Inequality Dynamics:

A Note on the Time-wise Decompositions of Entropy-based Measures of Inequality with a Special Focus on Theil Measures

by Pedro Conceição

pedroc@uts.cc.utexas.edu

UTIP and

Center for Innovation, Technology and Policy Research

Instituto Superior Técnico, Lisbon, Portugal

UTIP Working Paper No. 19

April 10, 2001

Abstract: This paper discusses the dynamic properties of generalized entropy measures of inequality, deriving an explicit expression linking changes in inequality with changes in the growth income. This establishes a time-wise decomposition for the family of generalized entropy measures of inequality, showing that the rate of change in the distribution of income can be divided into two parts. One is purely a function of the levels and rates of change of macroeconomic variables; a second part depends exclusively on the micro nature of the distribution process. We then focus on a subset of generalized entropy inequality indexes, focusing on Theil measures. By assuming that the distribution process follows a specific rule, we further simplify the time-wise decomposition formula, establishing a direct relationship between the rate of income growth and the rate of change in inequality as measured by Theil measures. A specific application of this formula to the simulation of an economy with a constant average positive growth rate produces behavior consistent with the Kuznets hypothesis.

1- INTRODUCTION

Theil (1967) inspired the development of the entropy-based measures of inequality. Shorrocks (1980) following the pioneering work of Bourguignon (1979), defined these measures, for $a \neq 0,1$, as

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

where y_i is the income of individual i and Y is the total income available in a society composed of n individuals. Theil's T measure corresponds to a=1:

[2] $$T = \sum_{i=1}^{n}\frac{y_i}{Y}.\log\left[\left(\frac{y_i}{Y}\right)\Big/\left(\frac{1}{n}\right)\right]$$

and a=0 corresponds to Theil's second measure:

[3] $$L = \sum_{i=1}^{n}\frac{1}{n}.\log\left\{\left[\left(\frac{y_i}{Y}\right)\Big/\left(\frac{1}{n}\right)\right]^{-1}\right\}$$

Any entropy measure of the form [1] is Lorenz consistent, fulfilling all the desirable properties of "well behaved" inequality measures. An important characteristic unique to these entropy based measures, immediately noted by Theil for the special case of his T and L indices, is that they are decomposable across groups that parse the individuals of the population into mutually exclusive, completely exhaustive, "bins". Overall inequality can be separated into a between group component and a within group component. If we consider that the population is divided into k groups, each with n_j individuals, $j=1, \ldots, k$, then, in general, the decomposition takes the form:

[4] $$E_a = Between\_groups + Within\_groups = E_a(1,\ldots j,\ldots,k) + \sum_{j=1}^{k} R_j E_a(j)$$

where the weights, R_j, for entropy measures of the type [1], are given by:

[5] $$R_j = \left(\frac{Y_j}{Y}\right)^a \left(\frac{n_j}{n}\right)^{1-a}$$

where Y_j is the total income and n_j the total population in group j. Shorrocks (1980, 1984) proved the strong result that any measure of inequality is decomposable in the form of [4] if and only if it is an entropy-based measure of inequality of the form [1]. More formally, Shorrocks showed that any inequality measure is Lorenz-consistent, normalized, continuous, and additively decomposable if and only if it is a multiple of a generalized entropy measure. Therefore, entropy-based measures of inequality are important whenever one is interested, whatever the reason, in finding or using a group-wise partition of an inequality measure. An important reason to use the decomposition properties of these inequality measures is associated with the possibility of constructing long and dense time-series of inequality, as explored in Conceição and Galbraith (2000).

The explicit formula for the decomposition of between and within group inequality is given by:

[6] $$E_a = E_a^{'} + E_a^{w} = \frac{1}{a(1-a)}\left(1 - \sum_{j=1}^{k} R_j\right) + \sum_{j=1}^{k} R_j E_a(j)$$

where

$$[7] \qquad E_a(j) = \frac{1}{a(1-a)}\left(1 - \sum_{i=1}^{n_j} r_i^{\,j}\right)$$

and

$$[8] \qquad r_i^{\,j} = \left(\frac{y_i}{Y_j}\right)^{a}\left(\frac{1}{n_j}\right)^{1-a}$$

Expressions [6] through [8] highlight the self-similar property of the generalized entropy measures at different levels of aggregation. In essence, the entropy measures quantify the extent to which the income shares (Y_j/Y for groups or y_i/Y_j for individuals in a group) differ from the population shares (n_j/n for groups and $1/n_j$ for individuals in a group)[1], as explored in Conceição and Ferreira (2000).

Mostly, the decomposition properties of the entropy-based measures have been explored in the literature in a static context: group and individual properties are defined for a set point in time. In this paper we explore the "decomposition" properties across time. Figure 1 helps to illustrate the conceptual reasoning behind the time-wise decomposition. The graph on top shows a hypothetical distribution of income of the

---

[1] Naturally, this interpretation is equally valid for inequality across individuals, where [1] can be written as $E_a = \frac{1}{a(1-a)}\left[1 - \sum_{i=1}^{n} r_i\right]$ with $r_i = \left(\frac{y_i}{Y}\right)^{a}\left(\frac{1}{n}\right)^{1-a}$ giving the disparity between the income and population shares of each individual in the context of the total number of individuals.

population at instant t_0. To the left, we divide this population into two groups, g_1 and g_2, making a static analysis (no changes through time are allowed). With the decomposition given by [4] we know how to relate the inequality in the initial distribution, with the inequality within and between the distribution of the populations in groups g_1 and g_2.

Consider now the graph below and to the right. There we have the original distribution for t_0 and the distribution of the entire population at an instant $t_1 > t_0$. Is there any way to establish a formal relationship between the inequality based entropy base measures for t_1 and t_0?

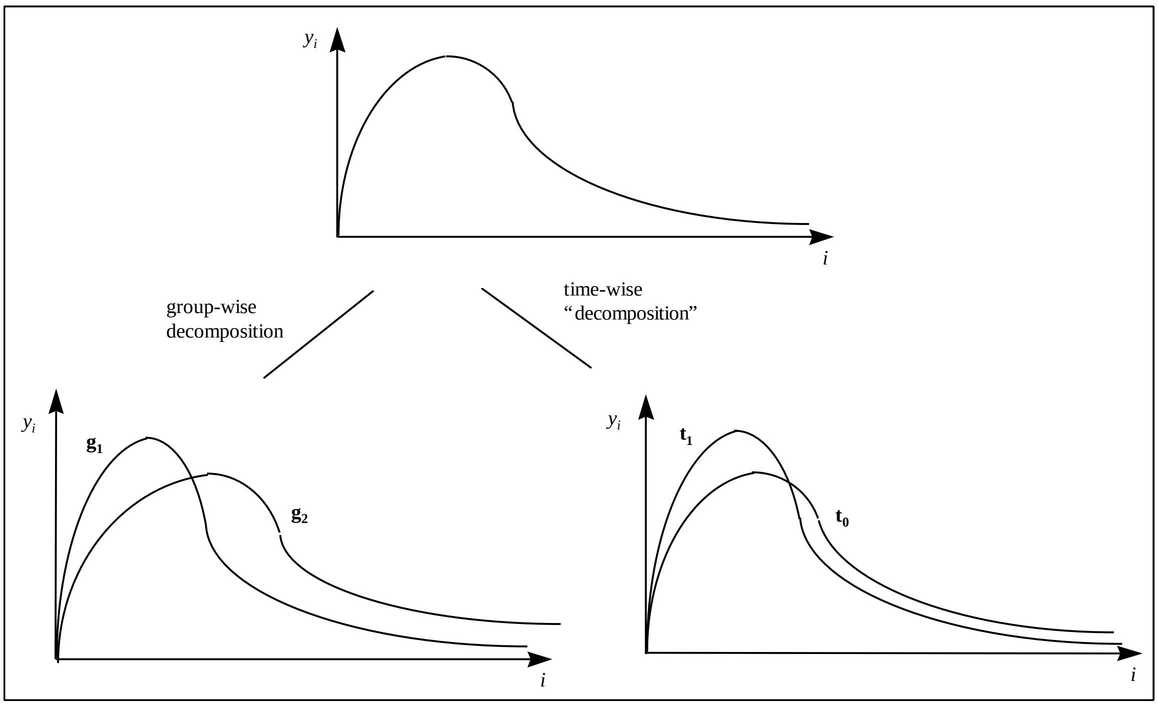

Figure 1- Group-wise decomposition and time-wise "decomposition" of income distribution.

