Ca' Foscari University of Venice

Department of Economics

Working Paper

Dino Rizzi

Tax Evasion Indices and Profiles

ISSN: 1827/3580
No. 37/WP/2012

Working Papers
Department of Economics
Ca’ Foscari University of Venice
No. 37/WP/2012
ISSN 1827-3580

Tax Evasion Indices and Profiles

Dino Rizzi
Department of Economics
Ca’ Foscari University of Venice

December 2012

Abstract
The aim of the paper is to apply definitions and graphical devices currently used in the economic literature on poverty to individual data on tax evasion. Starting from simple indices, the paper presents composite indices and profiles of tax evasion and compliance, based on the *three I’s* of tax evasion: incidence, intensity and inequality. In the field of tax evasion, a stream of literature produces potentially a large amount of individual micro data using agent-based models: the aim of the paper is to enrich the analysis offered by these models with indices that take into account the whole distribution of taxpayers’ evasion rates, rather than the usual average rate.

Keywords: tax evasion, indices, incidence, intensity, inequality

JEL Codes: H26

Address for correspondence:
Dino Rizzi
Department of Economics
Ca’ Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (++39) 041 2349167
Fax: (++39) 041 2349176
E-mail: dino.rizzi@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca’ Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

The Working Paper Series
is available only on line
(www.dse.unive.it/pubblicazioni)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca’ Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

1 Introduction

Figures about tax evasion are usually represented by very simple indices, namely the percentage of tax evaders, the share of unreported income, and the share of unpaid taxes. However, it is possible to borrow from the economic literature about poverty some definitions and graphical devices that can explain in a clear way some other features of tax evasion.

Since Sen (1976), the literature about poverty measurement derived many axioms, measures and graphical devices that can be widely used in analyses entailing distributional issues. An important stream of literature focused on poverty indices expressed in terms of normalized poverty gaps. In particular, it is worth recalling the Three I's of Poverty (TIP) curve defined by the components of Sen index: *incidence, intensity* and *inequality*. Following this approach, Shorrocks (1995) and Jenkins and Lambert (1997) have defined a curve considering these three I's, called the *poverty gap profile* or *TIP curve*. More recently, Zengh (2000), Xu and Osberg (2001) and Xu (2003) derived from the poverty gap profile a synthesis of the poverty indices proposed by Sen (1976), Thon (1979) and Shorrocks (1995): i.e. the so called Sen-Shorrocks-Thon poverty index.

In the field of tax evasion, if individual data are available, for instance from simulation models, it is straightforward to see some analogies with the measurement of poverty. In the same way as the *i*-th individual of a distribution is defined "poor" by comparing income y_i with a poverty line z (if $y_i < z$), an individual is regarded as "tax evader" if declaring a taxable income x_i lower than the true income m_i. The "poverty gap" is defined as the difference $g_i = z - y_i$; the "unreported income" can be seen as another type of "gap": $u_i = m_i - x_i$.

The analysis of tax evasion is anyway more complex, as for poverty the reference level is unique for all individuals, namely the poverty line z, while in the case of tax evasion the unreported income is computed by using the true income m_i that differs among individuals.

Therefore, while poverty gaps and incomes of the poor are ordered exactly in the opposite way (i.e. if $y_1 < y_2 < y_3 < ...y_q$, then $g_1 > g_2 > g_3 > ... > g_q$), in the case of tax evasion it is not possible to link the ranking of unreported incomes with the ranking of true incomes, as the tax evader can be either a poor or a rich person.

In the field of tax evasion individual micro data are rarely available. However, there is a stream of literature that produces potentially a large amount of individual data generated by agent-based models[1]. The aim of the paper is to enrich the results of these models by applying to tax evasion some definitions and graphical devices currently used in the economic literature on poverty.

Section 2 presents some simple indices of tax evasion and tax compliance. Borrowing the TIP curve concept, in Section 3 and 4 a tax evasion profile and a tax compliance profile are derived. Section 5 and 6 are devoted to building composite indices of tax evasion and tax compliance similar to the Sen-Shorrocks-Thon poverty index. An empirical exercise is presented in Section 7, based on individual data produced by a simple agent-based model of tax evasion.

2 Definitions and simple indices of tax evasion and compliance

There are many possible definitions of tax evasion and tax compliance, as we can consider the number of tax evaders, the amount of reported/unreported incomes and the amount of paid/unpaid taxes.

Consider a society with N individuals i=1, ..., N, with a distribution of true incomes $m = \{m_1, m_2, ..., m_N\}$, a distribution of reported incomes $x = \{x_1, x_2, ..., x_N\}$ and an associated distribution of unreported incomes $u = \{u_1, u_2, ..., u_N\}$. For each individual i let's assume $m_i > 0$, $0 \le x_i \le m_i$, so the unreported income is:

[1] $$u_i = m_i - x_i$$

and $0 \le u_i \le m_i$. The individual rate of evasion, i.e. the share of true income not declared, is:

[2] $$e_i = \frac{u_i}{m_i}$$

that is, by definition, $0 \le e_i \le 1$. The compliance rate can be defined as:

[3] $$d_i = 1 - e_i = \frac{x_i}{m_i}$$

The binary variable η_i can be used to identify tax evaders:

[4] $$\eta_i = \begin{cases} 0 & if \quad e_i = 0 \\ 1 & if \quad e_i > 0 \end{cases}$$

[1] The use of agent-based model can be traced back to Mittone and Patelli (2000). More recent contributions are to Hokamp and Pickhardt (2010), Davis et al. (2003) Korobow et al. (2007) Bloomquist (2006) Bloomquist (2011) Zaklan et al. (2009), Pickhardt and Seibold (2011).

in which the value "1" is associated to a "tax evader" and "0" is associated to a "full complier". The number of tax evaders, N_E, and the number of full compliers, N_D, are then:

[5] $$N_E = \sum_{i=1}^{N} \eta_i \qquad \text{and} \qquad N_D = N - N_E$$

With these elements we can define:

- the total amount of incomes (M) and the amount of tax evaders' incomes (M_E):

[6] $$M = \sum_{i=1}^{N} m_i \qquad \text{and} \qquad M_E = \sum_{i=1}^{N} \eta_i m_i = \sum_{i=1}^{N_E} m_i$$

- the amount of incomes declared by all individuals (X) and by tax evaders (X_E):

[7] $$X = \sum_{i=1}^{N} x_i \qquad \text{and} \qquad X_E = \sum_{i=1}^{N} \eta_i x_i = \sum_{i=1}^{N_E} x_i$$

- the amount of unreported income:

[8] $$U = \sum_{i=1}^{N} u_i = \sum_{i=1}^{N} (m_i - x_i) = M - X = M_E - X_E$$

With the above definitions, it is possible to compute some simple indices of tax evasion, as shown in Table 1. The first two indices can be related to a concept of *incidence*:

- the *incidence of evaders, H*, i.e. the share of tax evaders in the population (the *head count ratio* of tax evaders);
- the *incidence of evaders' reported incomes*, H_X, i.e. the share of income reported by tax evaders on the total amount of true income.

The other two indices in Table 1 are average rates of evasion and can be associated to a concept of intensity:

- the *intensity of tax evasion*, $\bar{e}$, i.e. the average rate of evasion of all individuals;
- the *intensity of tax evasion of evaders*, $\bar{e}_E$, defined as the average rate of evasion among tax evaders.

In the same way, Table 2 shows five simple *tax compliance* indices, which can be derived as the complement to 1 of the corresponding tax evasion indices.

The indices of tax evasion and compliance are computed by using the individual true income shares, $f_i^m = m_i / M$, and the cumulative true income shares, $F_i^m = \sum_{j=1}^{i} m_j / M = \sum_{j=1}^{i} f_j^m$.

Table 1 – Incidence and intensity of tax evasion indices

incidence of evaders (*head count ratio* of tax evaders):	$H = \frac{N_E}{N} = \sum_{j=1}^{N} \eta_i$
incidence of evaders' true income (share of true income of tax evaders*)*	$H_M = \frac{M_E}{M} = \sum_{i=1}^{N_E} \frac{m_i}{M} = \sum_{i=1}^{N_E} f_i^m$
intensity of tax evasion (average rate of evasion)	$\bar{e} = \frac{U}{M} = \sum_{i=1}^{N_E} \frac{u_i}{m_i} \frac{m_i}{M} = \sum_{j=1}^{N_E} e_j f_j^m$
intensity of tax evasion of evaders (average rate of evasion of tax evaders)	$\bar{e}_E = \frac{U}{M_E} = \frac{U}{M} \frac{M}{M_E} = \frac{1}{H_M} \sum_{i=1}^{N_E} e_i f_i^m = \frac{\bar{e}}{H_M}$

Table 2 – Incidence and intensity of tax compliance indices

incidence of compliers (*head count ratio* of full compliers):	$H^d = \frac{N_D}{N} = 1 - H$
incidence of compliers' true income (share of true incomes of full compliers*)*	$H_M^d = \sum_{i=1}^{N} \frac{m_i(1-\eta_i)}{M} = 1 - H_M$
intensity of tax compliance (average rate of compliance)	$\bar{d} = \frac{X}{M} = \sum_{i=1}^{N} d_i f_i^m = 1 - \bar{e}$
intensity of tax compliance of evaders (average rate of compliance of tax evaders)	$\bar{d}_E = \frac{X_E}{M_E} = \frac{1}{H_M} \sum_{i=1}^{N_E} d_i f_i^m = 1 - \bar{e}_E$

3 Tax evasion profiles

As done in poverty analyses[2], if individual o grouped data are available it is possible to draw a curve, the *tax evasion profile,* that summarizes the three '*I*'s of tax evasion: incidence, intensity and inequality.

The profile represents the cumulative values of a "tax evasion variable" $y = \{y_1,..., y_i,..., y_N\}$ (on the vertical axis) as a function of the cumulative share of a "reference variable" $w = \{w,..., w_i,..., w_N\}$ (on the horizontal axis). The profile has the typical concave shape (as it is actually a rotated Lorenz curve) if individuals are arranged in decreasing order according to a variable $z = \{z_1,..., z_i,..., z_N\}$ *that* represents a vector of individual measures of tax evasion: $z_i = y_i / w_i$.

[2] Jenkins-Lambert (1997), Xu-Osberg (2001).

In general, the profile of a tax evasion variable *y* can be defined by the following notation:

$$L(y, F_i^w | \tilde{z}) = \sum_{j=1}^{i} y_j \quad \text{with} \quad i = 1,..., N$$

in which:

- $F_i^w = \sum_{j=1}^{i} w_j = \sum_{j=1}^{i} f_j^w$ is the cumulative distribution of the reference variable *w*,
- $\tilde{z}$ means that individuals are arranged in decreasing order of *z* ($z_1 > ... > z_i > ... > z_N$).

In the case of poverty there are two possibilities, as the poverty profile can be defined by using either *absolute* poverty gaps or poverty gaps *normalized* with the poverty line. When dealing with tax evasion, more choices can be made:

- about the reference variable on the horizontal axis:
  - cumulative population share
  - cumulative true income share
  - cumulative taxes due
- about the tax evasion variable on the vertical axis:
  - cumulative unreported income share
  - cumulative unpaid taxes shares
- about the values of shares of the tax evasion variable:
  - absolute values
  - normalized values

Combining the 3 types of choices we can build up 12 profiles. To simplify the analysis I just focus on the *normalized unreported incomes profile*, based on the following choices:

- cumulative share of true income (in the horizontal axis);
- cumulative share of unreported income (in the vertical axis), normalized with the total amount of true incomes;
- individuals arranged in decreasing order according to the individual rates of evasion.

