax & transfer) system: change in standard deviation, change in coefficient of variation, Lorenz Dominance, and change in Gini Index. For the single-dimensional measures (std deviation, coefficient of variation, and Gini Index), the values are listed below where the baseline is btinc.

$$btinc_i = fmiwag_i + fmiint_i + fmiret_i + fmise_i + fmifrm_i \forall i$$

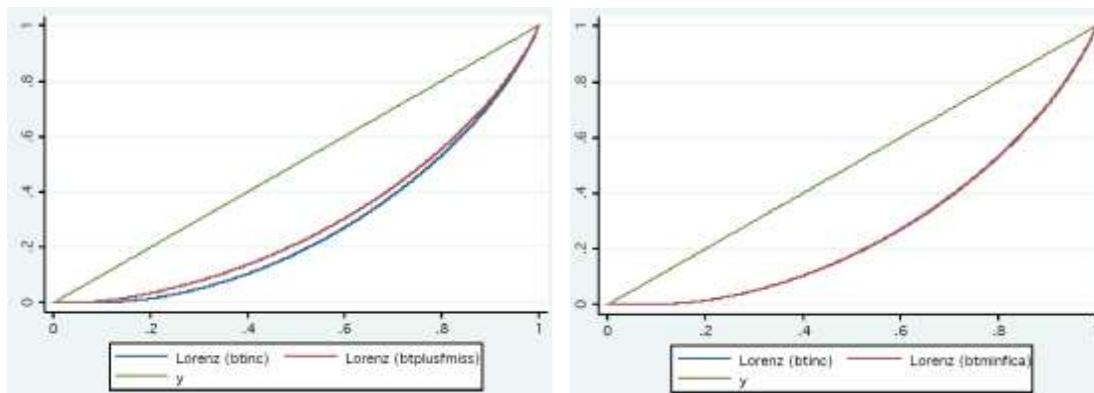
$$atinc_i = btinc_i - fedtax_i - fica_i - fedret_i - sttax_i - prptax_i + eitcrd_i + fmipa_i + fmiss_i \forall i$$