A way to establish a relationship between the measures of inequality for the two different time periods is by deriving an explicit expression of the rate of change:

[9] $$\frac{E_a(t_1) - E_a(t_0)}{E_a(t_0)} = f(\bullet, t_1, t_0)$$

The objective of the paper is to derive such an expression, making explicit the functional form to the right of [9]. We call expression [9] a time-wise decomposition in the sense that, as in the group-wise decomposition, we are able to have a relationship of the inequality measures in two different time periods; while in the group wise decomposition a static total inequality is decomposed into two static measures of inequality, in the time-wise decomposition two measures of inequality in two consecutive time periods are linked through the change (or the rate of change) in inequality.

We consider first, in section 2, the discussion of the time change of general entropy measures. In section 3 we focus on the Theil measures. The fundamental result of both of these sections is the decomposition of the change in inequality measures over time into changes at different levels of aggregation. In particular, we are able to separate changes exclusively associated with macro variables from micro behavior. We the present in two sections two application of these general results. In section 4 we focus on changes in Theil measures and show that by defining income distribution rules (which rule the micro behavior) we are able to express the changes in the Theil measures exclusively as a function of macro variables. One of the distribution rules analyzed produces a relationship between income growth and inequality change consistent with the Kuznets hypothesis. In section 5 we approach the problem of linking together the dynamics of the between-group component of Theil index at different levels of aggregation. Section 6 summarizes and concludes the paper.

2- TIME-WISE DECOMPOSITION OF GENERAL ENTROPY BASED MEASURES OF INEQUALITY

The results in this section are valid for any generalized entropy based measure of inequality for which $a \neq 0,1$. The special cases of the Theil measures will be analyzed in section 3. We begin by taking the time derivative of expression [6], which results in:

[10] $$\frac{dE_a}{dt} = \dot{E}_a = \sum_{j=1}^{k} \dot{R}_j \left[E_a(j) - \frac{1}{a(1-a)} \right] + \sum_{j=1}^{k} R_j \dot{E}_a(j)$$

Expression [10] reflects the usual way in which changes in inequality over time are normally understood. The last term on the right-hand side of the last equality sign represents the total change in inequality associated with changes within each of the groups, while the first term captures the dynamics of the between group changes.

Focusing, first, on this last component of [10], making explicit the time-derivative of R_j gives:

[11] $$\dot{R}_j = R_j \left[a(g_j - g) + (1-a)(p_j - p) \right]$$

where g_j is the rate of change of income for group j and g is the overall rate of change of income, p_j is group's j rate of population change and p is the overall rate of population change:

[12] $$\begin{cases} g_j = \dfrac{\dot{Y}_j}{Y_j} \quad , \quad g = \dfrac{\dot{Y}}{Y} \\ \\ p_j = \dfrac{\dot{n}_j}{n_j} \quad , \quad p = \dfrac{\dot{n}}{n} \end{cases}$$

Expression [11] reflects the commonly used decomposition of the changes in inequality across groups into an income effect and an allocation effect. The income effect, associated with changes in the average income of each group as well as the overall growth of income, is captured by $\alpha(g_j - g)$: for each group, the growth rate of income of that group is contrasted with the overall growth rate of income. The impact of the allocation effect is given by an expression that is formally equivalent to that of the income effect: (1 - α)(p_j- p). Each of these two effects is multiplied by R_j and enters in the summation of expression [10] "weighted" also by the within group inequality:

$$[13] \qquad \sum_{j=1}^{k} \dot{R}_j \left[E_a(j) - \frac{1}{a(1-a)} \right] = \sum_{j=1}^{k} \left\{ R_j \left[E_a(j) - \frac{1}{a(1-a)} \right] \left[a(g_j - g) + (1-a)(p_j - p) \right] \right\}$$

Expanding in [13] the terms with the growth rates of income and population leads to an expression that decomposes the change in inequality between groups into two components:

$$[14] \qquad \begin{aligned} &\sum_{j=1}^{k} \dot{R}_j \left[E_a(j) - \frac{1}{a(1-a)} \right] = \\ &= \sum_{j=1}^{k} R_j \left[E_a(j) - \frac{1}{a(1-a)} \right] \left[a g_j + (1-a) p_j \right] - \\ &- \left[E_a - \frac{1}{a(1-a)} \right] \left[a g + (1-a) p \right] \end{aligned}$$

The first component in [14] – that is, the second line of the expression after the equality sign – accounts for the dynamics of income and population at the group level. The third line in expression [14] captures the total changes in population and income. Therefore, this last component can be understood as a macro component of changes in inequality, while the second line in [14] can be considered a meso component (since it is between the macro level and the individual, or micro, level).

There is a clear symmetry between the structure of the meso and of the macro components. Each term in the summation of the meso component has the same structure of the macro component, except for the weight R_j. This structure corresponds to multiplying the level of inequality (minus the constant 1/(1-α)) by a linear combination of the rates of changes of income and of population, with the linear combination being given by multiplying the rate of change of income by α and the rate of change of population by (1-α).

The time derivative $\dot{E}_a(j)$ requires a further transformation before we can move into the actual computation of an explicit expression for the change in the within group inequality. The transformation is needed because the upper limit, n_j, for each j, in the summation that defines $E_a(j)$, given by [7], changes over time. In other words, the population of each group is not constant, and the upper limit in the summation is varying over time. To deal with this problem, we transform the discrete distribution of income across individuals defined in [7] into a continuous distribution.

For each group j, the distribution of income is assumed to be given by a density distribution function $f_j(z)$, where z is the parameter that represents the level of income, with $z \in [0, \infty[$. Therefore (omitting the explicit dependence on time to simplify the notation):

[15] $$\frac{Y_j}{n_j} = \int_0^{\infty} z f_j(z) dz$$

The challenge, now, is to transform expression [8], which is discrete in income, into a continuous expression in income. This transformation has to be followed by the transformation of the summation in [7] into a continuous integral over z. First, consider a neighborhood dz of z. Assuming that the neighborhood is small enough, the number of individuals that it contains is $n_j.f_j(z)dz$, since $f_j(z)$ can be considered constant for $z \in \left]z - \frac{1}{2}dz, z + \frac{1}{2}dz\right[$ (see Figure 2). Additionally, all individuals in this neighborhood can be assumed to earn the same income z, because dz is very small, and therefore the total income in the neighborhood is $z.n_j.f_j(z)dz$.

