

Tax Systems Simulation Using Integrated Administrative Data

Paolo Mariani

Dipartimento di Statistica, Via Bicocca degli Arcimboldi 8 20126 Milano, e-mail:
paolo.mariani@unimib.it

Mauro Mussini

Dipartimento di Statistica, Via Bicocca degli Arcimboldi 8 20126 Milano, e-mail:
m.mussini@campus.unimib.it

Biancamaria Zavanella

Dipartimento di Statistica, Via Bicocca degli Arcimboldi 8 20126 Milano, e-mail:
biancamaria.zavanella@unimib.it

Abstract: This paper presents the opportunities for measuring taxation effects by using combined administrative data from AMeRiCA Data Warehouse. Data combination provides demographic and income information on individuals and families resident in the Milanese area thereby permitting the evaluation of the impact of tax reforms on a specific taxpayer type as well as the simulation of the application of various tax systems. In our analysis we shall concentrate on Milanese families and we measure the impact of five different tax systems in force during 2005 from various European countries (France, Germany, Ireland, Italy and Spain).

Keywords: Administrative Data Combination, Income Taxation, Family Characteristics

1. Introduction

The use of administrative data for statistical purposes has become increasingly common for local area analysis (Brackstone, 1987; Tronti, 2008; Zavanella, 2004). This rising interest for administrative data is due to the opportunities for research provided by its availability. Firstly, unless a survey that includes all the relevant variables for the aims of a specific research, administrative data can be used to supplement this lack of information on the variables of interest. Secondly, the survey sample may be too small to obtain representative statistics when the focus of the research is on a local area such as a town or a province. An administrative agency usually only collects relevant data for its specific purposes. When a research issue requires answers available in separate administrative sources the linkage of data records can lead to more detailed statistical analysis than ones based on information provided by each individual administrative source (Martini, 1995).

The primary objective of this paper is to show that using combined administrative data may serve the scope of investigating the redistributive effect of taxation of the population resident in the Milanese area. To pursue this aim, we use data provided by the AMeRiCA Data Warehouse that combines administrative records from the tax register of the Milan Revenue Agency with personal data from the Milanese Registry Office. Data from AMeRiCA contains pre-tax income for each individual resident in the Milanese area and provides information concerning his family composition which is relevant to determine the actual income tax paid thereby permitting the evaluation of the effects which derive from

the application of different tax systems both on individuals and families. Our analysis concentrates on specific family types so as to compare the treatment of every family type considered after applying tax systems in force during 2005 from various European countries.

2. Data Description

Data used for the analysis derives from AMeRiCA which provides demographic and income information concerning individuals and households resident in the Milanese area. The structure of AMeRiCA is based on a Data Warehouse that combines administrative data records from the tax register of the Milan Revenue Agency with personal data available at the archives of the Milanese Registry Office. The Revenue Agency registers the Italian Personal Income Tax (IRPEF) in categories: pre-tax income by source (pre-tax income is defined as the sum of incomes from dependent employment, self-employment, company income, rental income, agricultural income, real estate income, and other incomes), taxable income, gross IRPEF, net IRPEF, tax allowance, deduction for dependent family members and other deductible costs. The archives of the Registry provide information about marital status, citizenship, address, family composition, residence and other personal data (gender, date of birth) for each registered individual. The combination of the two administrative sources enables AMeRiCA to render information on pre-tax income, income tax paid, tax allowance and deductions plus family size and composition of residents in the Milanese area.

We focus on household income data available for 2005. After deleting observations with non-positive pre-tax income, we refer to a population composed of 672020 homes.

3. Application

In this section we focus on the opportunities for the analysis of the redistributive effect of taxation that are offered by the use of the AMeRiCA tax records by applying the French, German, Irish, Italian and Spanish tax systems in force during 2005 (OECD, 2005) which present diverse definitions of taxable income, tax schedules, tax units, accounting for dependent family members (Baclet et al., 2005). Since the definition of taxable income differs among the considered tax systems, we therefore assume that pre-tax income equals taxable income. However, this assumption does not imply a limit for our analysis as we are not interested in drawing conclusions about the fairness of the different tax systems.

Given the differences in the aforementioned tax systems we simulate their application by starting from the pre-tax income of a tax unit. In so doing we consider the chosen tax unit for each country tax system; for example, in the case of French tax system family represents the tax unit whereas individual is the tax unit of Italian tax system. We then calculate the post-tax income by considering tax allowance related to the taxpayer's marital status and dependent family members. This procedure ensures that the tax schedule and the allowance for work, marriage and dependent family members are taken into account during the simulation for each country tax system. The analysis aims at comparing the redistributive effect due to different definitions both of the tax schedule and of the tax unit. We then differentiate between various family compositions in order to examine how the treatment of each type of household differs among the applied tax systems. We shall concentrate on eight different household types, namely single with no children (henceforth S), single parent with one, two, three children (hereafter SP1, SP2, SP3 respectively), a

couple with no children (henceforth CS), couple with one, two, three children (hereafter C1,C2,C3 respectively). A child is defined by being either natural or adopted but younger than 18 years old.