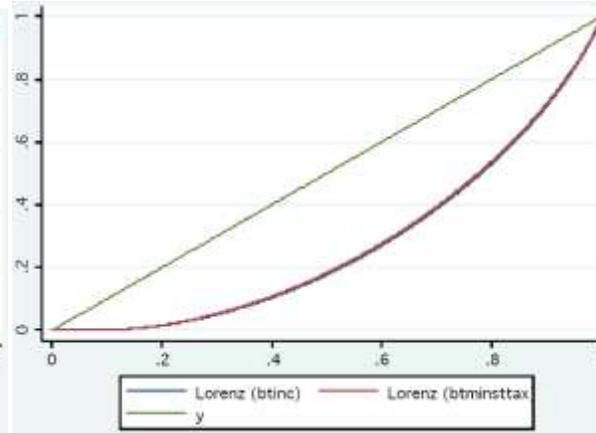
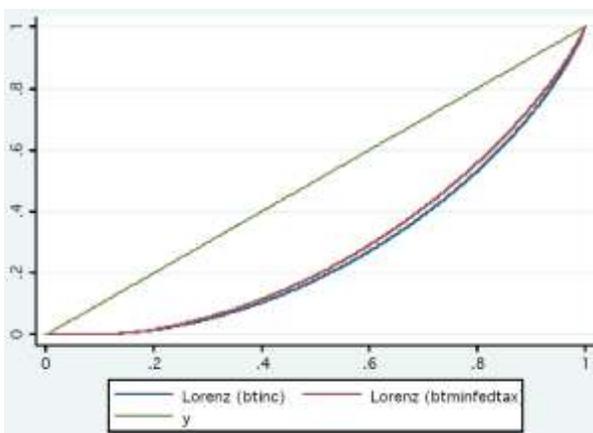
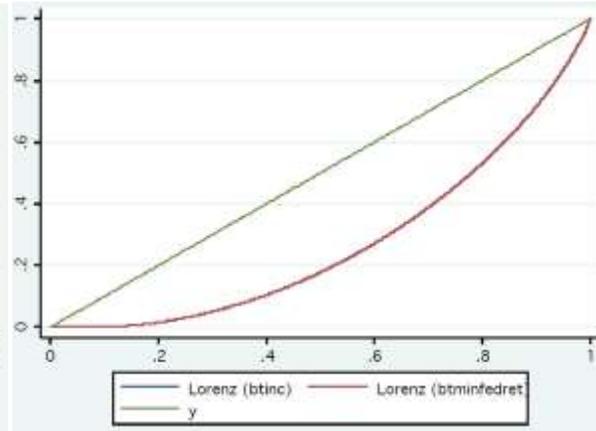
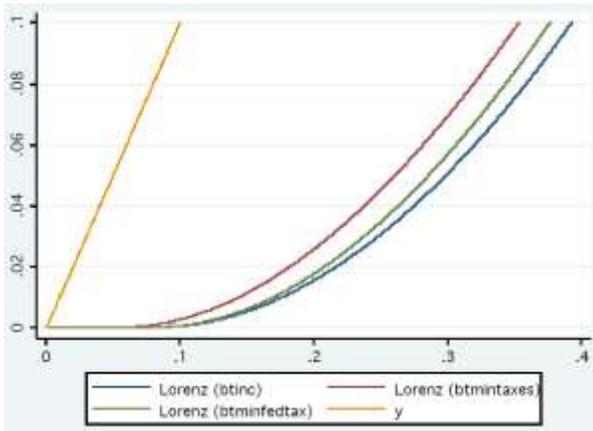
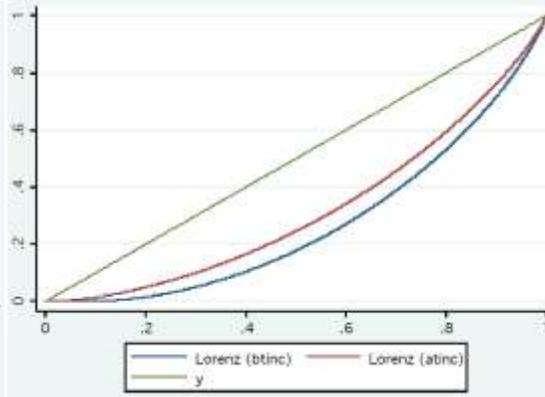
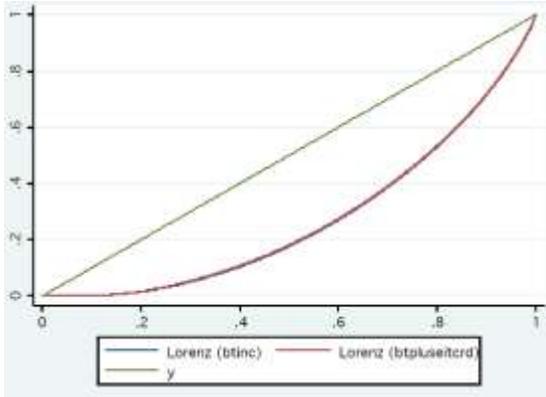
It is important to note that all progressivity measures were calculated as if they were the only tax and transfer component. For example, for Fica taxes (btminfica), the Gini Index is calculated with the Lorenz curve of before T&T income minus just fica payments.