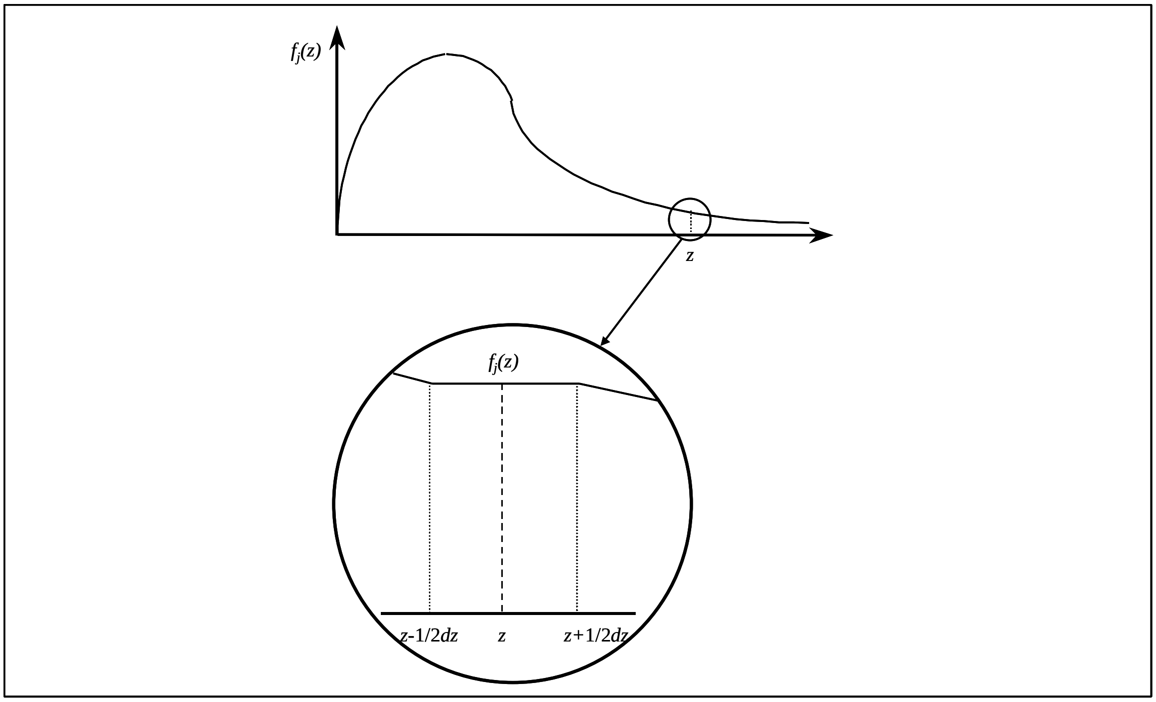

Figure 2- Defining a neighborhood of the density distribution function.

Each neighborhood has more than one individual, so the formula we need to use to compute inequality is the between-group expression (or “between-neighborhood”, to be more precise), where [5] is now transformed into a continuous formulation with reference to group j:

[16]
$$R_j(z,dz) = \left(\frac{zn_j f_j(z)dz}{Y_j}\right)^a \left(\frac{n_j f_j(z)dz}{n_j}\right)^{1-a} = \left(\frac{z}{Y_j}\right)^a \left(\frac{1}{n_j}\right)^{-a} f_j(z)dz =$$
$$= \left(\frac{zn_j}{Y_j}\right)^a f_j(z)dz$$

The “between-neighborhood” inequality is then given by:

[17] $$E_a(j)=\frac{1}{a(1-a)}\left[1-\int_0^{\infty}\left(\frac{zn_j}{Y_j}\right)^a f_j(z)dz\right]$$

The between-neighborhood inequality is, indeed, the total within group j inequality, as we indicate in [17], since we assume that all individuals in each neighborhood earn the same income z and, therefore, there is no within-neighborhood inequality. There is a subtle, but very important, difference between [17] and the discrete formulation given by [7], associated with the difference between [16] and [8]. While in [8] the ratio between the income share and the population share (to the α power) is divided by n_j, in [16] this division is absent. The reason is that the division by n_j in [8] normalizes the summation; for example if all the ratios between income and population are one, the result of the summation of all the ratios is n_j. With the division by n_j, the summation term in [7] is normalized to one, and the inequality measure is zero. With the continuous formulation this normalization is not required, since we are using a density distribution function, so that if for all z we have $z=Y_j/n_j$, the result of the integral is one, and the inequality measure is zero[2].

The time derivative of expression [17] can now be taken allowing for changes in the population for each group:

[18] $$\dot{E}_a(j)=a\left(p_j-g_j\right)\left[E_a(j)-\frac{1}{a(1-a)}\right]-\frac{1}{a(1-a)}\int_0^{\infty}\left(\frac{zn_j}{Y_j}\right)^a\frac{\partial f_j(z)}{\partial t}dz$$

[2] A formally more accurate formulation of the idea of perfect equality would be to say that if all individuals have the same income, which must necessarily be Y_j/n_j (the arithmetic mean income), then the distribution function is a Dirac delta (impulse) function centered on $z=Y_j/n_j$.

Again we can see in [18] the decomposition into a meso (group level) component and a micro (individual level) component. The first term after the inequality sign, which represents changes aggregated at the group level (meso component), has the same structure of the two components shown in [14], with the inequality at level *j* (plus a constant) being multiplied by a linear combination of the changes in income and population of group *j*. In fact, this specific linear combination happens to be the symmetric of the rate of change of average income in group *j*. The second term in [18], with the integral, represents changes at the individual level. The time (partial) derivative of the density distribution function captures the way in which the shape of the distribution changes over time: for each income level, *z*, $\partial f_j(z)/\partial t$ gives the instantaneous addition or reduction in the density of individuals at the income level *z*.

We can now introduce both expressions [18] and [14] in the previously derived general decomposition of changes in inequality given by [10] to obtain a formal time-wise decomposition of a general entropy based measure of inequality, with the macro, meso and micro components presented in successive lines, respectively, from the top:

$$
\begin{aligned}
\dot{E}_a &= -\left[E_a - \frac{1}{a(1-a)}\right]\left[ag + (1-a)p\right] + \\
&+ \sum_{j=1}^{k} R_j \left[E_a(j) - \frac{1}{a(1-a)}\right] p_j - \\
&- \frac{1}{a(1-a)} \sum_{j=1}^{k} \left[\int_0^{\infty} \left(\frac{z}{Y}\right)^{a} \left(\frac{1}{n}\right)^{1-a} n_j \frac{\partial f_j(z)}{\partial t} dz\right]
\end{aligned} \qquad [19]
$$

With [19] we capture the relationship between changes in the inequality measure with changes at the aggregate, group and individual levels. While the macro component is left unchanged from [14], the meso components of [14] and [18] interact in a way that leaves the final meso component in [19] only as a function of changes in population of group *j*.

This section accomplished the time-wise decomposition for all general entropy based measures of inequality with the exception of the Theil measures. The decomposition shows that it is possible to separate the changes over time in the inequality measure into changes at different levels of aggregation. In the next section we proceed along the same lines for Theil measures.

3- TIME-WISE DECOMPOSITION OF THEIL MEASURES

The general entropy measures become the Theil inequality measures, defined by [2] (Theil's T) and [3] (Theil's L) when a is, respectively, one and zero. The Theil measures are important because they are the only inequality entropy measures for which the summation of the weights R_j defined in [5] equals one. Consequently, it is only for the Theil measures that the within-groups inequality component of the decomposition defined by [6] is a weighted average of the inequality of all groups. Shorrocks (1980) noted that Theil (1967) had seen in this special feature of Theil measures the property of independence between the weights and the between group component of inequality. The Theil measures are, perhaps for these reason –which facilitates the interpretation given to the within-group/between-group decomposition of total inequality – the most commonly used entropy measures of inequality. This section performs essentially the same type of time-wise decompositions of section 2 but only for Theil measures.

The decomposition of the Theil measures [2] and [3] is given below, with [5] representing the decomposition of Theil's T and [6] of Theil's L. The income shares are given by $w_j = \dfrac{Y_j}{Y}$ and the population shares by $e_j = \dfrac{n_j}{n}$.

$$[20] \qquad \begin{cases} T = \sum_{j=1}^{k} w_j \log \dfrac{w_j}{e_j} + \sum_{j=1}^{k} w_j T_j \\ T_j = \sum_{i=1}^{n_j} \dfrac{y_i}{Y_j} \log\left[\left(\dfrac{y_i}{Y_j}\right) \Big/ \left(\dfrac{1}{n_j}\right)\right] \end{cases}$$

[21] $$\begin{cases} L = \sum_{j=1}^{k} e_j \log \dfrac{e_j}{w_j} + \sum_{j=1}^{k} e_j L_j \\ L_j = \sum_{i=1}^{n_j} \dfrac{1}{n_j} \log\left[\left(\dfrac{1}{n_j}\right) \Big/ \left(\dfrac{y_i}{Y_j}\right)\right] \end{cases}$$

The comparison between the two decompositions highlights the well-known fact that the role of the income (population) shares in T is played in L by the population (income) shares.

