*01.172 THEIL INDEXES Concept and Applications see details in ** 01.17

- **A. Concept:** Theil-T index assess how much a given income distribution (each person receive y_i of total income) is away of a perfect uniform distribution (each person receive 1/n of total income), or the redundancy degree in relation to the latter, weighting each observation by its share in total income.

$$T = \ln n - H(x) = \sum_i y_i \ln \frac{y_i}{1/n}$$

$0 \le T \le \ln n$, that is, we have $T = 0$ in the case of a perfect egalitarian distribution and $T = \ln n$ in the case of maximum inequality. Theil-T index assess how much a given income distribution (each person receive y_i of total income) is away of a perfect uniform distribution (each person receive 1/n of total income), or the redundancy degree in relation to the latter, weighting each observation by its share in total income. If in ln in *nits (natural logs units)*,

The second Theil measure of inequality is Theil-L index, defined by the following formula:

$$L = \sum_{i=1}^{n} \frac{1}{n} \log \frac{1/n}{y_i} = \frac{1}{n} \sum_{i=1}^{n} \log \frac{y_i}{1/n} (-1)$$

It inverts the redundancy comparison and weights. While in Theil T the inequality factors of weighting within the groups are the share of income, in Theil L the inequality factors of weighting within the groups are their respective population.

THEIL (General Entropy) INDEXES

C. General Entropy S- measure nests Theil T and Theil L as special cases.

$$S = \frac{1}{\varepsilon(1-\varepsilon)}\left[1 - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^{1-\varepsilon}\right]$$

OBS: **ε=0 Theil T; ε=1 Theil L; ε= -1 Coefficient of Variation**

B. Dual: $U_2 = \phi + (1-\phi)U_1$ allows to compare different inequality measures in the same 0 to 1 scale **The Dual of the Gini Index is the Gini Index G* = G (1-%) + %, % are** new 0s a way to proceed with maximum inequality (G=1) so is adding top incomes. One can use this formula for introducing both ends of income distribution. As the dual of any inequality measure since its dual transformation measures in the Gini scale. Applying this formula $U_2 = \phi + (1-\phi)U_1$ to the to the Theil –T we get $T2 = T1 - \ln(1-\phi)$. A fully decomposable overall measure of social welfare inspired on Sen (1973) is $SW = mean.(1 - U_{T1})$. Since the Theil L does not admit null values, it also does not admit a Dual measure.

D. Intra and Inter Groups Decomposition of Theil T (Theil L allows a similar formula)

$T = T_e + \sum_{h=1}^{K} Y_h T_h$ Where, $T_e = \sum_{h=1}^{k} Y_h \log \frac{Y_h}{\pi_h}$ is the Theil T between groups and $T_h = \sum_{i=1}^{n_h} \frac{y_{hi}}{Y_h} \log n_h \frac{y_{hi}}{Y_h}$ is the Theil intra groups. Therefore $\sum_{h=1}^{K} Y_h T_h$ is the weighted average of intra-groups Theil Ts. **Te / T is the Contribution of a certain characteristic to inequality (say how much schooling (or gender) explains <u>exactly</u> total inequality?). Alternative to mincerian regressions based decompositions.**

Theil Between groups ← **Theil Total** → **Theil Within groups**

Other application: Does per capita Household Income underestimates true inequality?

Theil Between Families (per capita) ← **Theil Total** → **Theil Within Families**

***Applying Decomposition to Inequality & Temporal Variability (Mobility, Risk or measurement error)**

Dispersion of Mean Income (earned over time) between people ← **Theil Total** → **Mean variability of income in relation to the mean across time for each person**

We have used the micro-longitudinal aspect of PME/IBGE to the Real Plan Stabilization. The main result here is that the fall of month-to-month inequality measures observed after the fall of inflation in 1994 drastically overestimates the fall of inequality when one compares it with mean earnings over four months. The greater fall of traditional inequality measures on a monthly basis in comparison with measures on a four-month basis is explained by the fall of the individual volatility measures following the sharp decline in inflation rates observed in this period. In sum, stabilization produced more stable earnings trajectories (i.e., lower temporal inequality (in fact, volatility) of individual earnings). On the other hand, the observed fall of inequality *stricto sensu* was much smaller than inequality measures based on monthly measures would have suggested. In sum, the post-stabilization fall in inequality for the group of population is much higher on a monthly basis (as traditionally used in Brazil) than when one uses mean earnings over four months. The fall of Theils is around 4 times higher when one uses the former concept.

Inequality of Opportunity in Brazil

François Bourguignon, Francisco H.G. Ferreira and Marta Menéndez (2007)

- **Drawing on the distinction between variables of 'circumstance' (not in control of the individual) and 'effort' (in control of the individual)** in John Roemer's work on equality of opportunity, their approach is to **simulate the reduction in earnings inequality which would attain if differences in circumstance variables were eliminated**.
- **The five observed circumstances** (father's and mother's education; father's occupation; race; and region of birth) **are found to account for between 10% and 37% of the Theil index, when accounting for possible biases**. **Parental education is the most important circumstance affecting earnings**, but the occupation of the father and race also play a role. On average, some 60% of the effect of these circumstances operates directly through earnings, while the remaining 40% or so operate by affecting the level of efforts expended by individuals. The decomposition is applied to the distribution of male earnings in urban Brazil in 1996.

