

INSTITUTO DE ESTUDIOS LABORALES Y DEL DESARROLLO ECONÓMICO (ielde)
Facultad de Ciencias Económicas, Jurídicas y Sociales
Universidad Nacional de Salta (UNSa)
Salta
Argentina

Documentos de Trabajo

Educational Inequality in Argentina: The best and worst

Melisa Morales
Corina Paz Terán

Junio de 2010
N° 5

ielde – Facultad de Ciencias Económicas, Jurídicas y Sociales - UNSa

<http://www.economicas.unsa.edu.ar/ielde>

UNSa: Av. Bolivia 5150, A4408FVY, Salta, Argentina

ISSN 1852-1118 (impreso), ISSN 1852-1223 (en línea)

Editor: Jorge A. Paz ielde@unsa.edu.ar

Educational Inequality in Argentina: The best and worst performers ¹

Melisa Morales
mmorales@iadb.org
Inter-American Development Bank

Corina Maria Paz Teran
cpaz@herrera.unt.edu.ar
Universidad Nacional de Tucumán

Abstract

What do we know about inequality in educational attainment across Argentina's cities? To answer this question, we present the education Gini coefficient for the period 2002-2007. Using microdata from the national household survey, we document the following results. First, educational inequality has declined in almost all metropolitan areas whereas it has increased in Posadas, Mar del Plata, Rosario and Formosa. Second, although there are no important differences in the average years of schooling across cities, great disparities exist with respect to the education Gini. Buenos Aires City is in a leading position, especially in relation to the northeast region of the country and, particularly, Posadas city.

Resumen

¿Qué sabemos sobre la desigualdad educativa en los aglomerados urbanos de Argentina? Para responder a la pregunta presentamos el Gini de educación para el periodo 2002-2007, usando datos de la Encuesta Permanente de Hogares (EPH). Los principales resultados son: la desigualdad educativa se ha reducido en casi todas las ciudades con excepción de Posadas, Mar del Plata, Rosario y Formosa en donde la desigualdad educativa ha aumentado. A pesar de que no existen importantes diferencias en los años promedio de educación a través de las ciudades, altas disparidades existen en lo que respecta al Gini de educación. La Ciudad de Buenos Aires presenta la distribución educativa más igualitaria del país. Su posición ventajosa es significativa principalmente respecto del Noreste del país, y en particular la ciudad de Posadas.

Keywords: C43, D3, J24

JEL: Gini, inequality, bootstrap

¹ Melisa Morales is Research Fellow at the Inter-American Development Bank and Corina Paz is Professor and Researcher at Universidad Nacional de Tucuman. They are grateful to Cecilia Adrogué, Suzanne Duryea, Victor J. Elias, Florencia Lopez Boo, Jorge Paz, Raul O. Sanchez, two anonymous referees, and participants at the XLIV Meeting of the Argentine Association of Political Economy (Mendoza, November 2009) for helpful comments to a related version of this paper. In particular, the authors warmly thank Victor J. Elias for his continuous and generous support. The views in this paper are solely the responsibility of the authors. Comments are very welcome.

1 Introduction

We propose to measure the level of inequality in educational attainment across Argentine cities. To this end, we focus on the widely used measure of inequality known as the Gini coefficient. The Gini coefficient allows us to study the distribution of a given variable. In this paper, we analyze the distribution of the variable years of schooling that we have chosen as a proxy of educational attainment.

It is worth mentioning that the education Gini we present here does not control for quality of education. Using microdata from Argentina's Permanent Household Survey (EPH), our analysis covers the total of metropolitan areas in the survey for the period 2002-2007. Basically, we examine the years 2002 and 2007 and the significance of the changes observed in the education Gini for that period. As we will explain later, the two years selected relate well with phases of crisis and recovery in Argentina's business cycle. Also, the dataset for 2007 is the latest available for the EPH survey as of July 2009.

The results are striking. Although differences in average years of schooling are not significantly great across metropolitan areas, important disparities remain with respect to the education Gini.² If we were to establish a ranking across metropolitan areas for 2007, first place would go to Buenos Aires City, which has the most egalitarian distribution of education. Localities that rank lowest are Posadas in northeast, Tucuman city in the northwest, Viedma in the Patagonia region, and San Luis in Cuyo.