The time derivative of Theil's T is given by:

[22] $$\dot{T} = \sum_{j=1}^{k} \left[\dot{w}_j \left(\log \frac{w_j}{e_j} + T_j \right) + w_j \frac{d}{dt} \log \frac{w_j}{e_j} \right] + \sum_{j=1}^{k} w_j \dot{T}_j$$

and of Theil's L by:

[23] $$\dot{L} = \sum_{j=1}^{k} \left[\dot{e}_j \left(\log \frac{e_j}{w_j} + T_j \right) + e_j \frac{d}{dt} \log \frac{e_j}{w_j} \right] + \sum_{j=1}^{k} e_j \dot{L}_j$$

Expressions [22] and [23] are similar to [10] in that they decompose changes in inequality into the changes due to the income and population effects at the group level (the first term to the right of the equality sign in each formula) and the changes of inequality within each group. The formulas are the same, with the role of the income shares and the population shares reversed when one goes from one to the other.

If we take only the time change associated with the groups, then we obtain for T:

$$\begin{aligned}&\sum_{j=1}^{k}\left[\dot{w}_j\left(\log\frac{w_j}{e_j}+T_j\right)+w_j\frac{d}{dt}\log\frac{w_j}{e_j}\right]=\\ \text{[24]}\qquad &=\sum_{j=1}^{k}w_j\left[g_j\left(\log\frac{w_j}{e_j}+T_j\right)-p_j\right]-\\ &-\left(gT-p\right)\end{aligned}$$

and for L:

$$\begin{aligned}&\sum_{j=1}^{k}\left[\dot{e}_j\left(\log\frac{e_j}{w_j}+L_j\right)+e_j\frac{d}{dt}\log\frac{e_j}{w_j}\right]=\\ \text{[25]}\qquad &=\sum_{j=1}^{k}e_j\left[p_j\left(\log\frac{e_j}{w_j}+L_j\right)-g_j\right]-\\ &-\left(pL-g\right)\end{aligned}$$

Again, as in [14], we see a clear decomposition of the changes into a meso and macro components. Besides reversing the role of the shares, we need now to reverse also the role of changes in population and changes in income. Therefore, while for T the macro component is given by multiplying T by the rate of change of income and subtracting the rate of change in population, for L the macro component is given by multiplying L by the rate of change of population and subtracting the rate of change of income. As in [14], we see that there is a formal similarity in the structure of the expressions for the meso and the macro components. For both T and L, the meso component is constituted by a summation of the components that would yield the macro component, were it not for the weights in the summation (which are the same weights used in the total within groups inequality).

To deal with the change of inequality within each of the groups we must, as in section 2, first begin by expressing the within group inequality in a continuous formulation. Following the same procedure as with the general entropy based measures of inequality, we obtain for the T measure:

[26] $$T_j = \int_0^\infty \frac{zn_j}{Y_j} \log \frac{zn_j}{Y_j} f_j(z) dz$$

and for the L measure:

[27] $$L_j = -\int_0^\infty \log \frac{zn_j}{Y_j} f_j(z) dz$$

Taking the time derivative of each measure:

[28] $$\dot{T}_j = -(g_j - p_j)(T_j + 1) + \int_0^\infty \frac{zn_j}{Y_j} \log\left(\frac{zn_j}{Y_j}\right) \frac{\partial f_j(z)}{\partial t} dz$$

and

[29] $$\dot{L}_j = (g_j - p_j) - \int_0^\infty \log\left(\frac{zn_j}{Y_j}\right) \frac{\partial f_j(z)}{\partial t} dz$$

The symmetry between the T measure and the L measure is not apparent in the comparison between [28] and [29]. The reason for this lack of symmetry is associated with the fact that the weights for the L measure are the population shares, the changes of which are embedded in the partial time derivative of the density distribution function. In other words, we cannot go as far in deriving an explicit expression for the time-wise decomposition of the L measure as we can for the T measure.

Still, as in [18], the rate of change in [28] and [29] is decomposed into a meso and micro components, with the micro component having the same structure (which is basically the integral of the time derivative of the density function weighted by the same weights that weigh the density function in the measures). The meso component is, again, a function of the change in average income in within the group. For the T measure, this change in average income is multiplied by one plus T_j, while for the L measure there is no further dependency on L_j.

Substituting [28] and [29], respectively, in [22] and [23], and using also the explicit expressions [24] and [25], results in the explicit time-wise decomposition for the T measure:

[30]
$$\begin{aligned}\dot{T} &= -g(1+T) + p + \\ &+ \sum_{j=1}^{k} w_j \left[g_j \log \frac{w_j}{e_j} + p_j T_j \right] + \\ &+ \sum_{j=1}^{k} \left[\int_0^\infty \frac{z}{Y} \log\left(\frac{zn_j}{Y} \right) n_j \frac{\partial f_j(z)}{\partial t} dz \right]\end{aligned}$$

and for the L measure:

[31]
$$\begin{aligned}\dot{L} &= -p(1+L) + g + \\ &+ \sum_{j=1}^{k} e_j p_j \left[\log \frac{w_j}{e_j} + L_j \right] - \\ &- \sum_{j=1}^{k} \left[\int_0^\infty \frac{1}{n} \log\left(\frac{zn_j}{Y} \right) n_j \frac{\partial f_j(z)}{\partial t} dz \right]\end{aligned}$$

The time-wise decomposition of changes in Theil measures partitioned changes in inequality into changes at different levels of aggregation. In particular, we are able to isolate changes associated with macro, meso and micro behavior, in the same fashion we accomplished for the general entropy based measures in section 2.

It is clear from [30] and [31] that further simplifications of the time-wise decomposition depend on providing and explicit functional form for the density distribution function or an explicit functional form for the changes in the distribution function over time, which is a weaker condition than defining explicitly the functional form for *f*. Making explicit the functional forms for *f* or for the partial time derivative of *f* corresponds, in the first case, to a definition of the way in which income is distributed in a static context and, in the second case, to a specification of a rule that determines how income is distributed from one instant to the next. In either case, we are entering into the definition of micro behavior. In section 4 we explore the impact of defining possible ways to describe micro behavior on the expressions for the time change of the Theil measures.

4- INCOME DISTRIBUTION RULES AND THE DYNAMICS OF THEIL MEASURES

In this section we focus on one application of the explicit formulas derived above for the Theil measures but with no partition of the population into groups. In this context, the expressions for the Theil measures given by [26] and [27] turn, for the T measure, to:

[32] $$T = \int_0^\infty \frac{zn}{Y} \log \frac{zn}{Y} f(z) dz$$

and for the L measure:

[33] $$L = -\int_0^\infty \log \frac{zn}{Y} f(z) dz$$

The time derivatives are, for the T measure:

$$\dot{T} = -\dot{m}(T+1) + \int_0^\infty \frac{zn}{Y} \log\left(\frac{zn}{Y}\right) \frac{\partial f(z)}{\partial t} dz \qquad [34]$$

and for the L measure:

$$\dot{L} = \dot{m} - \int_0^\infty \log\left(\frac{zn}{Y}\right) \frac{\partial f(z)}{\partial t} dz \qquad [35]$$

where $\dot{m}$ is the change in mean income, which is equal to ($g - p$).

Theil (1967) showed that by assuming an explicit expression for the density distribution function the Theil measures could often be expressed only as a function of the parameters that define the assumed distribution. Specifically, Theil experimented with two distributions functions that are commonly used to describe the distribution of income: the log-normal and the Pareto distribution functions. For the log-normal, Theil showed that the T measure is reduced to half the variance of the logarithm of income, and for the Pareto distribution the T measure is given by an expression that is a function only of the Pareto constant.

