Journal of Economics and Development Studies
September 2015, Vol. 3, No. 3, pp. 37-41
ISSN: 2334-2382 (Print), 2334-2390 (Online)
Copyright © The Author(s). All Rights Reserved.
Published by American Research Institute for Policy Development
DOI: 10.15640/jeds.v3n3a4
URL: http://dx.doi.org/10.15640/jeds.v3n3a4

Using Nonparametric Conditional Distributions To Visualise Low-Income Mobility

Antonio Fernández-Morales[1]

Abstract

This article introduces a graphical tool that can be very useful for income mobility analysis based on panel data. It is inspired in the concept behind transition matrices but uses nonparametric kernel estimation of the conditional distributions of income to allow the continuous treatment of the income variables. The resulting tool is very flexible, since it is possible to apply several degrees of resolution by varying the number of curves plotted and to add reference lines to locate particular initial incomes, like poverty lines. Thus, the methodology of this paper permits obtaining a visual representation of the mobility of the whole distribution that does not depend on particular mobility indexes, and facilitates focusing on specific parts of the distribution, especially within the low-income population.The method is illustrated with an application to four countries from the EU-SILC panel database, two Mediterranean and two Scandinavian, revealing differences in the probabilities of escaping poverty and the mobility in the lower tail of the distribution of income.

Keywords: Income mobility, kernel density estimation, conditional distributions, income distribution

1. Introduction

The increasing availability of income panel data sets is allowing a growing number of applied investigations of income mobility to be carried out. However, certain dynamic aspects of income distribution are not easy to capture with the traditional approach of estimating single mobility indexes. On the other hand, the transition matrix approach attempts to describe the movements and direction of changes in the distribution by dividing the population into classes, and obtaining conditional probabilities of moving to every income class depending on the initial class. However, the discrete nature of the transition matrix is a serious limitation that conditions the results to the number and definition of classes, and loses information, especially when the number of classes is small.

Although challenging, the graphical analysis of income mobility may provide helpful tools to better understand income mobility processes and complement available measures (van Kerm, 2009), and sometimes could be more effective than tabular presentations (Jäntti and Jenkins, 2013). Some recent examples include the transition probability plot (van Kerm 2011), mobility profiles (Jenkins and van Kerm, 2011), the mobility curve (Foster and Rothbaum, 2012) or the mobility probability curve (Cheong and Wu, 2014). Specifically, this paper proposes a graphical instrument that attempts to generate a representation of the mobility process between two periods using a continuous approach. It is inspired in the transition probability colour plot by van Kerm (2011), which is a visual representation of the transition matrix. The methodology proposed in this paper also adopts a squared plot region with both axes (0, 1) denoting initial and final population fractiles, but introduces a set of conditional distribution functions, which are obtained using bivariate kernel density estimations as in Trede (1998).

The paper is structured as follows: the proposed methodology is presented in section 2. Section 3 provides an empirical application to four countries from the EU-SILC panel database. A brief conclusion follows.

[1] Departamento de Economía Aplicada (Estadística y Econometría), University of Málaga, Spain, Calle El Ejido, 6, 29071 Málaga, Spain. Email: afdez@uma.es, Phone: +34 952137189 , Fax: +34 952131279

2. Methodology

The graphical representation proposed in this paper is inspired in the transition matrix in such a way that the vertical coordinates are related to the initial distribution of income, and the horizontal coordinates are related to the conditional distribution of incomes in the final period. Denote the joint distribution function of Y_0 and Y_1 by $F(y_0, y_1)$, where Y_1 represents the income in the final period and Y_0 the income in the initial period. We will assume that the bivariate density of Y_0 and Y_1, $f(y_0, y_1)$, exists. The distribution function of final income conditional on initial income y_0 is

$$F(y_1 \mid y_0) = \int_{-\infty}^{y_1} f(t \mid y_0)\,dt = \int_{-\infty}^{y_1} \frac{f(y_0, t)}{f(y_0)}\,dt \quad . \tag{1}$$

The marginal distribution and density functions of Y_0 are given by

$$F_0(y_0) = \int_{-\infty}^{y_0} \int_{-\infty}^{\infty} f(t, y_1)\,dt\,dy_1 \qquad f_0(y_0) = \int_{-\infty}^{\infty} f(y_0, y_1)\,dy_1 \quad . \tag{2}$$

Letting $p_0=F_0(y_0)$, which is placed on the vertical axis, the inverse $F_0^{-1}(p_0)$ gives the $p_0$quantile of the distribution of initial incomes. For a fixed value of $Y_1= y_1$, the function $F(y_1 \mid F_0^{-1}(p_0))$ gives the probability that Y_1 is less than or equal to y_1 in the final period, conditioned on the $p_0$quantile of the initial distribution of incomes. The values given by this function to every p_0 are used as horizontal coordinates to plot the curve of conditional probabilities corresponding to the chosen y_1.