In Figure 1 the horizontal axis represents the cumulative share of true incomes:

[9] $$F_i^m = \sum_{j=1}^{i} \frac{m_j}{M} = \sum_{j=1}^{i} f_j^m$$

For the *i-th* individual, the normalized unreported income is:

[10] $$\frac{u_i}{M} = \frac{u_i}{m_i} \frac{m_i}{M} = e_i f_i^m$$

that is the individual rate of evasion e_i weighted with the true income share f_i^m. The variable in the vertical axis is the normalized unreported incomes profile:

[11] $$L\left(\frac{u}{M}, F_i^m \middle| \tilde{e}\right) = \sum_{j=1}^{i} \frac{u_j}{M} = \sum_{j=1}^{i} e_j f_j^m$$

that is a function of F_i^m for the data set $u = (u_1, u_2, ..., u_N)$ normalized with the total amount of true income *M*, where $\tilde{e}$ refers to the sequence of e_i's arranged in decreasing order.

For all evaders, i.e. for all $i = 1, ..., N_E$ which have $e_i > 0$, the cumulative share of true income (the abscissa of point *A* in Figure 1) is equal to the *incidence of evaders' true income*:

[12] $$\sum_{j=1}^{N_E} \frac{m_j}{M} = \frac{M_E}{M} = H_M$$

and the maximum value of the profile is equal to the *intensity of tax evasion*, i.e. the average rate of evasion for all individuals (the ordinate of point *A* in Figure 1):

[13] $$L\left(\frac{u}{M} / M, H_M \middle| \tilde{e}\right)) = \sum_{j=1}^{N_E} e_j f_j^m = \bar{e}$$

The curve starts from (*0,0*) and increases as subsequent tax evaders are added. Each segment represents a tax evader. For individual *i*, the slope of the segment is equal to her rate of evasion:

[14] $$\frac{L(u/M, F_i^m|\tilde{e}) - L(u/M, F_{i-1}^m|\tilde{e})}{F_i^m - F_{i-1}^m} = \frac{\sum_{j=1}^{i} e_j f_j^m - \sum_{j=1}^{i-1} e_j f_j^m}{\sum_{j=1}^{i} f_j^m - \sum_{j=1}^{i-1} f_j^m} = e_i$$

So segments have decreasing slopes as subsequent tax evaders have decreasing rate of evasion. When e_i becomes zero, subsequent individuals are full compliers and therefore no unreported income is added, therefore the curve becomes horizontal at the right of point *A*. At that point we can see the indices of tax evasion: on the horizontal axis the dotted line shows the incidence of evaders' true income (H_M), while on the vertical axis the dotted line shows the intensity of tax evasion ($\bar{e}$).

The *inequality among rates of evasion* is highlighted by the concavity of the arc $0\widehat{A}B$.

The slope of the line from the origin to the point *A* is the ratio:

$$slope\ OA = \frac{\bar{e}}{H_M} = \frac{H_M \bar{e}_E}{H_M} = \bar{e}_E$$

that is the average rate of tax evasion of evaders. The slope of the line *OB* is equal to the intensity of tax evasion, $\bar{e}$.

As an example, consider the micro data of Table 3. There are 10 people with true income ranging from 10 to 100. Four of them are tax evaders and their rate of evasion varies from

30% to 75%. With these figures, the incidence of tax evaders is 50%, the incidence of evaders' true income is 45% and the intensity of tax evasion is 27% (see Table 4). Figure 1 shows the normalized unreported income profile.

Table 3 – Example of individual data on tax evasion

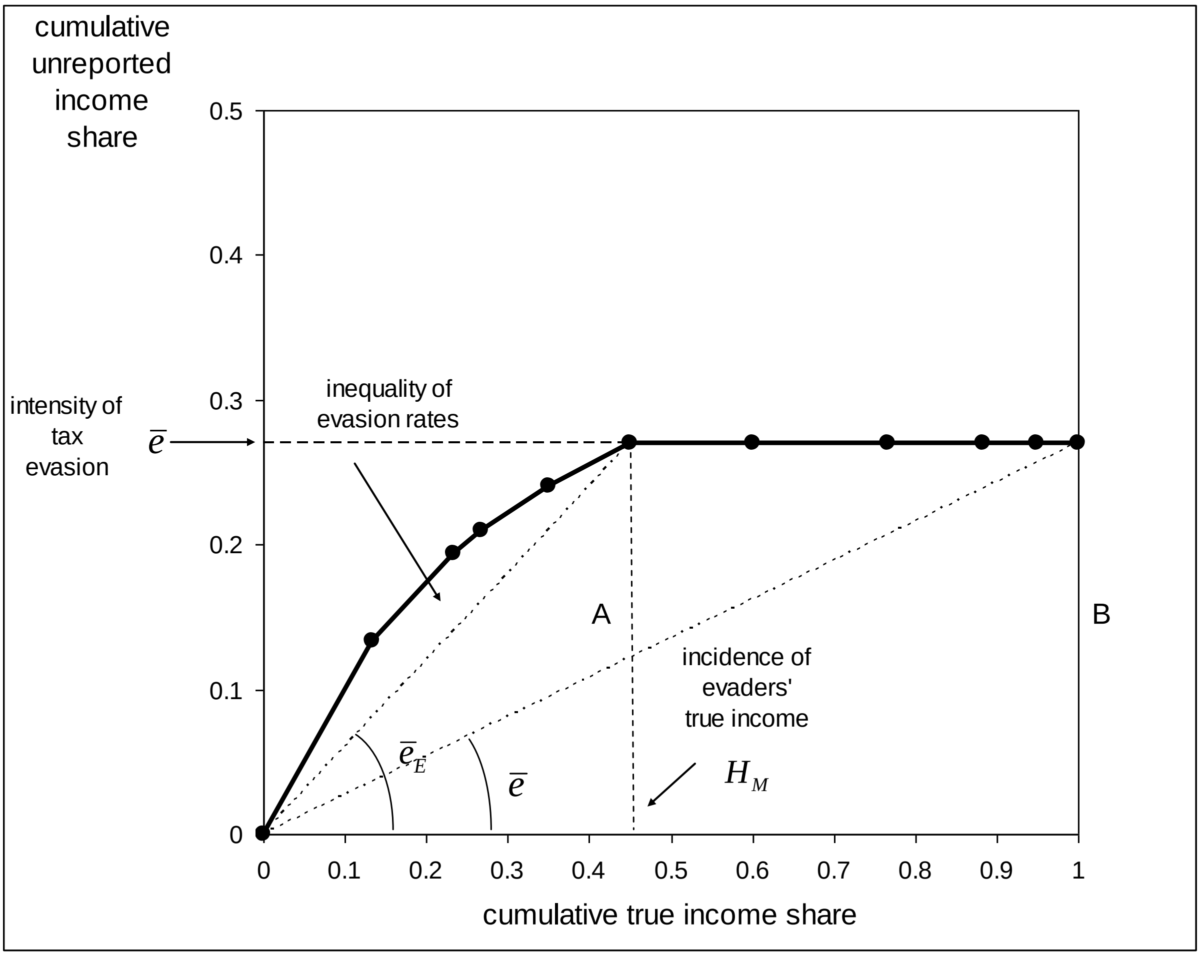
Individuals	*True income*	*Declared income*	*Unreported income*	*Rate of tax evasion*	*Tax evaders (=1)*	*Cumulative true income share*	*Cumulative normalized unreported income share*
i	m_i	x_i	u_i	$e_i = \frac{u_i}{m_i}$	η_i	F_i^m	$L(u/M, F_i^m \mid \tilde{e})$
1	80	0	80	1.00	1	0.133	0.133
2	60	24	36	0.60	1	0.233	0.193
3	20	10	10	0.50	1	0.267	0.210
4	50	32	18	0.36	1	0.350	0.240
5	60	42	18	0.30	1	0.450	0.270
6	90	90	0	0	0	0.600	0.270
7	100	100	0	0	0	0.767	0.270
8	70	70	0	0	0	0.883	0.270
9	40	40	0	0	0	0.950	0.270
10	30	30	0	0	0	1.000	0.270
Total	600	438	162		5		