Gasparini (2007), also using the Permanent Household Survey, found that "the proportion of highly educated people has significantly increased during the last 14 years in Argentina. While in 1992 17.8 percent of adults aged 25 to 65 had more than 13 years of formal education, that share increased to 21.3 percent in 1998 and to 24.7 percent in 2003, and to 27 percent in [2006]". He also has noted that the education Gini for the country has fallen slightly during the last 14 years.

Accordingly we would expect a welcomed fact: the decline in the education Gini. Regardless of the situation, we believe the analysis here provides a useful framework for exploring the inequality in educational attainment within Argentina. This paper may be the first attempt to measure the education Gini across the different cities of Argentina.

² It is worth noting that the differences in average years of schooling across cities are not significantly great if we exclude the case of Buenos Aires City.

The remainder of this paper is as follows. The next section describes the education Gini coefficient. In Section 3, we present the data used, focusing particularly on the variable years of schooling. Section 4 discusses our findings and Section 5 provides concluding remarks.

2 What is the education Gini?

The Gini coefficient is well established as a popular inequality measure; it has been extensively used to analyze income inequality. The coefficient allows us to measure the dispersion of a distribution, ranging from 0 (complete equality) to 1 (complete inequality). In other words, a larger Gini implies greater inequality. As mentioned above, our goal is to analyze the distribution of years of schooling by means of the education Gini. We have used years of schooling as a proxy of educational attainment.³

Based on equation (1) from Deaton (1997), the education Gini coefficient adopted here is in (2). Deaton (1997) defined income Gini as "the ratio to the mean of half of the average overall pair of the absolute deviations between [all possible pairs of] people."

$$G = \frac{1}{\bar{y}N(N-1)} \sum_{i>j} \sum_j |y_i - y_j| \quad (1)$$

where G is the Gini coefficient; N is the total number of observations; \bar{y} is the mean of income; and y_i and y_j are dollar values of income of individuals.⁴

In this sense, the education Gini formula is as follows:

$$G(s) = \frac{1}{\bar{s}N(N-1)} \sum_{i>j} \sum_j |s_i - s_j| \quad (2)$$

where $G(s)$ is the education Gini coefficient; \bar{s} is the mean of the variable years of schooling; N is the total number of observations; s_i and s_j are years of schooling of individuals.

As regards the interpretation of the education Gini coefficient, Thomas et al. (2001) are very clear: "As an analogue to Deaton's definition, education Gini measures the ratio to the mean (average years of schooling) of half of the average schooling deviations between all possible pairs of people."

³ Extensive literature on Gini coefficient can be found. See, for example, Chakravarty (1990) for a comprehensive survey on the Gini and other measures of inequality.

⁴ Note that there are $\frac{N(N-1)}{2}$ distinct pairs in all.

Worldwide, there are few papers that analyze the education Gini. Most have studied the coefficient across countries and followed the equation (3).⁵

$$G(s) = \frac{1}{\bar{s}} \sum_{i=2}^n \sum_{j=1}^{i-1} p_i |s_i - s_j| p_j \quad (3)$$

, where $G(s)$ is the education Gini; \bar{s} is the average years of schooling of the population; s_i and s_j are the years of schooling at different educational attainment levels; n is the number of levels in attainment data; p_i and p_j are the proportions of population with those levels. Seven levels of education usually are measured (illiterate, partial primary, complete primary, partial secondary, complete secondary, partial tertiary, and complete tertiary).⁶

We argue that the one chosen in the current paper -- equation (1) -- is more accurate for measuring education Gini when microdata are available. Some precision is lost in working with aggregate data because the reduced number of categories of level of education variable tends to underestimate the results.⁷

As far as we know, only Gasparini (2007) has estimated the education Gini coefficient for Argentina using microdata.⁸ As mentioned in the introduction, he found that the coefficient has slightly decreased in the last 14 years. Specifically, education Gini for 15 cities in 1995 was around .235; .233 for 28 cities in 1998; and the coefficient fell from .219 in 2003 to .215 in 2006.

3 The Data

The data used come from the Permanent Household Survey (*Encuesta Permanente de Hogares*, EPH), which allows a complete study of the education Gini for the whole population of the country. The survey is the main household survey in Argentina, and it is developed by Argentina's National Institute of Statistics and Censuses (INDEC). Currently, the EPH covers only metropolitan areas in six regions.⁹

⁵ See Thomas et al. (2001) for a detailed review of the previous few studies that analyzed the education Gini using enrollment, financing, and attainment data. See also, for example, Thomas et al. (2002) and Qian and Smyth (2008).