4. Main Results

In order to estimate the redistributive effects that derive from the different tax systems, we calculate a set of common summary measures of inequality in the income distribution when moving from pre-tax to post-tax household income distribution. Table 1 shows the results of the different tax systems when considering equivalent household income distribution. The equivalent income is obtained dividing the corresponding nominal income by a divisor (S) that is calculated according to an equivalence scale of Cutler-Katz type (Cutler and Katz, 1992):

$$S = (NA + NC * \alpha)^\beta \quad (1)$$

where NA is the number of adults of the family, NC is the number of children, while the parameters are $\alpha=0.5$ and $\beta=0.6$.

Table 1: *Inequality and redistributive indexes by country tax systems. Milan. 2005.*

Index	France	Germany	Ireland	Italy	Spain
$G^{\text{pre-tax}}$	0.4923	0.4923	0.4923	0.4923	0.4923
$G^{\text{post-tax}}$	0.4580	0.4271	0.4308	0.4513	0.4573
RS	0.0391	0.0659	0.0640	0.0421	0.0351
K	0.2285	0.2480	0.2660	0.1242	0.1720
RE	0.0343	0.0652	0.0615	0.0410	0.0350
R	0.0048	0.0007	0.0025	0.0011	0.0001

Source: own elaborations of the AMeRICA data.

We notice that the Reynolds-Smolensky (RS) index shows higher values for the German and the Irish systems than for the other. This implies that the impact of redistribution is higher by applying these two tax systems. The redistributive impact derives from both the incidence of aggregate tax rate and from tax progression measured by the Kakwani index (K). Looking at Table 1, we note that K is higher for the Irish tax system rather than for the German one, and thus, the higher redistributive effect generated by the German system is due to the higher aggregate tax rate of the German tax system. In fact the German aggregate tax rate equals 20.6 per cent and therefore higher than the Irish one (19.2 per cent). As the Italian tax system shows lower progressivity index (0.1242), we can therefore conclude that the shift from proportional taxation to progressive taxation is slighter than for the other tax systems. However, the RS index is higher for the Italian tax system rather than for the French and Spanish tax systems. This is due to the Italian aggregate tax rate (25.5 per cent) which is higher than the aggregate tax rates of the other country tax systems. The post-tax Gini index ($G^{\text{post-tax}}$) shows higher value for the French tax system than for the other. The French tax system exhibits the highest reranking index (R) whereas reranking is almost negligible for the Spanish tax system. As the actual redistributive effect (RE) is given by the difference between RS and R, we then notice that the Spanish actual redistributive effect is higher than the French one.

Table 2 shows the actual average tax rates by the aforementioned household types. By comparing the actual average tax rates related to various family types, we notice that

according to the French, the German and the Irish tax systems, SP2 and SP3 face higher actual average tax rates than those of couples with the same number of dependent children, C2 and C3 respectively, although their pre-tax average incomes are definitely lower than the pre-tax average incomes of C2 and C3.

Table 2: *Actual average tax rates by household types. Milan. 2005.*

	S	CS	C1	C2	C3	SP1	SP2	SP3
France	12.6%	11.1%	13.5%	15.6%	17.0%	9.8%	17.8%	18.2%
Germany	22.8%	20.3%	18.7%	19.0%	21.9%	19.4%	21.6%	23.7%
Ireland	19.9%	16.8%	20.6%	21.8%	23.1%	32.2%	34.5%	35.3%
Italy	25.2%	25.1%	26.1%	27.7%	30.4%	25.4%	27.8%	29.0%
Spain	16.9%	17.0%	18.0%	18.0%	19.3%	17.0%	17.8%	18.2%

Source: own elaborations of the AMeRiCA data.

Although the Irish aggregate tax rate is lower than the aggregate tax rate of the Italian tax system its average tax rates referred to single parent family types are higher than the corresponding Italian ones. The Spanish tax system has a more generous treatment towards single parent households. In fact the Spanish actual average tax rates related to SP1, SP2 and SP3 are at the same level of those of couples with the same number of dependent children.

References

- Baclet A., Dell F., Wrohlich K. (2005) Income Taxation and Households Size: Would French Family Splitting Make German Families Better Off? IZA DP No. 1894.
- Brackstone G. J. (1987) Issues in the Use of Administrative Records for Statistical Purposes, *Survey Methodology*, 13, 1, 29-43.
- Cutler D. M., Katz L. F. (1992) Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980's, *American Economic Review*, 82, 2, 546-551.
- Martini M. (1995) Necessità e possibilità di un registro statistico delle imprese in Italia, in: *Il registro statistico europeo delle imprese. Esperienze e metodi per la sua costruzione in Italia*, Biffignandi, S. & Martini, M. (Eds.), FrancoAngeli, 141-166.
- OECD (2005) Social Policies. Benefits and Wages: Statistics. Country specific files. Retrieved July 6, 2008, from http://www.oecd.org/document/29/0,3343,en_2649_34637_39618653_1_1_1_1,00.html.
- Tronti L. (2008) I dati amministrativi per le ricerche sui mercati del lavoro locali: il progetto Guida, paper presented at the Nona Conferenza Nazionale di Statistica, Roma. Retrieved April 8, 2009, from <http://www.istat.it/istat/eventi/conferenze/nonaconf/relazioni.html>.
- Zavanella B. (2004) “Nuovi dati” per l’analisi statistica territoriale: le opportunità offerte dai sistemi informativi della pubblica amministrazione, in: *Qualità degli archivi amministrativi e qualità dell’informazione statistica*, Aimetti, P. & Zavanella, B. (Eds.), FrancoAngeli, 131-149.