Our approach in this section differs from Theil's in that instead of assuming that the income is distributed according to some specific functional form, we define a rule for the dynamic changes in the distribution function. Therefore, our approach is more general, because the results continue to be valid with an arbitrary distribution function. After defining the dynamic distribution rule (we will use only distribution rule from now on, but it should remain clear that it is a dynamic rule) it is possible to express, in some cases, the changes in Theil measures exclusively with macro variables.

To specify an income distribution rule in [34] and [35] is equivalent to define an explicit functional form for $\frac{\partial f(z)}{\partial t}$. This time derivative represents, as mentioned above, the change in the density of individuals for the level of income *z*. A simple income distribution rule can be defined as[3]:

[36] $$\frac{\partial f(z)}{\partial t} = b(z-m)f(z)$$

where β is any real constant. The rule defined by [36] describes a dynamic behavior in which the density change for a certain level of income *z* depends on the deviation of that income level *z* from the mean of overall income. The constant of proportionality is the factor *b*, the sign of which determines if the distribution rule is "progressive" or "regressive". If *b* is negative, then the distribution rule is progressive: there is an increase in the density of individuals with incomes below the mean, and a decrease in the density of individuals with incomes above the mean. If *b* is positive the opposite occurs, and the distribution rule is regressive.

Including the distribution rule [36] into expression [35], which defines the rate of change of the L measure, produces a particularly simple formula, where the changes in the L measure are given as a function only of macro variables[4]:

[3] The explicit dependency on time was omitted to simplify the notation, but it is important to bear in mind that, with the exception of *z*, which plays the role of a parameter, and β, by definition a constant, everything in [36] is time-dependent, including the mean and the density distribution function.

[4] With distribution rule [36] it is not possible to reduce the change in the T measure to an expression that is only a function of macro variables; therefore, we discuss only the changes in the L measure associated with the income rule [36].

[37] $$\dot{L}(t) = \dot{m}(t) - bm(t)[T(t) + L(t)]$$

It is important to note that with the distribution rule [36] the change in mean income is defined endogenously, since:

$$\dot{m} = \frac{d}{dt}\int_0^\infty zf(z)dz = \int_0^\infty z\frac{\partial f(z)}{\partial t}dz = b\int_0^\infty z(z-m)f(z)dz = b\left[\int_0^\infty z^2 f(z)dz - m\int_0^\infty zf(z)dz\right]$$

it follows that:

[38] $$\dot{m}(t) = bs^2(t)$$

where σ^2 represents the variance of the distribution *f*. Therefore, a progressive income distribution rule (β negative) is accompanied by a negative growth of average income, and a regressive distribution rule by a positive growth of average income. The income distribution rule [36] is a possible way to formalize the trade-off between inequality and growth, which is part of some theoretical and empirical work on the relationship between economic growth and economic inequality. However, it is not a very flexible income distribution rule, precisely because the growth of average income is defined endogenously; the rule does not allow for the possibility of income growth being determined by other exogenous factors, beyond the distribution rule itself.

Before considering a more flexible distribution rule, it is worthwhile noting that it is possible to express the change in the L measure associated with the income distribution rule [36] exclusively with entropy based measures of inequality and the mean income. Since E_2 is proportional to the coefficient of variation, it is easy to show that with the income distribution rule [36]:

[39] $$\dot{L} = bm\left[2mE_2 - (T + L)\right] \Leftrightarrow$$
$$\Leftrightarrow \dot{E}_0 = bm\left[2mE_2 - (E_1 + E_0)\right]$$

We consider now a new distribution rule defined by:

[40] $$\frac{\partial f(z)}{\partial t} = b\,\frac{(z-m)}{z}\frac{\dot{m}}{m}\,f(z) + \frac{\dot{m}}{m}\,f(z)$$

Income distribution rule [40] is structurally similar to rule [36] and can be interpreted in the same way: the gain, or loss, in population density at income level z depends on the constant β and also on the "distance" of this level of income z from the mean income. However, there are subtle, but important, differences. First, we consider now a relative "distance", since the difference between z and the mean income is divided by z. Secondly, the relative distance is multiplied by the growth rate of mean income. Finally, we add one more term with the growth rate of mean income multiplied by the density function. This means that all levels of income gain or loose density proportionally to the existing density level and the rate of growth of mean income, with the proportionality being defined by the constant β and the relative distance from the mean.

It is easy to see that the income distribution rule [40] does not determine the growth of mean income endogenously. In fact, for the class of income distribution rules defined by:

[41] $$\frac{\partial f(z)}{\partial t} = \left[\frac{(z-m)}{z}g(\bullet) + 1\right]\frac{\dot{m}}{m}\,f(z)$$

where $g(\bullet)$ is any function that does not depend on *z*, the growth rate of income does not depend directly from the distribution rule. The reason is that:

$$\dot{m} = \int_0^\infty z \frac{\partial f(z)}{\partial t} dz = \int_0^\infty [(z-m)g(\bullet)+1] \frac{\dot{m}}{m} f(z)dz = g(\bullet)\frac{\dot{m}}{m}\int_0^\infty (z-m)f(z)dz + \dot{m} = \dot{m}$$

since $\int_0^\infty (z-m)f(z)dz = \int_0^\infty zf(z)dz - m = 0$. In the case of the specific distribution rule [40] we are considering the special situation where $g(\bullet) = b$. Including the explicit distribution rule [40] into the expression for the change in the T measure [34] leads to[5]:

[42] $$\frac{\dot{T}}{T} = \frac{\dot{m}}{m}\left[1 + b\left(1 + \frac{L}{T}\right) - m\left(1 + \frac{1}{T}\right)\right]$$

Expression [42] provides an explicit relationship between the growth rate of the T measure and the growth rate of mean income. The relationship is complex, despite the fact that the income distribution rule defined in [40] was relatively simple. In fact, there is almost a trade-off between the simplicity of the income distribution rule and the simplicity of the relationship between the rates of growth of mean income and of the T measure. One way to express formally this trade-off is to make explicit the rate of change of the T measure, given the income distribution rule in the more general form [41]:

[43] $$\frac{\dot{T}}{T} = \frac{\dot{m}}{m}\left[1 + g(\bullet)\left(1 + \frac{L}{T}\right) - m\left(1 + \frac{1}{T}\right)\right]$$

---

[5] With distribution rule [40], it is the T measure that allows for a full expression only as a function of macro variables; with this distribution rule, that is not possible for the L measure.

We can now “choose” the function $g(\bullet)$ so that we obtain a simple expression when we make explicit this functional form in [43]. This may entail that the income distribution rule becomes complex and hard to interpret. For example, if we make $g(\bullet)=\frac{T(b-1)}{T+L}$ then the income distribution rule is rather complex, but [43] turns to the simple form:

[44] $$\frac{\dot{T}}{T}=\frac{\dot{m}}{m}\left[b-m\left(1+\frac{1}{T}\right)\right]$$

With this specific transformation we eliminate the dependency on the L measure. Therefore, it is easier to see the dynamic impact of the distribution rule [41] on the relationship between income growth and inequality change, which depends now only on β and on the growth rate of mean income (which must be provided exogenously). Depending on the value of β, and on an initial level of income and inequality, the growth rate of T can be either positive or negative. If, for some initial conditions, the rate of growth of T is positive (meaning that T is increasing) and if we assume also a positive rate of growth of mean income, after a certain amount of time the growth rate of T will become negative (since μ is growing and 1/T is also growing, because T is decreasing). From then on, growth in mean income will be accompanied by decreasing inequality.

The evolution of the relationship between income growth and the rate of change of inequality described in the last paragraph corresponds to the relationship between income and inequality of the Kuznets hypothesis. Consider, as an illustration, a situation where both β and the growth rate of mean income are fixed and defined exogenously as in the note to Figure 3. The initial levels of mean income and inequality are, respectively, one and 0.1. In these conditions, the income growth is inequality augmenting in a first stage, but a threshold level of mean income is reached after which income growth is inequality reducing. This behavior is consistent with the Kuznets hypothesis, as illustrated graphically in Figure 3.