⁶ Thomas et al. (2002) show the negative relationship between average years of schooling and the education Gini using equation (3).

⁷ A lower number of categories will imply higher underestimation in the results.

⁸ Gasparini (2002) also calculated the Gini coefficient of the distribution of probabilities of attending high school in Greater Buenos Aires during the period 1980-2000, albeit in that paper his approach was to measure unfairness in school attendance instead of inequality. See Gasparini (2002) for a comprehensive discussion about the difference between unfairness and inequality.

⁹ The six regions established in the survey are: northwest (NOA), northeast (NEA), Cuyo,

The education Gini was obtained for all the metropolitan areas of the EPH for the years 2002 and 2007. The first year represents the last crisis phase of the business cycle of the country, while 2007 is related to the last phase of economic recovery. Changes in education Gini coefficients across areas will be reported for this quinquennium in the next section along with the bootstrapped standard errors computed with 100 replications.

Our study includes 83,313 observations in 2002 across the whole country, and it covers data referred to May wave. The sample in 2007 includes 63,438 observations and the data corresponds to the first quarter of the year, which is the last dataset published by INDEC. The survey covers 29 and 32 metropolitan areas in 2002 and 2007, respectively.¹⁰ For 2002 the EPH used responses to the version called punctual EPH, which offers data for the waves of May and October. It is a previous version with respect to the second one, 2007-- which represents the last version published by INDEC, called the continuous EPH.¹¹

Since the data belong to two different versions of the survey, the equivalence of definitions must be established to compare the same variables. For our results, we used a sample of individuals aged 25-65.¹² To generate the years of schooling variable, we have used the questions "Do you attend school currently?" and "What is the last year of formal education you have completed?" We truncated years of schooling at 17 years because the survey in 2002 does not capture years of graduate education. Our reason for choosing this procedure, instead of using the variable level of education reported by INDEC, was given in Section 2.

We used the 2007 data, which has mentioned previously, was the latest available when writing this document. We selected the 2002 dataset mainly because international evidence that examined the education Gini have traditionally provided the information for quinquennium (Thomas et al., 2001 and 2002; Zhang et al., 2002). The 2002 dataset has exactly the same variables to construct years of schooling as the 2007 dataset, however the 2002 dataset responds to the "punctual" version. We did not use the 2003 dataset because the 2003 data lack these variables so that we would not be able to obtain years of schooling.

Pampeana, Patagonia and Greater Buenos Aires (GBA).

¹⁰ For 2007 we were able to get the education Gini for three additional areas: San Nicolas-V.Constitucion (Pampeana region), Rawson-Trelew, and Viedma-Carmen de Patagones (both in Patagonia region).

¹¹ The methodological change from punctual to continuous occurred in 2003 when INDEC started providing four quarterly datasets, and two semester datasets. See, www.indec.gov.ar, for detailed information on EPH.

¹² Weighted population was used.

4 Results

In Table 1 the education Gini coefficient is reported along with the standard errors. Standard errors were estimated by bootstrapping techniques, which provide interval estimations for the $G(s)$.¹³ As expected, a negative relationship between education Gini and years of schooling is found (Figure 1, based on Table 1).¹⁴ An implication of this finding is that, by increasing the length of schooling, a more egalitarian education distribution is reached. All the metropolitan areas increased the average years of schooling of the economically active population for the period under study. However, not all these cities had the same performance during the period. While San Luis increased the average years of schooling from 9.4 in 2002 to 10.5 in 2007 (11.3 percent), Tucuman did the same by 8.5 percent, Concordia by 3.5 percent, and Posadas by 3.2 percent. This last city is one of the four cities that have increased the education Gini, along with Mar del Plata, Rosario, and Formosa. In other words, these four cities were not able to improve their distribution of educational attainment.

In addition, Formosa had the lowest increase in average years of schooling in Argentina (from around 9.6 in 2002 to 9.7 years in 2007, or 1.57 percent). On the other hand, Buenos Aires city exhibited the lowest education Gini as well as the highest average years of schooling. Note that the Greater Buenos Aires (GBA) includes only one metropolitan area but the survey provides two measures separately for Buenos Aires City and the rest of GBA.¹⁵ Clearly, a remarkable difference exists between Buenos Aires City and the rest of GBA, with $G(s)$ equals to .1483 and .2138, respectively.

¹³ The bootstrapped standard errors were computed with 100 replications. The implementation of bootstrap followed Mills and Zandvakili (1997); Gasparini and Sosa Escudero (2000).