Note: People ranked according to non-increasing rate of evasion

Table 4 – Indices of tax evasion and tax compliance

$H = \frac{N_E}{N} = \frac{5}{10} = 50\%$	$H^d = 1 - H = 50\%$
$H_M = \frac{M_E}{M} = \frac{270}{600} = 45\%$	$H_M^d = 1 - H_M = 55\%$
$\bar{e} = \frac{U}{M} = \frac{162}{600} = 27\%$	$\bar{d} = 1 - \bar{e} = 73\%$
$\bar{e}_E = \frac{U}{M_E} = \frac{162}{270} = 60\%$	$\bar{d}_E = 1 - \bar{e}_E = 40\%$

Figure 1 – A tax evasion profile



4 Tax compliance profiles

A profile similar to the one introduced in the previous section can be derived by considering the tax compliance rate. The profile represents the cumulative values of declared income $x = \{x_1,...,x_i,...,x_N\}$ normalized with the total amount of true income (on the vertical axis) as a function of the cumulative share of the true income $m = \{m_1,...,m_i,...,m_N\}$ (on the horizontal axis). Individuals are arranged in increasing order according to their tax compliance rate: $d_i = x_i / m_i$.

The profile is defined by:

$$L\left(\frac{x}{M} / M, F_i^m \middle| d\right) = \sum_{j=1}^{i} \frac{x_j}{M} = \sum_{j=1}^{i} d_j f_j^m \quad \text{with} \quad i = 1,..., N \qquad [15]$$

in which the set of reported incomes, x, is arranged in increasing order of d ($d_1 < ... < d_i < ... < d_N$).

By construction, the slope of the profile is the individual measure of tax compliance:

[16] $$\frac{L(x/M,F_i^m|d)-L(x/M,F_{i-1}^m|d)}{F_i^m-F_{i-1}^m}=\frac{x_i}{m_i}=d_i$$

As shown in Figure 2, the curve starts from the origin and increases as subsequent individuals are added. Each segment is associated with an individual and its slope represents the individual tax compliance rate d_i, so individuals with higher slopes comply more. If some individuals are full compliers the profile has a slope equal to 1 at the right of some point *A*, because $d_i = 1$ (i.e. $e_i = 0$) for all subsequent individuals. Therefore all evaders are plotted on the left of point *A* (with $d_1 < ... < d_{N_E} < 1$) and all full compliers are plotted on the right (with $d_{N_E+1} = ... = d_N = 1$). The abscissa of point *A* shows the *incidence measure of tax evasion* H_M, while the ordinate shows the share of reported incomes of evaders on the overall true income:

[17] $$L\left(\frac{x}{M},H_M|d\right)=\sum_{j=1}^{N_E}\frac{x_j}{M}=\sum_{j=1}^{N_E}d_j f_j^m=\bar{d}_E H_M$$

