Age Distribution of Taxes and Social Benefits by Income Deciles: Evidence from Mexico

Iván Mejía-Guevara

(Extended abstract: March 2010)

Abstract

The high unequal distribution of income in some developing countries is an important restriction for their social and economic development. How this income inequality might influence the mechanisms of taxation and distribution of social benefits is a matter of important consideration for public policy. I focus in this paper on this problem and the effects of the distribution by age for those transfers. In the context of the National Transfer Accounts methodology, and performing analysis for survey data, I show evidence from Mexico about the relationship between income inequality and the age allocation of transfers. The evidence indicates that only the richest decile has significant impact on age distribution of taxes and social contributions.

Introduction

Income inequality in Mexico has increased substantially in last decades. This unequal distribution restricts the opportunities for large segments of the population to get access not only to labor and financial markets, but also limits the access to basic services such as health, housing, education, among others. In (ECLAC 2004) the United Nations' Economic Commission for Latin America and the Caribbean (ECLAC) points out that the high percentage of resources concentrated in the richest ten percent of the population is the hallmark of income distribution in Latin America. In the case of Mexico, it is reported there that the income of the richest decile was around fifteen times bigger than the income of the four poorest deciles in 2004. Although this was not the biggest figure in the region since

many other countries presented higher ratios, it is evident that income disparities in Mexico are substantial.

The analysis of income concentration is very important in many aspects of an economy. In particular, the fiscal policy generates transfers in the form of taxes and public expenditure programs, which importantly affects levels of income and expenditure at the household level (SHCP 2004). Therefore, an estimation of the distribution of payment of taxes and the benefits obtained via public expenditure is relevant to quantify the scope and effectiveness of the mechanisms of taxation and public consumption.

I focus in this article on the distribution of tax payments; in particular, income taxes and social security benefits. The analysis of these two kinds of transfers is given in the context of the National Transfer Accounts (NTA) project, a framework that allows for intergenerational reallocation among different age groups while being consistent at the aggregate level (Lee et al.; Mason et al. 2009). What is particularly important in this framework is the age allocation of transfers both public and private. The selection of only public transfers (income tax, social security contributions and benefits) is also relevant due to important differences in the construction of their age profiles found by Mejía-Guevara (2009) when comparisons between estimates from the entire population and those obtained by considering the distribution of income to form deciles of individuals were performed.

Hypothesis

The findings reported in this article are based on an exhaustive statistical analysis which takes into account the complex sample design of the National survey of income and expense for homes in 2004, called ENIGH-2004 (Encuesta Nacional de Ingresos y Gastos de los Hogares). This survey is the base for the construction of transfers and social security benefits as well as for the most important profiles constructed for Mexico using the NTA methodology.

The objective of the analysis is to contrast differences in the shape of age profiles by different subpopulations (income deciles). In particular, the differences between the richest decile and lower income deciles. This analysis is conducted in terms of differences of means for subpopulations since the a-th element (value of the profile for individuals of age a) of the age profile, for the subpopulation s, is defined as:

$$\overline{y}_{sa} = \frac{\sum_{i=1}^{n_{sa}} w_{sai} y_{sai}}{\sum_{i=1}^{n_{sa}} w_{sai}} \qquad (1)$$

where w_{sai} is the population weight for the individual i of age a in the subpopulation s. The population weight is included in (1) to indicate that the estimates presented here are obtained from a survey population, but they need to represent the total population in Mexico. The definition of y_{sai} is similar, but "y" stands for the value corresponding to the variable of interest; for instance, the income tax paid by that particular person (or the groups of individuals represented by w_{sai}). The term n_{sa} is the sample population of individuals of age a in the subpopulation s; therefore, $N_{sa} = \sum_{i=1}^{n_{sa}} w_{sai}$ is the estimated total population of age a in the subpopulation s. The proportion of the entire subpopulation s, defined in terms of (1) can be computed as follows:

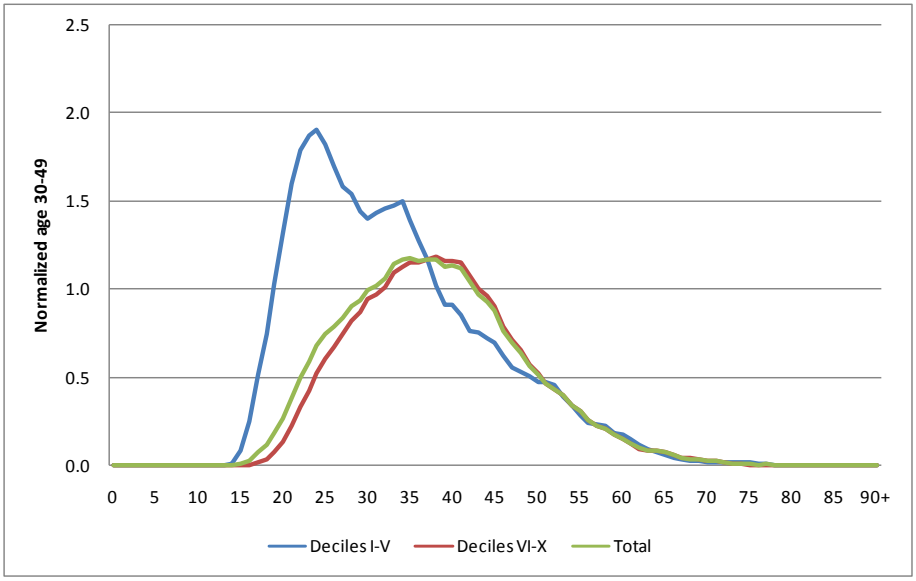
$$\overline{y}_{s} = \frac{\sum_{a=0}^{A} N_{sa} \overline{y}_{sa}}{\sum_{s=1}^{S}\sum_{a=0}^{A} N_{sa}} = \frac{\sum_{a=0}^{A}\sum_{i=1}^{n_{sa}} w_{sai} y_{sai}}{\sum_{s=1}^{S}\sum_{a=0}^{A}\sum_{i=1}^{n_{sa}} w_{sai}} \qquad (2)$$

where A is the age of the oldest person in the survey population and S(=10) stands for the ten deciles defined previously as subpopulations and $N = \sum_{s=1}^{S} \sum_{a=0}^{A} N_{sa}$ is the total population.

Therefore, it is supported here that differences in shape for age profiles in the subpopulations lead to different allocation by age for the entire population, but that difference is significant when discrepancies in subpopulations (income deciles) are remarkable.

Figure 1 displays, for instance, the age allocation of income taxes for deciles one to five (I-V), deciles six to ten (VI-X) and for total taxes. The differences observed there are important for the deciles' distributions, although total taxes' allocation looks pretty alike with that of the richest deciles. However, what is most important at the end is to determine how those differences are relevant when the respective macroeconomic adjustments are implemented to reflect the real values of the contributions. A statistical analysis will shed some light on this in what follows.

Figure 1. Age Distribution of income taxes by income deciles (I-IV, VI-X and Total): Mexico, 2004.



Source: Mejía-Guevara (2009)

Agenda

This article is organized as follows. The next section describes salient characteristics of the data used for the analysis and the methods for complex survey designs implemented in

this paper. Then I present the principal results derived from a survey data regression analysis. The paper concludes with a discussion about the implications of the findings.