¹⁴ The correlation coefficient is -0.8419 (2007).

¹⁵ Buenos Aires City refers to *Ciudad Autonoma de Buenos Aires* and the rest of GBA to *Partidos del Gran Buenos Aires*.

Table 1. Education Gini, average years of schooling and bootstrapped standard errors, 2002 and 2007

<i>Metropolitan Areas</i>	2002			2007			%	
	s	G(s)	s.e.	s	G(s)	s.e.	s	G(s)
BUENOS AIRES								
Ciudad de Buenos Aires	12.7	0.1631	0.0037	13.2	0.1483	0.0044	3.7	-9.1
Partidos del Conurbano	9.4	0.2228	0.0024	9.8	0.2138	0.0027	4.4	-4.1
PAMPEANA								
Gran Parana	10.3	0.2092	0.0049	11.1	0.1861	0.0060	7.3	-11.1
Gran La Plata	10.9	0.2175	0.0051	11.8	0.1935	0.0058	8.5	-11.0
Gran Cordoba	10.6	0.2234	0.0047	11.6	0.1956	0.0052	9.4	-12.4
Bahia Blanca-Cerri	10.3	0.2073	0.0047	10.8	0.1988	0.0062	4.1	-4.1
Gran Santa Fe	10.3	0.2222	0.0046	10.9	0.2079	0.0064	6.2	-6.5
Mar del Plata-Batan	10.2	0.2020	0.0048	10.6	0.2130	0.0069	4.6	5.4
Gran Rosario	10.0	0.2183	0.0045	10.6	0.2191	0.0054	5.9	0.4
Santa Rosa-Toay	9.9	0.2410	0.0058	10.5	0.2192	0.0068	6.3	-9.0
San Nicolas-V.Constitucion				9.9	0.2230	0.0058		
Rio Cuarto	9.8	0.2449	0.0060	10.4	0.2395	0.0066	5.6	-2.2
Concordia	9.0	0.2578	0.0059	9.3	0.2542	0.0080	3.5	-1.4
CUYO								
Gran San Juan	10.1	0.2200	0.0041	10.7	0.2049	0.0045	6.2	-6.9
Gran Mendoza	10.3	0.2236	0.0039	11.1	0.2097	0.0055	7.5	-6.2
San Luis-El Chorrillo	9.4	0.2406	0.0061	10.5	0.2119	0.0058	11.3	-11.9
NEA								
Corrientes	10.3	0.2339	0.0062	11.1	0.1966	0.0057	7.8	-15.9
Formosa	9.6	0.2308	0.0053	9.7	0.2393	0.0067	1.0	3.7
Gran Resistencia	9.3	0.2677	0.0070	10.1	0.2456	0.0073	8.6	-8.3
Posadas	9.3	0.2456	0.0060	9.6	0.2563	0.0099	3.2	4.4
NOA								
La Rioja	9.9	0.2099	0.0042	10.7	0.1956	0.0046	8.1	-6.8
Salta	10.4	0.2140	0.0045	11.1	0.1978	0.0055	6.7	-7.6
Sgo del Estero-La Banda	10.0	0.2197	0.0057	10.5	0.2008	0.0046	5.0	-8.6
Gran Catamarca	10.2	0.2181	0.0051	10.8	0.2042	0.0054	5.9	-6.4
Jujuy-Palpala	10.2	0.2221	0.0055	10.7	0.2148	0.0070	4.9	-3.3
Gran Tucuman-Tafi Viejo	9.9	0.2482	0.0052	10.7	0.2239	0.0048	8.1	-9.8
PATAGONIA								
Ushuaia Rio Grande	10.7	0.1931	0.0042	10.9	0.1829	0.0059	2.5	-5.3
Rio Gallegos	10.2	0.2009	0.0038	10.9	0.1874	0.0055	6.7	-6.7
Comodoro Rivadavia-R. Tilly	9.8	0.2200	0.0049	10.2	0.1950	0.0059	4.5	-11.4
Neuquen-Plottier	10.1	0.2361	0.0060	10.9	0.2063	0.0077	8.7	-12.6
Rawson-Trelew				9.8	0.2304	0.0068		
Viedma-C. de Patagones				9.9	0.2443	0.0074		

Source: Author's based on EHP Survey.