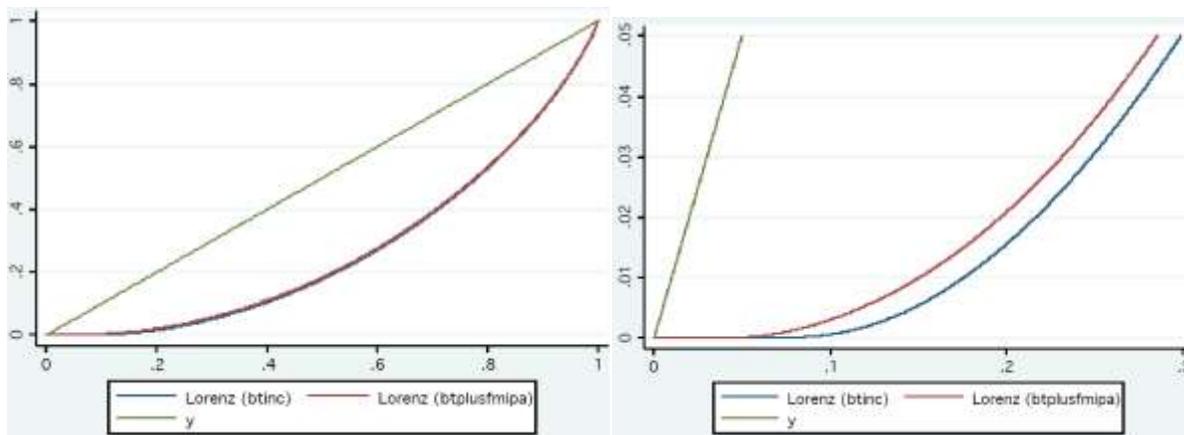
Variable:	Mean:	Std Dev:	Kern Dens:	Coeff of Var:	Lorenz:	Gini:
btinc	38941.88	33398.8	Done	0.85765762	Done	0.4564
taxinc	23147.4	26403.54	Done	1.14066979	Done	0.5961
atinc	32698	22479.06	Done	0.68747507	Done	0.3709
btmintaxes	29809.95	23859.13	N/A	0.80037471	Done	0.438
btminfedtax	33904.19	26649.21	N/A	0.78601524	Done	0.431
btminfica	37334.39	32223.97	N/A	0.86311762	Done	0.4579
btminfedret	38858.88	33331.31	N/A	0.85775272	Done	0.4563
btminsttax	37502.93	33331.31	N/A	0.88876549	Done	0.4507
btminprptax	37977.05	32934.2	N/A	0.86721322	Done	0.4614
btplustransfers	41829.93	31956.36	N/A	0.7639592	Done	0.4012
btpluseitcrd	39066.75	33305.48	N/A	0.85252753	Done	0.4534
btplusfmipa	39215.75	33150.12	N/A	0.84532669	Done	0.4487
btplusfmiss	41431.18	32332.79	N/A	0.78039752	Done	0.4119

Concerning data treatment, some agents were found to have negative after T&T income, usually due to large property taxes as well as other taxes. For the most part, these agents were the elderly dissaving. This data analysis suggests that policy should be concerned with, and analyze redistribution effects on permanent income, as it is a more lifestyle-relevant variable. Nonetheless, since negative income values distort Lorenz curves and Gini Index calculation, the procedure was adopted to truncate these observations. As an alternative, I could have shifted all income values to make them all positive, but this method has fallacies of its own. Any translation or shift changes relative income values and thus the curvature of the Lorenz Curve and Gini Index calculation. In addition, some income variables evidenced top coding, but just for some observations. There is difficulty in trying to identify which individuals were from the population that was eventually top-coded, and which were not. With this information, I could have used a two-step procedure to create a predicted income distribution, but without it, I resorted to leaving the variable as is, with the top-coding. Another option would have been to multiply all top-coded values by a greater-than-one constant to get away from the obvious under-estimation of the distribution. The danger is that the multiplication leads to a worse over-estimation of the distribution.

The best tool for analyzing progressivity of a tax and transfer component is the Lorenz Curve. It allows the econometrician to see the changes in progressivity and where in the distribution that redistribution is most prevalent.







Perhaps due to the way that I dealt with the top-coded incomes and truncated negative after tax income, I found that no T&T system components yielded crossing Lorenz curves. Every component of the T&T system was therefore found to be either purely progressive or purely regressive, and for purposes of this exercise, Lorenz dominance was a complete metric, although it generally is not. Of interest in the Lorenz curves is where the progressivity or regressivity is most prevalent, hence some of the zoomed in Lorenz Curves.

Results

The density estimates all evidence income distributed in a lognormal fashion. The taxable income distribution appeared to have slightly more dispersion than the before t&t distribution, but it was not clear. Also not clear, but suggested was that the after t&t income had the least dispersion of the three. I would look at the density estimates for general characteristics of the data such as evidence of topcoding and effects of truncation, but to assess the progressivity of various t&t instruments, the Lorenz curves provide a more powerful analytical tool.

Many of the t&t instruments had negligible effect on the Lorenz curves indicating that those instruments were flat taxes and that they therefore generate government revenue without meddling in the redistributive objective of government. Other instruments, especially the transfers and the fedtax and the entire system, place a higher emphasis on the redistributive objective, tax more progressively according to all metrics of progressivity, and result in a more equalized income distribution. The government definitions of taxable income include regressive concepts however since it was found that the distribution of taxable income was Lorenz dominated by the before t&t Lorenz curve. The after t&t distribution Lorenz dominated both and implies that they overall government t&t system is redistributive. Given these results, there were no inconsistencies among the different progressivity measures. What was more progressive according standard deviation, was also more progressive according to coefficient of variation, Gini Index, and Lorenz dominance.