Data and Methods

Model specification

A multiple linear regression analysis is used in this article to contrast the association among income deciles and taxes to income, as well as social contributions and benefits. However, the ENIGH does not contain information reported for income taxes and social security contributions. It does in the case of some social cash benefits, like the programs OPORTUNIDADES and PROCAMPO[1]. I focus on these two programs in this paper for the distribution of social benefits. In the case of taxes and social contributions, the approach followed here consists on estimating these contributions using information of the corresponding taxable income; i.e., those concepts of income associated with the payment of payroll taxes and social security contributions. Other source of income taxes -besides payroll taxes- is obtained from enterprises, but there is no information available in the ENIGH to get proper estimation for them. Therefore, the focus in this paper is the estimation of income taxes only paid by individuals. A proper estimation of income taxes and social contributions is highly difficult in practice since there are many taxation rules established in the Mexican Laws (SEGOB 2009, 2009a), which depend on the levels of income, but also on many particular exemptions. Instead, in this paper it is assumed that both income taxes (called ISR in Mexico) and social security contributions (SSC) are proportional to their respective taxable income. No specific flat rates are define here, since the interest here is the age distribution of

[1] PROGRESA is an important program designed as a mechanism to combat poverty in Mexico, whereas PROCAMPO was designed to transfer resources to support the economy of rural producers.

those contributions nor a proper estimation for those taxes, which in terms of the NTA project is obtained by adjusting the taxable income proportional to the corresponding macroeconomic level. The taxable income for ISR is defined as Y_{ISR}, whereas Y_{SSC} defines taxable income for SSC and T_{SB} stands for social benefits. All Y_{ISR}, Y_{SSC} and T_{SB} are dependent variables for the regression models and represent quarterly values reported in the ENIGH.

As I pointed out above, the main interest here is to test the significance of income deciles and their possible relationship to the age distribution of taxable income. Then population is divided by their level of income in deciles. The proper approach is to use SUDAAN® which provides quantile estimations that take into account the complex survey design (Heeringa et al 2010). However, an alternative approach was implemented by using SAS® procedure PROC UNIVARIATE which allows for the estimation of weighted quantile estimates. The quantile estimations were performed using income variables related to the definition of labor income in Mexico. The categorical variable D_Y represents the separation of the population into deciles. This variable defines a covariate for the respective linear regression models of Y_{ISR}, Y_{SSC} and T_{SB}. The second independent variable is AGE, which is used in the specification of the model to control the effect of deciles on the taxable income and cash transfers, respectively. A third dependent variable is D_E which is used to control the effects of different levels of education over the corresponding dependent variables.

A set of multi-parameter Wald tests were performed in Stata® to assess the significance of the bivariate relationships among independent and dependent variables. The first bivariate relationships were tested between Y_{ISR}, Y_{SSC} and AGE, while the second set was as follows: Y_{ISR} vs. D_Y and Y_{SSC} vs. D_Y. These tests were implemented by using a linear regression model in each case.

Table 1. Multi-parameter Design-based Wald tests among dependent and independent variables.

Bivariate Test	Wald Test	Prob >F	Bivariate Test	Wald Test	Prob >F
Y_{ISR} vs. AGE	1,228.22	0.0000	Y_{ISR} vs. D_Y	21,993.73	0.0051
Y_{SSC} vs. AGE	1,214.06	0.0000	Y_{SSC} vs. D_Y	16,696.57	0.0058

The Wald tests are displayed in Table 1. Evidence in all cases was found against the null hypothesis that the parameters associated with the covariates are zero. After a careful analysis, the interaction between independent variables was also included which leaded to the final specification of the model, as follows:

$$y_M = \beta_0 + \beta_1 age + \beta_2 D_Y + \beta_3 age * D_Y + \beta_4 D_E \quad (3)$$

where y_M is the dependent variable and D_Y is the categorical variable defined before for income decile groups and D_E stands for level of education, which is also considered as a continuous variable. The term M stands for ISR, SSC or SB, depending of the model which is being estimated.