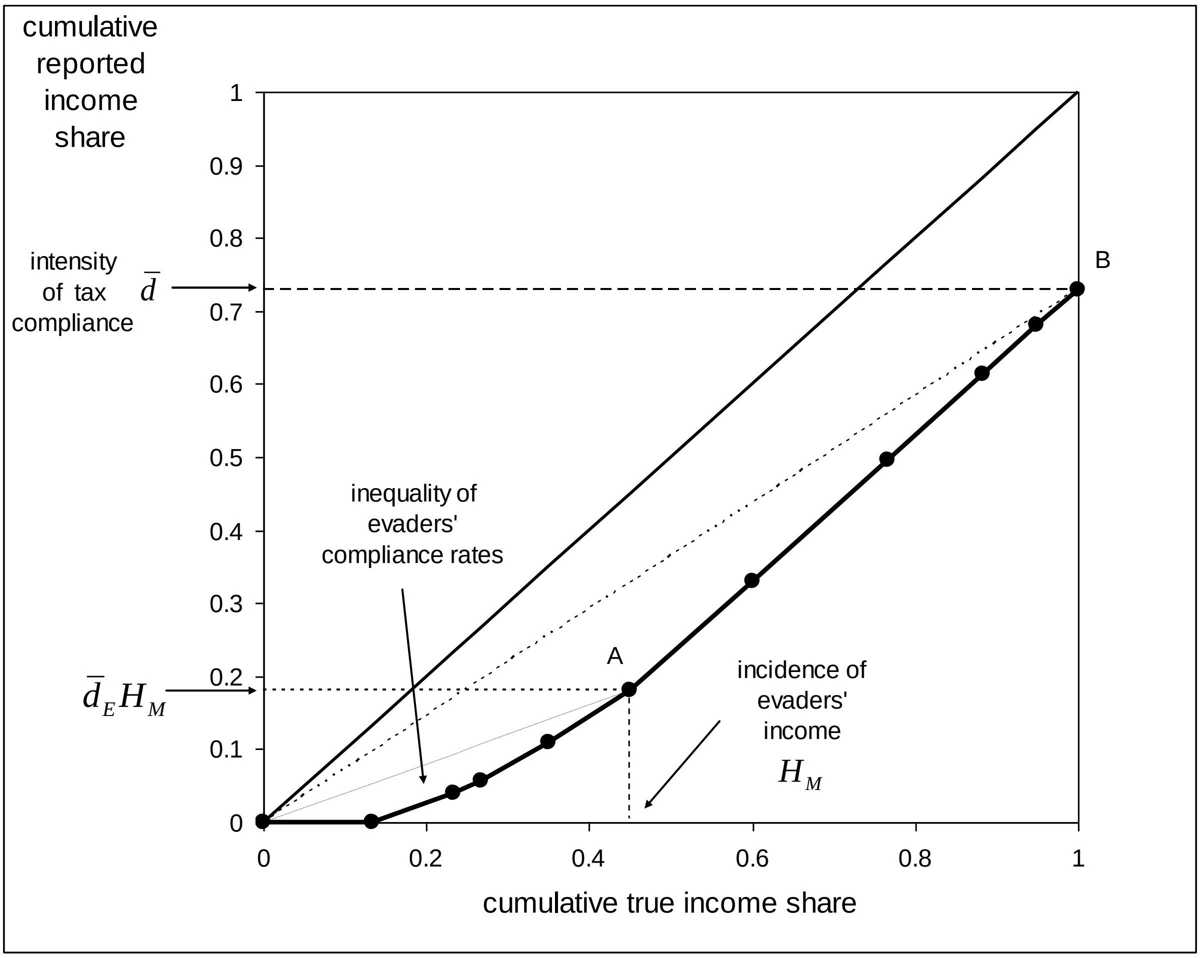
The slope of the straight line connecting the origin with point *A* is the average value of *compliance* among tax evaders: $\bar{d}_E$. The profile ends at point *B*, where all individuals are added to the profile. The ordinate of point B is:

[18] $$L(x/M,F_N^m|d)=\sum_{j=1}^{N}\frac{x_j}{M}=\sum_{j=1}^{N}d_j f_j^m=\frac{X}{M}=\bar{d}$$

The slope of the straight line connecting the origin with point *B* is the average value of compliance among all individuals: $\bar{d}=\sum_{i=1}^{N}d_i f_i^m$, that is an *intensity measure of tax compliance.* The *inequality of individual measure of tax compliance* is shown by the concavity of the profile.

If all individuals are tax evaders and have the same value of the compliance rate ($d_1 = ... = d_N = \bar{d} > 0$) the profile is the dotted line connecting the origin with point *B*.

Figure 2 – A tax compliance profile



As an example, consider again the micro data presented in Table 2: the intensity of tax compliance is 73% ($\bar{d} = X / M = 438/600$). If all the positive values of the tax compliance rate are equal, then the profile is the straight line connecting the origin with point *A*, and its slope is equal to the average rate of compliance of tax evaders: $\bar{d}_E = X_E / M_E = 108/270 = 40\%$.

5 A composite index of tax evasion

A composite index of tax evasion can be derived from the unreported income profile simply by normalizing the area below the curve with the maximum value of the area.

The maximum area below a tax evasion profile is reached when "*all evade all taxes*", i.e. when the incidence of evaders' income is $H_M = 1$ and the intensity of unreported income is $\bar{e} = 1$ (see fig. 1). In this case the tax evasion profile is a straight line from (0,0) to (1,1) and the area below the curve is equal to 0.5 (i.e. $H_M \times \bar{e} / 2 = 1 \times 1/2$).

The tax evasion profile defined in Section 3 is an inverted Lorenz curve that represents the cumulative share of normalized unreported income:

[19] $$L\left(\frac{u}{M}, F_i^m \middle| \tilde{e}\right) = \sum_{j=1}^{i} \frac{u_j}{M}$$

where individuals are arranged in non-increasing order with respect to their rate of evasion e_i, i.e. if $e_1 \geq e_2 \geq ... \geq e_i \geq ... \geq e_N$.