Table 5. The Contribution of Unequal Opportunities to Earnings Inequality, Urban men in Brazil, Actual and Counterfactual Theil coefficients (and ratios) for 5-year cohorts of men.

	b1936_40	b1941_45	b1946_50	b1951_55	b1956_60	b1961_65	b1966_70
Total Observed Inequality (1)	**0.873**	**0.997**	**0.759**	**0.655**	**0.706**	**0.580**	**0.566**
PANEL 1: "Complete" (observed) opportunity share of earnings inequality							
(Upper bound estimate)	(0.692)	(0.675)	(0.644)	(0.531)	(0.572)	(0.415)	(0.507)
Mean estimate (2a)	**0.654**	**0.656**	**0.619**	**0.519**	**0.562**	**0.407**	**0.494**
(Lower bound estimate)	(0.642)	(0.632)	(0.609)	(0.516)	(0.555)	(0.402)	(0.485)
(Upper bound estimate)	(0.264)	(0.366)	(0.199)	(0.212)	(0.214)	(0.307)	(0.143)
Mean share ((1)-(2a))/(1)	**0.251**	**0.343**	**0.184**	**0.208**	**0.205**	**0.298**	**0.128**
(Lower bound estimate)	(0.207)	(0.323)	(0.151)	(0.190)	(0.190)	(0.284)	(0.104)

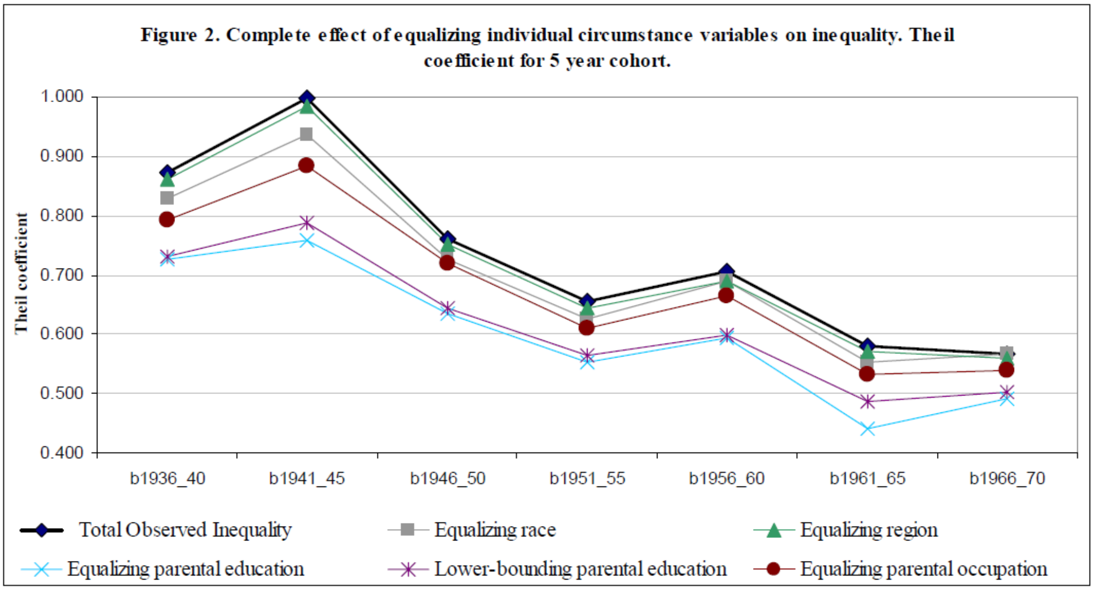
For men born between 1941-45, elimination of inequality due to observed circumstances reduces the Theil index from 0.997 to some value between 0.632 and 0.675, with a mean estimate of 0.656. These estimates indicate that 32%-37% of earnings inequality in this cohort is accounted for by unequal opportunities – due only to those five observed circumstance variables.

Table 6. Contribution of individual circumstance variables to earnings inequality; by cohort.

	b1936_40	b1941_45	b1946_50	b1951_55	b1956_60	b1961_65	b1966_70
Total Observed Inequality	**0.873**	**0.997**	**0.759**	**0.655**	**0.706**	**0.580**	**0.566**
Equalizing race	0.830	0.936	0.727	0.626	0.689	0.552	0.567
Equalizing region	0.860	0.983	0.751	0.645	0.689	0.571	0.560
Equalizing parental education	0.726	0.759	0.634	0.553	0.595	0.440	0.491
Lower-bounding parental education	0.730	0.787	0.643	0.565	0.599	0.486	0.503
Equalizing parental occupation	0.793	0.883	0.719	0.611	0.666	0.532	0.540

NOTE: Mean estimates from the 90 counterfactual distributions corresponding to the "90th confidence interval" of unbiasedness of the coefficients.

- The complete effect of equalizing each individual circumstance variable, while controlling for all others, is shown above for the Theil coefficient separately for each cohort.
- Parental education plays the largest role in determining inequality, across all cohorts. If a lower bound (of six school years) is imposed, as if schooling were compulsory (*de facto*, rather than merely *de jure*) until a certain age, the contribution of parental education to reducing earnings inequality is not much smaller. This suggests that it is the inequality of education at the bottom of the distribution that matters most to explaining the contribution of opportunities to earnings inequality.



- The most promising policies for reducing inequality of opportunities in Brazil might be those aimed at reducing the effect of parental education on the child's schooling and earnings.