Furthermore, a subclass indicator variable was constructed to ensure an unconditional subclass analysis as opposed to a conditional subclass one, which is not recommended because the risk of losing important design features due to deletion of elements that do not belong to the subclass (West et al 2008). The subclass consists of those individuals in the age group 12-90 who received health benefits from public institutions or contribute to the social security pension system. Then, the unconditional subclass analysis will incorporate into the analysis only the people working in the formal sector, which certainly make social contributions and pay income taxes. This consideration is very important given the high fiscal evasion in Mexico, but also because it provides approximate information about the extent of that evasion and levels of social contributions.

It should be emphasized that many concepts of taxable income are used for both ISR and SSC, which lead to very similar values for Y_{ISR} and Y_{SSC}. An unconditional descriptive analysis performed for these dependent variables is shown in Table 2. The SAS® procedure PROC SURVEYMEANS was used to perform the estimations in Table 2 (mean, linearized standard errors and confidence intervals).

Table 2. Unconditional descriptive analysis of dependent variables: Y_{ISR} and Y_{SSC}.

Variable	N	Mean*	Standard Error**	95% Confidence	
Y_{ISR}	13,781	17,634.0	417.74	16,474.60	18,794.28
Y_{SSC}	16,518	16,518.0	393.24	15,425.83	17,609.44

* Weighted estimate.

** Taylor Series Linearization.

From Table 2 we can confirm the similarity of these two dependent variables since their means and standard errors are closed from each other, which is confirmed by noting that their confidence intervals are overlapped. Although the mean of Y_{SSC} is lower than the mean of Y_{SSC}, its standard error is also smaller. Then, the confidence interval of the later is a subset of the confidence interval of the former.

References

ECLAC [Economic Commision for Latin America and the Caribbean]. 2004. *Social Panorama of Latin America 2004*, United Nations.

Heeringa, Steve, Brady West, and Patricia Berglund. Applied Survey Data Analysis. Chapman and Hall, Publication Date: 03/12/2010.

INEGI [Instituto Nacional de Estadística y Geografía]. 2008. *Encuesta nacional de ingreso y gasto de los hogares 2004* [National survey of income and expense for homes, 2004] www.inegi.gob.mx (accessed June 20th 2008).

Lee, Ronald, Sang-Hyop Lee, and Andrew Mason. 2008. Charting the economic lifecycle. In A. Prskawetz, D. E. Bloom, and W. Lutz, (eds.), *Population aging, human capital accumulation, and productivity growth*. Population and Development Review 33(Suppl.). New York: Population Council. Also NBER Working Paper No. 12379.

Mason, Andrew, Ronald Lee, An-Chi Tung, Mun-Sim Lai, and Tim Miller. 2009. Population aging and intergenerational transfers: Introducing age into national accounts. In D. Wise (ed.), *Developments in the Economics of Aging*. Chicago: National Bureau of Economic Research and University of Chicago Press (forthcoming).

Mejía-Guevara, Iván. 2009. The role of public sector in the national transfer accounts of Mexico. Paper presented at the Sixth Meeting of the Working Group on Macroeconomic Aspects of Intergenerational Transfers, UC Berkeley, CA, 9–10 January 2009.

SEGOB [Secretaría de Gobernación]. DOF [Diario Oficial de la Federación].2009. *Ley del Impuesto Sobre la Renta*. http://www.ordenjuridico.gob.mx/Federal/Combo/L-78.pdf (accessed July 22th 2009).

———. 2009a. *Ley del Seguro Social*. http://www.ordenjuridico.gob.mx/Federal/Combo/L-88.pdf (accessed July 22 2009).

SHCP [Secretaría de Hacienda y Crédito Público]. 2004. *Distribución del pago de impuestos y recepción del gasto público por deciles de hogares y personas: resultados para el año 2004* [Distribution of the payment of taxes and the reception of public expenditure by deciles of households and individuals: results for 2004]. Mexico City. www.shcp.gob.mx (accessed July 12th 2009).

West, B. T., Berlung, P., and Heeringa, S. G. 2008. A closer Examination of Subpopulation Analysis of Complex Survey Data, The Stata Journal, 8(4), 520-531.