The area below the profile in Figure 1 can be computed with the geometric method: with *N* individuals, the area is a sum of *N* trapezoids. For the *i-th* trapezoid the parallel bases are the values of the ordinates, i.e. $L(u/M, F_i^m|\tilde{e})$ and $L(u/M, F_{i-1}^m|\tilde{e})$, while the height is $f_i^m = F_i^m - F_{i-1}^m$. For ease of notation we define $L_u(F_i^m) = L(u/M, F_i^m|\tilde{e})$, so the area below the profile is:

$$a = \frac{1}{2}\left[L_u(F_1^m) + 0\right]f_1^m + \frac{1}{2}\left[L_u(F_2^m) + L_u(F_1^m)\right]f_2^m + ... + \frac{1}{2}\left[L_u(F_N^m) + L_u(F_{N-1}^m)\right]f_{N-1}^m$$

By rearranging and substituting $L_u(F_i^m) = \sum_{j=1}^{i} u_j / M$:

[20] $$a = \frac{1}{2M}\sum_{i=1}^{N} u_i\left(2(1 - F_i^m) + f_i^m\right)$$

To obtain the composite index of tax evasion we normalize the area *a* with respect to the maximum value of the area when "*all evade all taxes*", (equal to 0.5):

[21] $$E = \frac{1}{M}\sum_{i=1}^{N} u_i\left(2(1 - F_i^m) + f_i^m\right)$$

From Fig. 1 it is also possible to compute the *Gini inequality index of rates of evasion*. As the maximum value of the profile is $\bar{e}$ instead of 1, the Gini index can be defined as:

[22] $$G\left(\frac{u}{M}, F_m, \tilde{e}\right) = 1 - 2\frac{a}{\bar{e}} = 1 - \frac{1}{\bar{e}M}\sum_{i=1}^{N} u_i\left(2(1 - F_i^m) + f_i^m\right)$$

Substituting in [21], the composite index of tax evasion becomes:

[23] $$E = \bar{e}\left[1 - G(u/M, F_m, \tilde{e})\right]$$

It is also known that the Gini index changes the sign when elements are arranged in the opposite order, so that $G(u/M, F_m, \tilde{e}) = -G(u/M, F_m, e)$, therefore the composite index of tax evasion can be written as:

[24] $$E = \bar{e}\left[1 + G(u/M, F_m, e)\right]$$

As $\bar{e} = H_M \bar{e}_E$, this composite index is a simple transposition of the Sen-Shorrocks-Thon poverty index (Shorrocks, 1995, Xu-Osberg, 2001) to tax evasion:

[25] $$E = H_M \bar{e}_E \left[1 + G\left(u / M, F_m, e\right)\right]$$

So, the index E incorporates the *three I's of tax evasion*:

- the incidence of evaders' true income: H_M,
- the intensity of tax evasion among evaders: $\bar{e}_E$,
- the index of inequality of rates of evasion: $G\left(u / M, F_m, e\right)$.

6 A composite indices of tax compliance

Similarly, a composite index of tax compliance can be derived from the reported incomes profile by normalizing the area below the curve with the maximum value of the area.

In this case, the maximum area below a tax compliance profile (see Figure 2) is reached when "*all pay all taxes*", i.e. when the incidence of compliers' income is $H_M^d = 1$ and the intensity of reported income is $\bar{d} = 1$, so that the tax compliance profile is a straight line from (0,0) to (1,1) and the area below the curve is equal to 0.5. In Figure 2 the profile can be seen as a Lorenz curve $L(x / M, F_i^m | d) = \sum_{j=1}^{i} \frac{x_j}{M}$, which represents the cumulative share of normalized reported income when individuals are arranged in increasing order with respect to their rate of compliance d_i, i.e. if $d_1 \leq d_2 \leq ... \leq d_i \leq ... \leq d_N$.

Following the same procedure of the previous Section the composite index of tax compliance can be written as:

[26] $$D = \frac{1}{M} \sum_{i=1}^{N} x_i \left(2(1 - F_i^m) + f_i^m\right)$$

As before, it is also possible to compute the *Gini inequality index of rates of compliance*, taking into account that the maximum value of the profile is $\bar{d}$ instead of 1:

[27] $$G\left(x / M, F_m, d\right) = 1 - 2\frac{a}{\bar{d}} = 1 - \frac{1}{\bar{d}M} \sum_{i=1}^{N} x_i \left(2(1 - F_i^m) + f_i^m\right)$$

Substituting in [27], the tax compliance index is:

[28] $$D = \bar{d}\left[1 - G\left(x / M, F_m, d\right)\right]$$

Again, the composite index of tax compliance is very similar to the Sen-Shorrocks-Thon poverty index, as it depends on the average rate of compliance $\bar{d}$ (an intensity measure) and on the Gini index of normalized reported incomes.

Looking at the Gini indexes, it can be shown that:

[29] $$G(x/M, F_m, d) == \frac{\bar{e}}{1-\bar{e}} G(u/M, F_m, e)$$

i.e. the Gini index of unreported incomes is connected to the Gini index of reported incomes by means of the average rate of evasion $\bar{e}$. It is worthwhile noticing that substituting $G(x/M, F_m, d)$ in [29] we obtain:

[30] $$D = 1 - \bar{e}[1 + G(u/M, F_m, e)] = 1 - E$$

so the compliance index is the complement to 1 of the tax evasion index.

7 An example with ABM simulations

In order to provide an example of tax evasion profiles and indices in comparing different situations, we generate some distributions of rates of evasion by using a simple agent-based model.

Following the standard Allingham-Sandmo-Yitzhaki[3] framework where tax payers maximize their expected utility we define the individual utility as:

[31] $$U_i(y_i, e_i) = (1-e_i)^{k_i} y_i^{1-\rho_i}$$

where:

- y_i is net income;
- e_i is the tax evasion rate, defined as the share of income not reported ($0 \le e_i \le 1$);
- k_i is an individual parameter representing the attitude to comply;
- ρ_i is the individual risk aversion parameter.

The income tax is applied at a constant rate t to the exogenous amount of income I_i, which is not known to the Government. Hence, the amount of taxes paid by individual i is $T_i = t(1-e_i)I_i$. The Government controls taxpayers with probability p: if an individual is audited, the tax evasion is certainly discovered and the tax payer has to pay a fine of f times the amount of evaded tax: $F_i = fte_iI_i$.

If the taxpayer is not controlled, the net income is

[32] $$W(e_i) = I_i - T_i = I_i[1 - t(1-e_i)]$$

If, instead, the taxpayer is audited, the net income is:

[3] Allingham and Sandmo (1972), Yitzhaki (1974). See Pyle (1991) and Sandmo (2005) for a review of the subsequent literature.