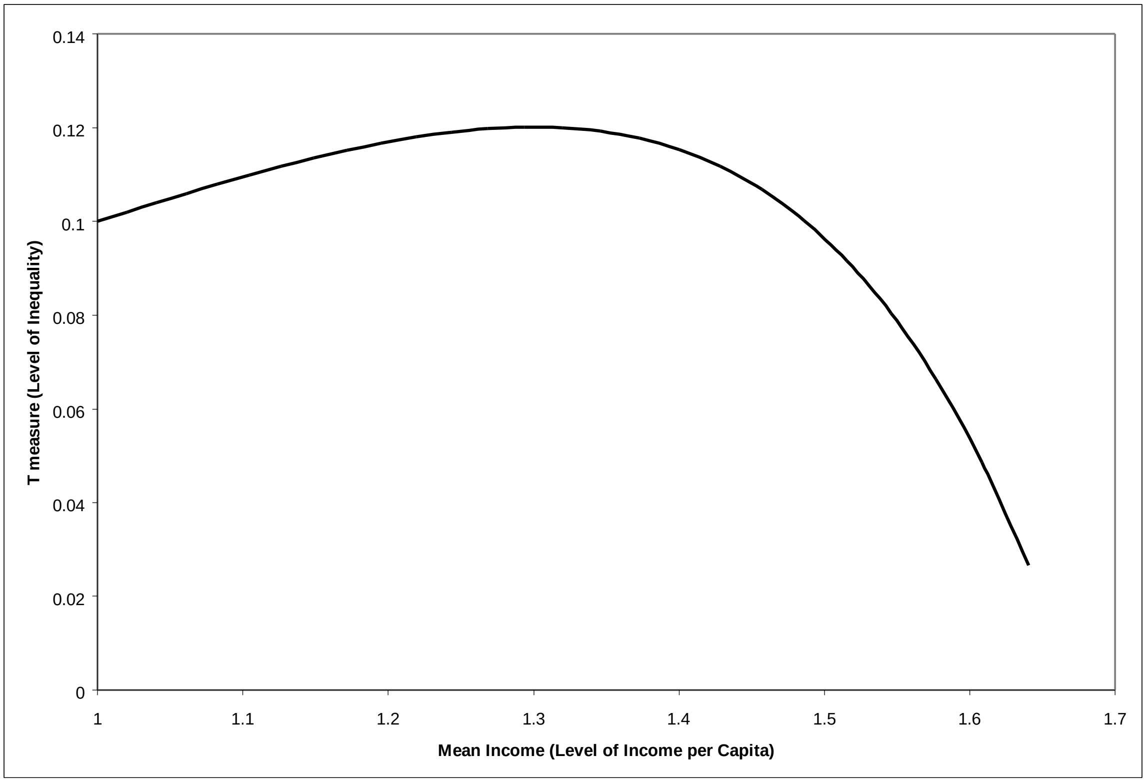

Figure 3- The Kuznets Hypothesis

Simulation of the evolution of the relationship between income per capita and inequality predicted by [44], considering $\frac{\dot{m}}{m} = 2\%$ and b = 12.

We used [44] to illustrate how we can choose a micro dynamic income distribution rule that produces macro behavior consistent with one theory – in the case of Figure 3, with the Kuznets hypothesis. Another application for [44] consists of determining empirically values for b for a given context, from known values of growth and changes and levels of inequality. More generally, we could try to determine the micro distribution rule of the form [41] that is consistent with a certain macro behavior (in other words, instead of looking for a parameter, defining the function g).

In conclusion, we provide in this section a template to map dynamic micro behavior into macro manifestations of the relationship between income growth and inequality changes. Further, we showed that by specifying a rather general dynamic income distribution rule the expression for the time changes of Theil measures can be expressed, in certain cases, exclusively as a function of macro variables.

5- DYNAMICS OF THE BETWEEN-GROUP THEIL COMPONENT AT DIFFERENT LEVELS OF AGGREGATION

In the dynamic analysis of this section, we will be looking only at inequality across groups, that is, we will not analyze dynamics of inequality across individuals. However, there will be an added level of complexity associated with the fact that we will consider a sequence of nested and hierarchic grouping structures. A static analysis of the behavior of the Theil index in sequences of grouping structures was presented in a previous paper (Conceição, Galbraith and Bradford, 2000), and we extend, with this section, the results of that paper adding the dynamic dimension. We will consider only fixed grouping structures, so that the number and relationship of the different levels of aggregation do not change over time.

With a fixed grouping structure the sources of between-group inequality variation over time are associated with income and population effects. Each effect is reflected in the way the Theil index responds to changes in the income and population shares. To analyze the Theil index response to income and population changes over time with generality we take the time derivative of the between group component only of the Theil, that is, of the first summation of the equation on top in [20]:

[45] $$\dot{T}' = \sum_{j=1}^{m} \left(\frac{\partial T'}{\partial w_j} \dot{w}_j + \frac{\partial T'}{\partial e_j} \dot{e}_j \right)$$

The first term in the summation corresponds to the income effect, while the second to the population effect. The rates of change of the shares are given by:

[46]
$$\begin{cases} \dot{w}_j = w_j\left(g_j - g\right) \\ \dot{e}_j = e_j\left(p_j - p\right) \end{cases}$$

The intuition behind [46] is now immediate: the change in the shares of group j is proportional to the difference between the rate of change of income (population) in group j and the overall rate of income (population) change. If group's j income changes at the same rate as the change in overall income, then the share remains the same. If the growth rates differ, the change depends also on the level of the shares.

The dependency of the between group Theil index on income and population changes is given by:

[47]
$$\begin{cases} \dfrac{\partial T'}{\partial w_j} = \log\left(\dfrac{w_j}{e_j}\right) + 1 \\ \dfrac{\partial T'}{\partial e_j} = -\dfrac{w_j}{e_j} \end{cases}$$

which in conjunction with [46] gives:

[48]
$$\dot{T}' = \sum_{j=1}^{m} w_j \left\{ \left(g_j - g\right)\left[\log\left(\frac{w_j}{e_j}\right) + 1\right] - \left(p_j - p\right) \right\}$$

Expression [48] shows how the between group Theil index reacts to changes in income and population. Essentially, the change in the between group Theil index depends on the difference between the rates of change of the shares of income and population of each group with the overall rates of change of income and population. If a group's share of income and population change at the same rate as the change in overall income and population, then this group does not contribute to changes in the Theil index (between group will be omitted in the remaining of the section, but it should be understood that we are considering only the between group component).

If g_j is different from g or p_j is different from p, then group j contributes to changes in the Theil. The way in which changes in income and population in group j affect the Theil index depends on whether group j is a "poor" or a "rich" group (poor in the sense that the share of income is substantial lower than the share of population). The multiplicative factor $[\log(w_j/e_j)+1]$ determines whether (g_j-g) contributes positively or negatively to the change in Theil. If $[\log(w_j/e_j)+1]>0$ (meaning that we are dealing with a "rich" group) then if $g_j>g$ the income effect increases the Theil index, because a "rich" group gains income at a rate higher that the overall population, increasing inequality.

Since (p_j-p) is preceded by a minus sign, the population effect works in a symmetric way (if a "rich" group gains population at a rate higher than the overall population growth, then inequality decreases). When $[\log(w_j/e_j)+1]<0$ we are dealing with "poor" groups, and the effect of income and population changes is opposite of the one described above.

The effect of $[\log(w_j/e_j)+1]$ goes beyond determining the sign that precedes (g_j-g), since it also weighs more heavily changes at the lowest shares of income. Figure 1 shows the dependency of $[\log(w_j/e_j)+1]$ on w_j, assuming that $e_j=0.5$.