To generate a relatively simple two-dimensional plot, a finite set of values for Y_1 must be specified in order to plot a curve for every Y_1 in the set. As our focus is on the lower tail of the income distribution, a vector of fractions of the median, Me_1, (e.g. $0.1Me_1$, $0.2Me_1$, …, $0.6Me_1$, …), seems to be an adequate choice. This set includes the standard relative poverty line, 60% of the median income, but also alternative poverty thresholds, such as 40%, 50%, or even 70% of the median, like those provided by Eurostat in their at-risk-of-poverty rates (ARPR).

The functions $F(y_1 \mid y_0)$ and $F_0(y_0)$ are needed to produce the two-dimensional plot described above. Like in Trede (1998), in this paper we will opt for a nonparametric kernel estimation of these functions. Kernel density estimation is widely used in income distribution studies (Jenkins 1995, Trede, 1998, Pittau and Zelli, 2004, 2006, Cheong and Wu, 2014). The approach followed is the same as in Trede (1998), although we introduce sample weights.

Let (y_{0i}, y_{1i}), i=1, 2, …, n, be the sample incomes of n individuals in periods 0 and 1, and $w_1, w_2, \ldots, w_n$ their corresponding weights, which are normalised so they add up to 1. Assuming a multiplicative kernel for simplicity, according to Trede (1998), an estimator for $F(y_1 \mid y_0)$ can be obtained using a bivariate kernel density estimator (weighted in our case) in expression (1):

$$\hat{F}_K(y_1 \mid y_0) = \frac{\sum_{1=1}^{n} w_i K\left(\frac{y_0 - y_{0i}}{h_0}\right) G\left(\frac{y_1 - y_{1i}}{h_1}\right)}{\sum_{1=1}^{n} w_i K\left(\frac{y_0 - y_{0i}}{h_0}\right)} \quad , \tag{3}$$

where the kernel function is denoted by $K(\cdot)$, its cumulative distribution by $G(\cdot)$, $G(x) = \int_{-\infty}^{x} K(t)\,dt$, and the respective bandwidths by h_0 and h_1. In addition, the estimator of $F_0(y_0)$ is based on the denominator of (3). The kernel function chosen for the application in section 3 is the bivariate normal with Silverman's optimal bandwidths.

Finally, the nonparametric estimation of the function $F_0(y_0)$ is numerically inverted to generate the pairs ($F(y_1 \mid F^{-1}_0(p_0))$, p_0) that constitute the curves associated to each value of Y_1. Thus, the proposed graphical instrument consists of a series of curves representing the nonparametric estimations of the conditional distributions of $F(y_1 \mid F_0^{-1}(p_0))$ for a set of fixed values of Y_1.

For a given value y_1, each point of its corresponding curve indicates in the abscissa the estimated probability of a final income not greater than y_1, conditional to an initial income $y_0 = F_0^{-1}(p_0)$, where p_0 is the ordinate. In this type of graphical representation, the extreme situation consisting of a distribution of incomes in the final period independently of the initial distribution of incomes will result in a series of vertical lines. In contrast, the case of complete rigidity, where all the individuals have the same income in the initial and final periods, will generate a set of horizontal discontinuities to vertical segments starting at (0,1) and ending at (1,0). In general, the position, slope and distance between the curves in every region of the graphical representation may provide useful indications about income mobility.

3. An application to four European countries from the EU-SILC panel database

In order to provide an empirical illustration of the proposed graphical tool to visualise low-income mobility, two countries from Southern Europe, Italy and Spain, and two northern ones, Finland and Sweden, have been selected from the EU-SILC database. The analysis carried out in this paper is based on two consecutive years of the longitudinal EU-SILC files (UDB 2009-2 from 01/03/12), establishing the initial period, t=0, in 2008 and the final one, t=1, in 2009. It is worth mentioning that there are important methodological differences among national sources in the EU-SILC database. Particularly, concerning our example, Italian and Spanish data are obtained by means of surveys, while Finnish and Swedish ones come form register sources. Therefore, part of the differences in income mobility found in the results may be attributable to this circumstance.