[33] $$Z(e_i) = I_i\left[1 - t(1 + fe_i)\right]$$

Given the tax rate, the probability of control and the fine, each taxpayer chooses the share of income to unreport, e_i, in order to maximize the expected utility:

[34] $$EU(e_i) = pU_i\left(Z_i, e_i\right) + (1-p)U_i(W_i, e_i)$$

Two tax evasion profiles are simulated. Simulation *A* is characterized by lower tax rate, probability of control and fine but higher attitude to comply with respect to simulation *B* (see Table 5). Using the terminology of Kirchler et al. (2008), the society described by simulation *A* has more "trust" and less "power" than that described by simulation *B*. The resulting indices and profiles are shown in Figure 3, while Figure 4 shows the tax compliance profiles.

With respect to simulation *B*, in simulation *A* the composite index of tax evasion *E* (23.2% versus 30.5%), incidence and intensity indices are lower, while the inequality among the rates of evasion is higher (the Gini index is 81.5% versus 47.1%). From Figures 3 and 4 we can see that the profiles intersect, so it would be possible to study the conditions for *tax evasion dominance* (see Jenkins and Lambert, 1997).

Table 5 – Parameters and results of the simulated tax evasion profiles

Parameters			
income distribution	y_i	$\ln N(30000,2)$	
risk aversion distribution	ρ_i	$U(0.0,1.0)$	
		Simulation A	*Simulation B*
attitude to comply distribution	k_i	$U(0.0,0.4)$	$U(0.0,0.2)$
tax rate	t	20%	30%
probability of control	p	2%	5%
fine	f	200%	500%
Results			
Incidence of evaders	H	33.0%	70.0%
Incidence of evaders' true income	H_M	27.3	71.4%
Intensity of tax evasion	$\bar{e}$	12.8	20.7%
Intensity of tax evasion of evaders	$\bar{e}_E$	46.8%	29,0%
Gini index of rates of evasion	$G(u/M, F_m, e)$	81.5%	47,1%
Composite index of tax evasion	E	23.2%	30.5%
Composite index of tax compliance	D	76.8%	69.5%

Figure 3 – Simulated tax evasion profiles

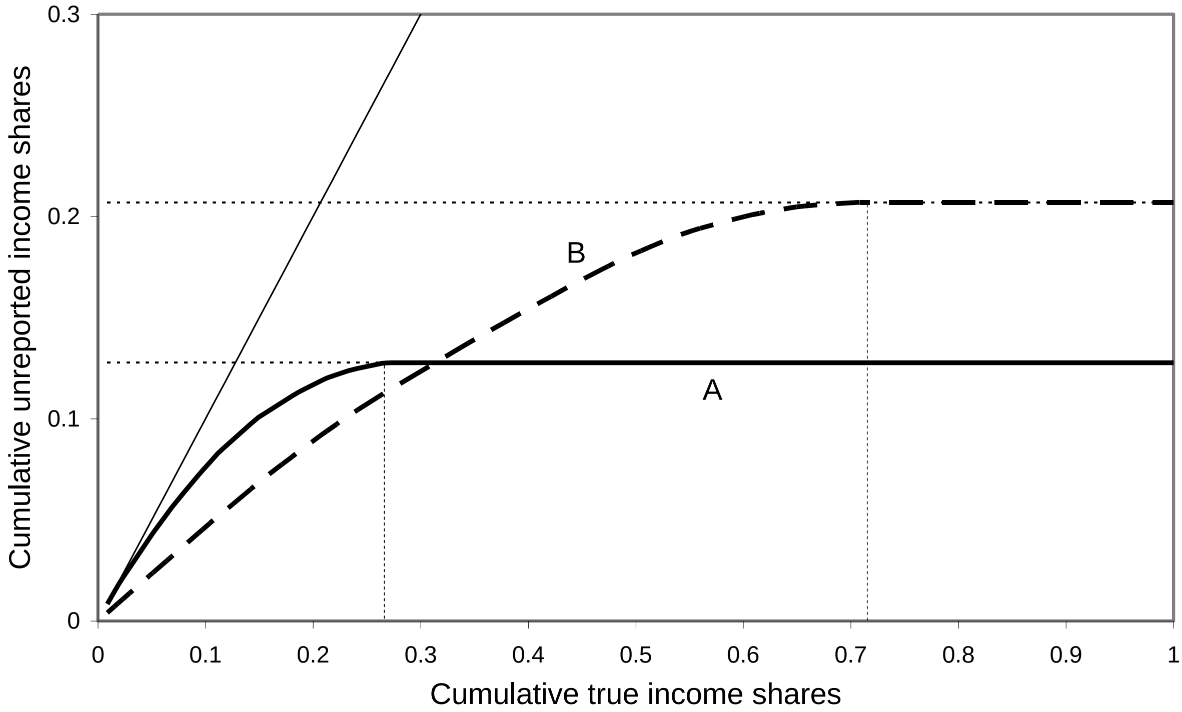
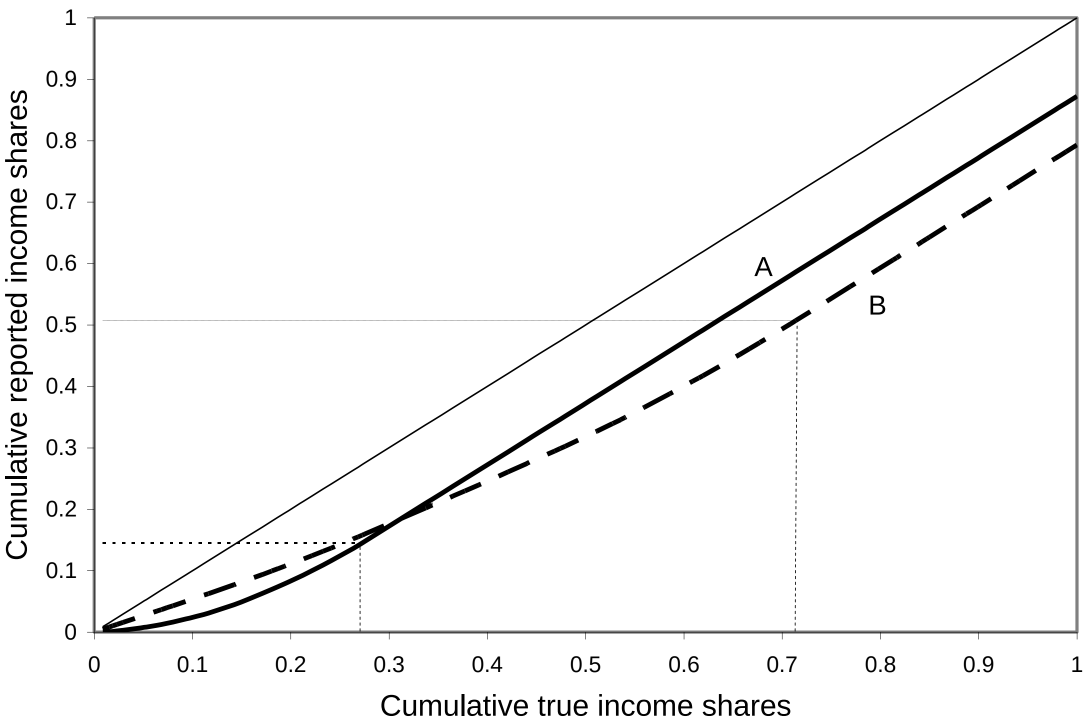