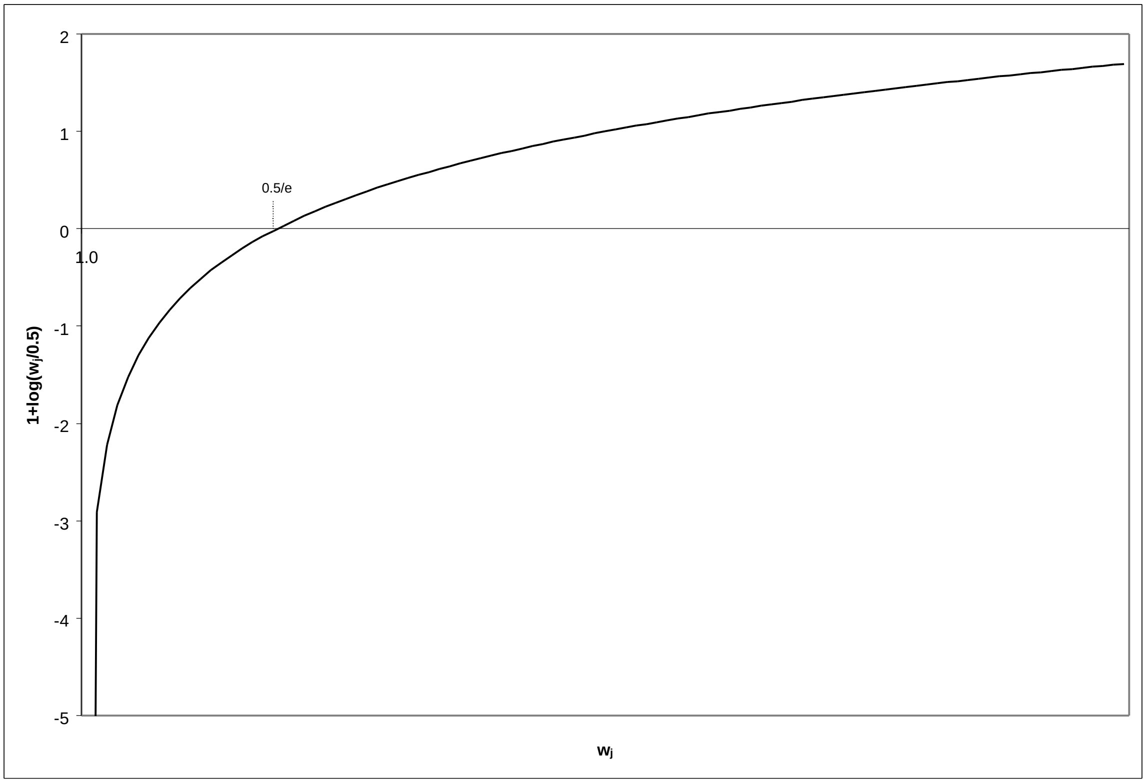

Figure 1- Multiplicative Factor of the Income Effect: Dependency on the Income Share

When the $w_j=e_j/e$ then $[\log(w_j/e_j)+1]=0$; wage shares lower than e_j/e lead to a negative $[\log(w_j/e_j)+1]$ and the reverse happens when the w_j is higher than e_j/e[6]. Additionally, the shape of the curve is such that when $[\log(w_j/e_j)+1]$ is negative (that is, when we are dealing with "poor" groups) the expression $[\log(w_j/e_j)+1]$ weighs more heavily on (g_j-g) than when $[\log(w_j/e_j)+1]$ is positive ("rich" groups).

What is the meaning of the "cut-off" point $w_j=e_j/e$ for which $[\log(w_j/e_j)+1]=0$? This is the point at which group *j*'s contribution to the Theil index attains its minimum, as illustrated in Figure 2. Conceição and Ferreira (2000) provide an intuitive interpretation of this "cut-off" point.

[6] Here "e" is Neper's number.

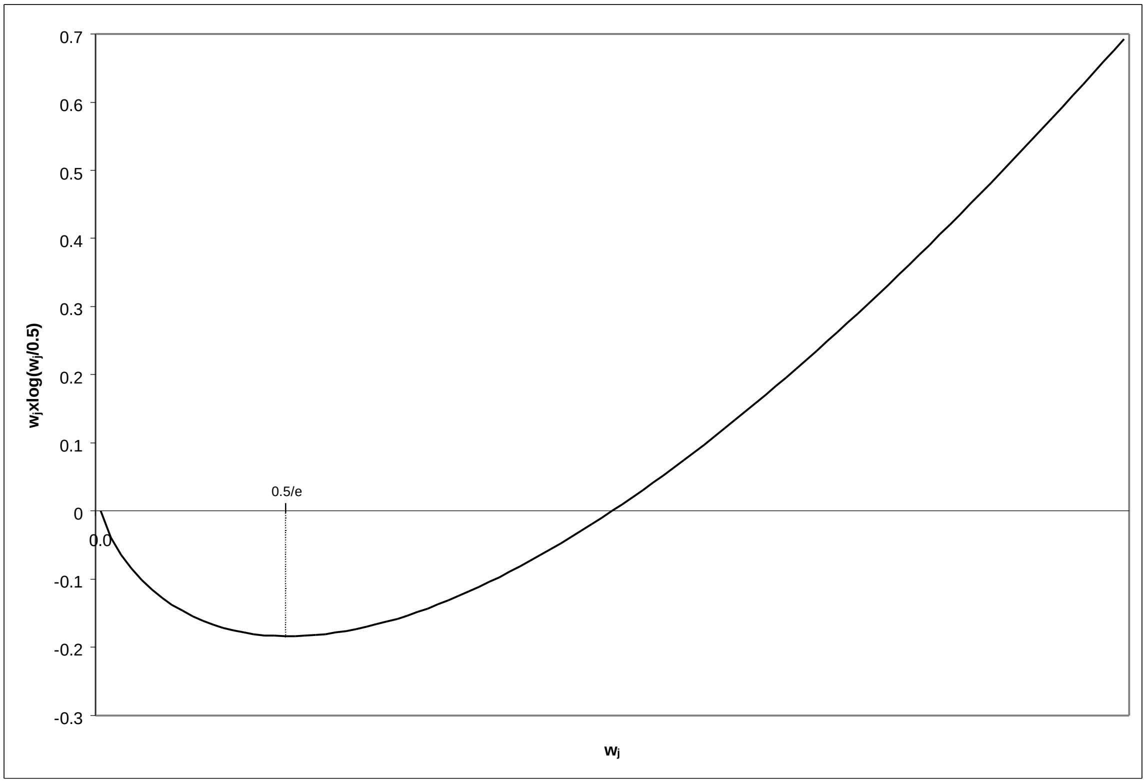

Figure 2- Group *j*'s Contribution to the Theil Index: Dependency on the Income Share

Yet another way to express the change over time of the Theil index is given by:

[49] $$\dot{T}' = p - gT' - \sum_{j=1}^{m} w_j \left[p_j - g_j \log\left(\frac{w_j}{e_j}\right)\right]$$

which shows, again, that changes in the between group component of the Theil can be separated into a "macro" component (*p-gT'*) and a "micro" component, which depends on the dynamics of the distribution of income and population across groups. This decomposition is explored in Conceição and Galbraith (forthcoming). The "micro" component is composed of two summations: $\sum_{j=1}^{m} w_j p_j - \sum_{j=1}^{m} w_j g_j \log(w_j/e_j)$. The structure of the micro component mirrors the structure of the macro component. The first summation adds the group's weighted rates of population change, where the weights are the income shares. The second summation is a modified Theil index, where the weights of $\log(w_j/e_j)$ are, instead of w_j, the changes in the wage shares (note that $w_j(t+1)=w_j(t)+g_j w_j(t)$; the weights are, then: $w_j(t+1)-w_j(t)$). So, in a way, the micro component is, as the macro component, a difference between the rates of population change and the rate of income change combined with the Theil index.

The general results on the between-group dynamics of the Theil index are important to formalize the relationship between the dynamics of the Theil index at different levels of aggregation. A result derived in Conceição, Galbraith and Bradford (2000) is that the relationship between the levels of the between-group Theil index at consecutive levels of aggregation can be expressed as:

[50] $$T'_{s+1} = T'_s + T_s$$

where s indexes a grouping structure in a sequence of nested and hierarchic grouping structures. For example, s could represent the regional aggregation of individual data at the state level; a lower level of aggregation is indexed by s+1, and this could correspond to having the data aggregated and the county level, continuing with the same example.