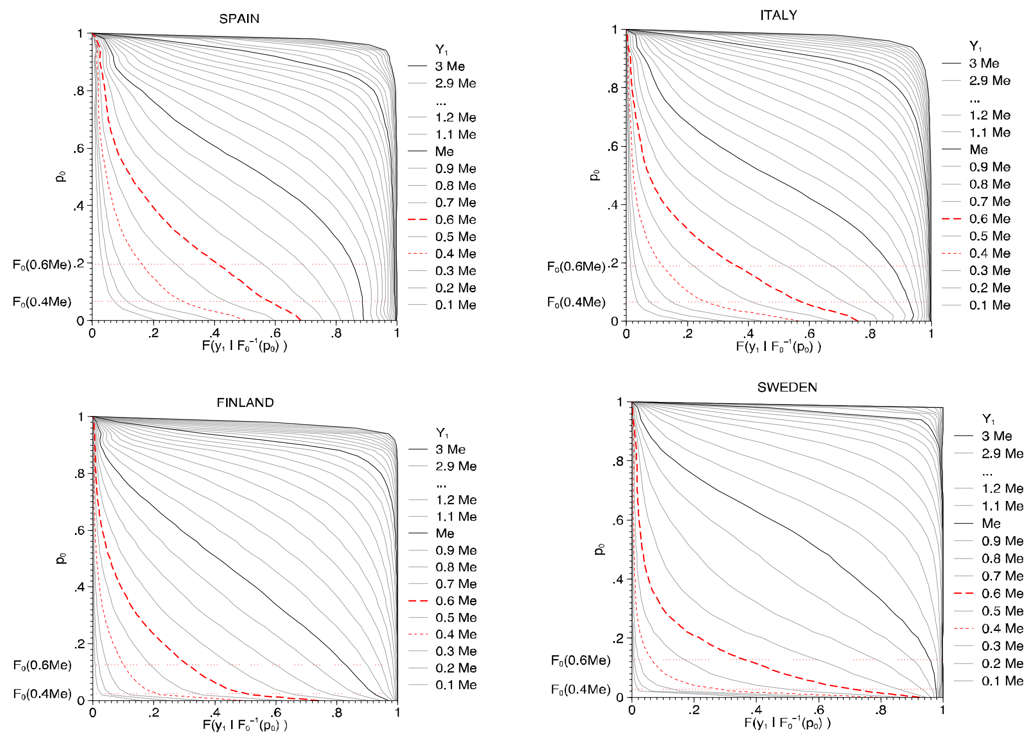
The income variable consists of the equivalised disposable household income, which is defined as the aggregate across all adult household members of all money income receipts during the reference period, with direct tax payments deducted from the total, adjusted with the modified OECD equivalence scale.

The results are shown in Figure 1. A set of 30 values of Y_1 has been defined for each country (0.1 Me_1, 0.2 Me_1, …, 3 Me_1). Slightly bold lines are used to represent the curves corresponding to Me_1, $2Me_1$ and $3Me_1$ in order to facilitate the interpretation. Due to the importance of the 0.6 Me_1 and 0.4 Me_1 poverty lines, their curves are plotted using dashed lines. Moreover, two horizontal reference lines (dotted) have been added to each graphic to indicate the position of the ARPR poverty rates with two different poverty lines, 60% and 40% of the median at the initial period, i.e. $ARPR_0(0.6) = F_0(0.6Me_0)$ and $ARPR_0(0.4)=F_0(0.4Me_0)$.

The image of low-income mobility appearing in Figure 1 shows some interesting differences between the Mediterranean and the northern countries. Firstly, both Spain and Italy have an initial $ARPR_0(0.6)$ near 0.2 (consistent with the results of previous research, like Fernández-Morales *et al.*, 2013), while this indicator is significantly lower for Finland and Sweden at below 0.13 (the same difference occurs for the initial $ARPR_0(0.4)$). Moreover, curves Y_1=0.6 Me_1 and Y_1=0.4 Me_1 appear farther left on the graphic in Finland and Sweden than in Spain and Italy, thus indicating, in general, that the former countries have lower probabilities than the latter countries of being under the corresponding poverty line in the second year for the same initial position in the distribution of income at least for the region over p_0=0.1. However, in the lowest part of the graphic, especially for the initial 5% poorer population, these curves for Finland and Sweden become more horizontal, revealing less mobility and greater probabilities in the case of Sweden of being under the mentioned poverty lines in the second year than the corresponding estimated values for Spain and Italy.