Figure 4 – Simulated tax compliance profiles



8 Conclusions

In this paper we adapted some concepts, indices and graphical representation that were originally developed for poverty analysis to the analysis of tax evasion. The composite index of tax evasion and the tax evasion profile can have a practical use for analyzing individual level data produced by agent-based models of tax evasion, increasingly used in the literature.

The tax evasion profile can also be used if aggregate data about tax evasion are available for some taxpayers' characteristics. For instance, if we knew that in a given country the self-employed earned the 30% of gross income with a tax evasion rate of 50%, while the employees earned the remaining 70% with a tax evasion rate of 5%, then we could easily draw the tax evasion profile and compute the average tax evasion rate (18.5%), the Gini index of inequality (11.35%) and composite index of tax evasion (20.6%).

Future research could address the adapation to tax evasion of other theorethical results borrowed by poverty analysis, such as conditions for tax evasion dominance and ethical issues.

References

Allingham M. G., Sandmo, A. (1972), Income Taxation: A Theoretical Analysis, *Journal of Public Economics*, 1, 323-338.

Bloomquist K. M. (2006), A Comparison of Agent-Based Models of Income Tax Evasion, *Social Science Computer Review*, 24(4), 411-425.

Bloomquist K. M. (2011), *Agent-Based Simulation of Tax Reporting Compliance*, mimeo, Shadow 2011, Münster.

Davis J.S., Hecht G., Perkins J.D. (2003), "Social Behaviors, Enforcement and Tax Compliance Dynamics", *Accounting Review*, Vol. 78(1), pp. 39-69.

Foster J., Greer J., Thorbecke E. (1984), "A Class of Decomposable Poverty Measures", *Econometrica*, 56: 761-766.

Jenkings S.P., Lambert P.J. (1997), "Three 'I's of Poverty Curves, with an Analysis of UK Poverty Trends", *Oxford Economic Papers*, 49, pp. 317-327.

Hokamp S., Pickhardt M. (2010), Income Tax Evasion in a Society of Heterogeneous Agents – Evidence from an Agent-Based Model, *International Economic Journal*, 24(4), 541-553.

Kirchler E., Hoelzl E., Wahl I. (2008), Enforced versus voluntary tax compliance: The slippery slope framework, *Journal of Economic Psychology*, 29, 210-225.

Korobow A., Johnson C., Axtell, R. (2007), An Agent-based Model of Tax Compliance with Social Networks, *National Tax Journal*, Vol. 60(3), 589-610.

Mittone L., Patelli P. (2000), "Imitative behaviour in tax evasion", in Stefansson B. and Luna, F., (eds.), *Economic simulations in swarm: Agent-based modelling and object oriented programming*, 133-158, Amsterdam: Kluwer.

Pickhardt M., Seibold G. (2011), *Income tax evasion dynamics: Evidence from an agent-based econophysics model*, CAWM Discussion Papers 53, Center of Applied Economic Research Münster (CAWM), University of Münster.

Pyle D.J. (1991), The Economics of Taxpayer Compliance, *Journal of Economic Surveys*, 5(2), 163-198.

Sandmo A. (2005), The Theory of Tax Evasion: A Retrospective View, *National Tax Journal*, 58(4), 643-663.

Sen A.K. (1976), "Poverty. An Ordinal Approach to Measurement", *Econometrica*, 44, pp. 219-231.

Shorrocks A.F. (1995), "Revisiting the Sen Poverty Index", *Econometrica*, 63, 5, pp. 1225-1230.

Thon D. (1979), "On Measuring Poverty," *Review of Income and Wealth*, 25, pp. 429–40.

Yitzhaki S. (1974). A Note on Income Tax Evasion, *Public Finance Quarterly*, 15(2), 123-137.

Xu K., Osberg L. (2001), "How to Decompose the Sen-Shorrocks-Thon Poverty Index: A Practitioner's Guide", *Journal of Income Distribution*, 10, 1-2, pp. 77-94.

Xu K. (2003), *How has the Literature on Gini's Index Evolved in the Past 80 Years?* , April. Dalhousie University, Economics Working Paper. Available at SSRN: http://ssrn.com/abstract=423200 or doi:10.2139/ssrn.423200

Zaklan G., Westerhoff F., Stauffer D. (2009), Analysing Tax Evasion Dynamics via the Ising Model, *Journal of Economic Interaction and Coordination*, 4(1), 1-14.

Zeng B. (2000), "Poverty Orderings", *Journal of Economic Surveys*, 14, 4, pp. 427-466.