Therefore, expression [50] is important because it shows that the information gain associated with calculating the Theil index at a lower level of aggregation – moving down from the state to the county level, for example – is a single additive factor. Since the Theil index is always positive or zero, we can also see that at a lower level of aggregation the Theil index will always be equal or higher than that computed at a higher level of aggregation. Additionally, expression [50] shows that, unless the distribution of income across the groups at level s (state) within each s+1 group (county) is homogeneous, the level of the Theil index at the two levels can be quite different.

However, our interest in this paper is in the dynamics of inequality. From [50] it is obvious that the rate of change at the lower level of aggregation is also the simple summation of the change at a higher level plus an additional term. In other words, the informational gain on the dynamics of the Theil index associated with a less aggregated grouping structure is given by:

[51] $$\dot{T}'_{s+1} - \dot{T}'_s = \dot{T}_s$$

The sign and scale of the term $\dot{T}_s$ is now ambiguous, because the information gain associated with considering a lower level of aggregation can either increase or decrease the rate of change at the higher level. The explicit expression for $\dot{T}_s$ is given by:

[52] $$\dot{T}_s = \sum_{i_1=1}^{m} \cdots \sum_{i_s=1}^{m_{i_1 \ldots i_{s-1}}} \frac{Y_{i_1 \ldots i_s}}{Y} p_{i_1 \ldots i_s} - gT_s + \sum_{i_1=1}^{m} \cdots \sum_{i_{s+1}=1}^{m_{i_1 \ldots i_s}} \frac{Y_{i_1 \ldots i_{s+1}}}{Y} \left[p_{i_1 \ldots i_{s+1}} - g_{i_1 \ldots i_{s+1}} \log\left(\frac{w_{i_1 \ldots i_{s+1}}}{e_{i_1 \ldots i_{s+1}}} \right) \right]$$

where

[53] $$\begin{cases} w_{i_1 \ldots i_{s+1}} = Y_{i_1 \ldots i_{s+1}} / Y_{i_1 \ldots i_s} \\ e_{i_1 \ldots i_{s+1}} = n_{i_1 \ldots i_{s+1}} / n_{i_1 \ldots i_s} \\ g_{i_1 \ldots i_s} = \dot{Y}_{i_1 \ldots i_s} / Y_{i_1 \ldots i_s} \\ p_{i_1 \ldots i_s} = \dot{n}_{i_1 \ldots i_s} / n_{i_1 \ldots i_s} \end{cases}$$

Expression [52] shares the general structure of the rate of change of the between-group Theil shown in [49], but is slightly more complex, given that T_s is not a "pure" between-group Theil. In the context of [52], micro now means at level s+1. Therefore, the first two terms to the right of the equality sign in [52] still reflect macro behavior. The first term is a weighted summation of the population shares at level s, with the weights being the income shares of all the groups in s. The second term is the overall income growth rate "corrected" by T_s, following the structure of the macro term in [49]. The micro term is the weighted summation of the differences between the growth in population shares and corrected wage shares at level s+1. The weights are the income shares of the groups at level s. The correction in the change rate of the wage shares is given by logarithm of the ratio between the shares of each group s+1 in each s and the corresponding population shares.

The information gain on the dynamics of the Theil index when one moves to a less aggregated grouping structure is zero whenever $\dot{T}_s$ is zero. From [52] it can be shown that if the following expression occurs for every group in s then $\dot{T}_s = 0$:

$$[54] \qquad g^{i_1 \ldots i_s} T^w - p_{i_1 \ldots i_s} = \sum_{i_{s+1}=1}^{m_{i_1 \ldots i_s}} w_{i_1 \ldots i_{s+1}} \left[g_{i_1 \ldots i_{s+1}} \log\left(\frac{w_{i_1 \ldots i_{s+1}}}{e_{i_1 \ldots i_{s+1}}} \right) - p_{i_1 \ldots i_{s+1}} \right]$$

Expression [54], once again, can be interpreted as equality between macro behavior (in this case, at level s) and micro dynamics (at level s+1). Formally, [54] is structurally similar to [49] (if we make $\dot{T}' = 0$ in [49]), and this interpretation follows from the discussion after [49].

While [54] is a sufficient condition for $\dot{T}_s = 0$, it is not a necessary condition. In fact, even if [54] is not verified for every group in s, $\dot{T}_s$ can be zero due to the interaction of negative and positive rates of change in [52]. Therefore, the ability to move further with general analytic results is limited by the possibility of complex interactions across the rates of growth of the shares of population and income of different groups at different levels of aggregation. Still, we were able to establish and understand the general structure of these interactions and to derive a formal relationship between the dynamics of the between group Theil index at different levels of aggregation.

A practical application of these results is associated with the decision over which level of aggregation is low enough to have a good representation of the inter-individual inequality from the between group Theil index. As suggested in Conceição and Galbraith (2000), the usage of the between industry Theil index can be used to construct long and dense time-series of inequality. But at which level of industrial classification should we stop to capture the dynamics of inequality? Do we need to consider a five-digit SIC code? Or is a four-digit grouping structure good enough? Empirical results, shown as illustrative, in Conceição, Galbraith and Bradford (2000) indicate that even a two-digit level could provide a good indication of the dynamics of a five-digit level of aggregation. The results derived in this section can be used to formally account for the information gain, or loss, associated with problems such as these.

6- CONCLUSIONS

This paper explores the dynamic property of general entropy-based measures of inequality, with a special focus on Theil measures. We derive explicit expressions for the time change of general entropy inequality measures and for Theil measures. In both instances, we show that it is possible to achieve a time-wise decomposition of changes in inequality whereby the impact of changes of macro variables and micro behavior are separated. This time-wise decomposition is parallel, in concept, to the static group-wise decomposition that is exclusive of entropy based inequality measures. To achieve a total decomposition of changes of inequality accounting both for income and population changes, we transform the definition of the distribution of income from a discrete formulation into a continuous formulation.

We then provide two applications of the generic results derived. First we show that by providing dynamic income distribution rules – rules that define the way in which the distribution of income changes over time at the micro level – some of the expressions for the time changes of the Theil measures can be simplified further, and be expressed exclusively as a function of macro variables. The specification of dynamic distribution rules is a weaker (more general) condition than specifying an explicit income distribution function. We illustrate the possibility of deriving expressions for dynamic micro behavior consistent with macro manifestations of the relationship between income growth and time changes in inequality, using the explicit example of the Kuznets hypothesis.

A second application is associated with a formal expression linking the dynamics of the Theil index at different levels of aggregation in a sequence of nested and hierarchic grouping structures. A practical application of these results is the quantification of the information loss that results from computing the Theil index at a higher level of aggregation.

REFERENCES

Theil, H. (1967). *Economics and Information Theory*. Chicago: Rand McNally and Company.

Bourguignon, F (1979). "Decomposable Inequality Measures", *Economotrica*, vol. 47.

Sen, A. K. (1973). *On Economic Inequality*, Oxford: Claredon Press.

Conceição, P., Galbraith, J. K. (2000), "Constructing Long and Dense Time-Series of Inequality Using the Theil Index", *Eastern Economic Journal*, 26(1): 61-74.

Conceição, P. P. Ferreira (2000), *The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications*, University of Texas Inequality Project Working Paper No. 14; available on the Internet at: http://utip.gov.utexas.edu.

Conceição, P., J. K. Galbraith, P. Bradford (2000). *The Theil Index in Sequences of Nested and Hierarchic Grouping Structures*, University of Texas Inequality Project Working Paper No. 15; available on the Internet at: http://utip.gov.utexas.edu.

Shorrocks, A. F. (1980). "The Class of Additively Decomposable Inequality Measures", *Econometrica*, vol. 48.

Shorrocks, A. F. (1984). "Inequality Decomposition by Population Subgroups", *Econometrica*, vol. 52.