In addition, looking at curve Y_1=0.4 Me_1 we can see that for initial incomes around the standard poverty line, 0.6 Me_0, the greatest probability of falling under the deeper $0.4Me_1$ poverty line in the following year (the abscissa of the intersection between the reference line $F_0(0.6Me_0)$ and the curve $0.4Me_1$) corresponds to Spain, followed by Italy, and the smallest one to Sweden. On the contrary, for those individuals with initial incomes around $0.4Me_0$, the greatest probability of escaping this poverty threshold corresponds to Finland. In this case, however, the smallest probability is found in Sweden, according to the lower degree of mobility observed in the lower percentage of population below the 0.4 Me_0 poverty line in this country.

Figure 1. Conditional distributions of equivalised household income, 2008-2009: Spain, Italy, Finland and Sweden.



4. Conclusion

The graphical instrument introduced in this article, which depicts income mobility between two periods through the estimated conditional distribution of income in the second period, may constitute a very useful tool for the analysis of income mobility. The estimation of the initial and the conditional distribution of incomes is performed using bivariate kernel density estimation, thus allowing the continuous treatment of the variable. The resulting tool is very flexible, since it is possible to apply several degrees of resolution by varying the number of curves (30 curves are plotted in the examples) and to add reference lines to locate particular initial incomes.

The examples that illustrate the use of the proposed graphical tool indicate that, in general, there were higher probabilities of escaping poverty in Finland and Sweden than in Spain and Italy in the observed period (2008-2009), thus implying lower mobility out of poverty in the latter countries. However, when focusing on the very low tail of the income distribution, Sweden exhibits the lowest probability of escaping poverty, which is indicative of a low level of mobility for this part of the distribution. Therefore, the methodology of this paper permits obtaining a visual representation of the mobility of the whole distribution that does not depend on particular mobility indexes, and facilitates focusing on specific parts of the distribution, especially within the low-income population.

Acknowledgement

The author gratefully acknowledges financial support from the *Plan Propio de Investigación para Ayudas a Estancias en Centros de Investigación TIPO 1-A, (Università degli Studi di Roma La Sapienza) of the University of Malaga (Spain).*

References

Cheong, T.S. and Wu, Y. (2014). Convergence and transitional dynamics of China's industrial output: A county-level study using a new framework of distribution dynamics analysis. Discussion paper 14.21, Business School, Univeristy of Western Australia.

Fernández-Morales, A., García-Lizana, A. and Martín-Reyes, G. (2013). Poverty and recession in Euro Area. Revista de Economía Mundial, 33, 153-178.

Foster, J. and Rothbaum, J. (2012). Mobility curves: using cut-offs to measure absolute mobility. George Washington University, mimeo.

Jäntti, M., Jenkins, S. P. (2013). Income mobility. SOEP papers on Multidisciplinary Panel Data Research, no. 607.

Jenkins, S. P. (1995). Did the middle class shrink during the 1980s? UK evidence from kernel density estimates.Economics Letters 49, 407-413.

Jenkins, S. P., Van Kerm, P. (2011). Trends in individual income growth: measurement methods and British evidence. ISER Working Paper Series 2011-06, Institute for Social and Economic Research, University of Essex, Colchester.

Pittau, M.G. and Zelli, R. (2004). Testing for changing shapes of income distribution: Italian evidence in the 1990s from kernel density estimates. Empirical Economics 29, 415-430.

Pittau, M.G., Zelli, R. (2006). Trends in income distribution in Italy: A non-parametric and semi-parametric analysis. Journal of Income Distribution 15, 90-118.

Trede, M. (1998). Making mobility visible: a graphical device. Economics Letters 59, 77–82.

Van Kerm, P. (2009). Income mobility profiles.Economics Letters 102, 92-95.

Van Kerm, P. (2011). Picturing mobility: Transition probability colour plots. United Kingdom Stata Users' Group Meetings 2011, 18, Stata Users Group.