Figure 1. EDUCATION GINI COEFFICIENT AND AVERAGE YEARS OF SCHOOLING

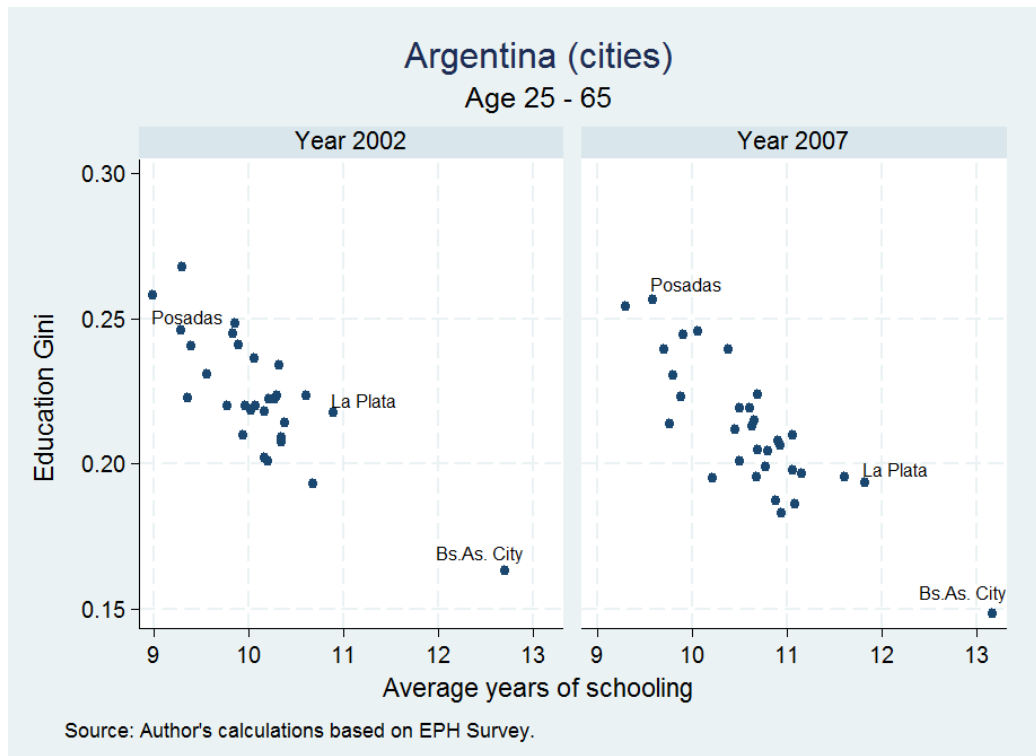


Figure 2. EDUCATION GINI COEFFICIENT, 2007

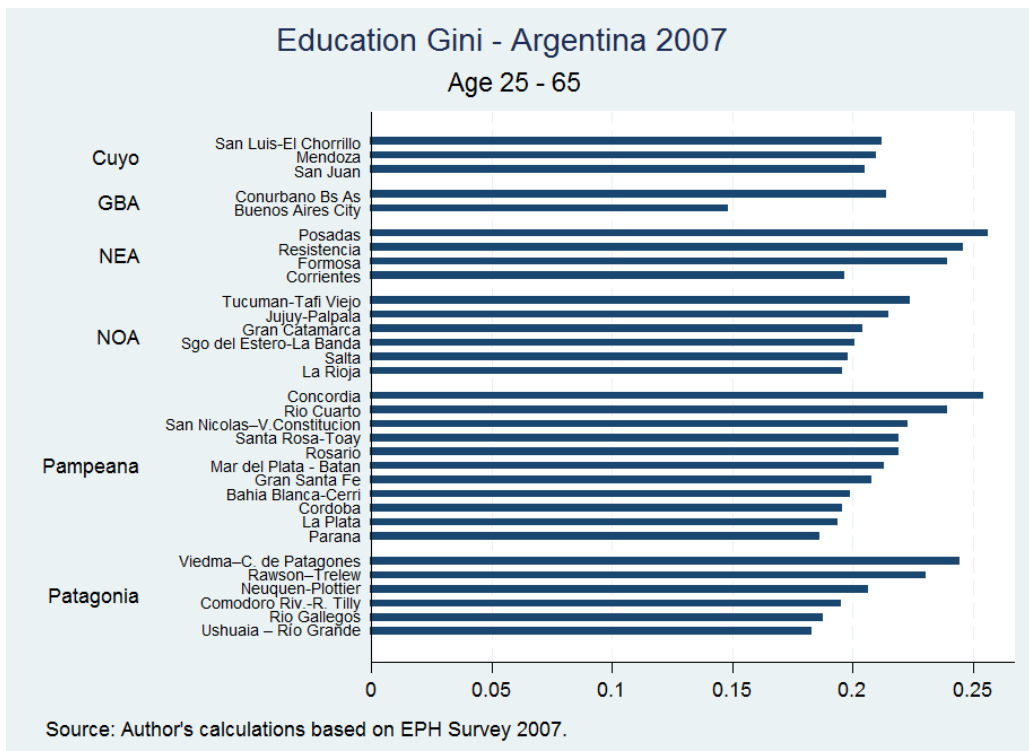
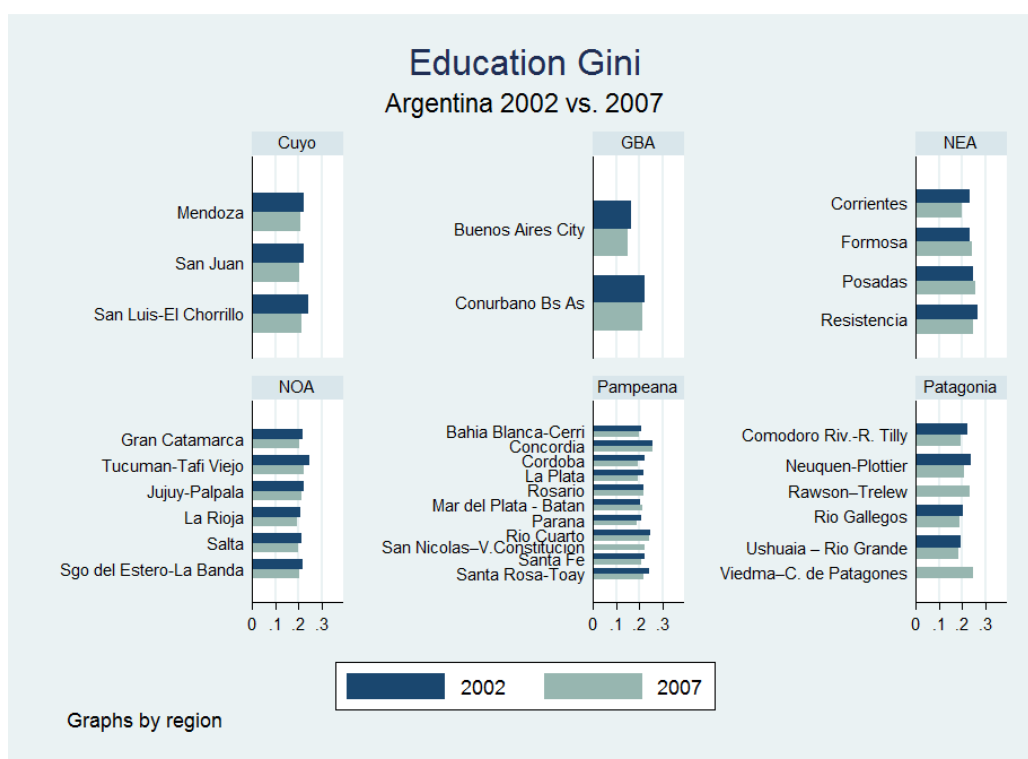


Figure 2 depicts the education Gini coefficient for 2007 by metropolitan area, and Figure 3 compares the education Gini for 2002 and 2007 in the six regions of the country. The cities

with the least egalitarian distribution of education in 2007 were San Luis in Cuyo, Tucuman in NOA, Concordia in the Pampeana region, Viedma in Patagonia, and Posadas in the northeast region of the country, which also showed the highest education Gini in the country. On the other hand, Cuyo showed the least dispersion among its three metropolitan areas. Rio Cuarto, in the Pampeana region, was one of the five cities with the least egalitarian education distribution with a $G(s) = .2395$, while Cordoba exhibited an education Gini equal to .1956. Both cities belong to the same province, Cordoba, and had made different progress in lessening education inequality. While Cordoba city increased the average years of schooling from 10.6 to 11.6 years (9.4 percent), Rio Cuarto moved from 9.8 to 10.4 years (5.6 percent).

Figure 3. EDUCATION GINI COEFFICIENT, 2002 vs. 2007



5 Concluding Remarks

The Gini coefficient is well established as a popular inequality measure which it has not been widely used to analyze education inequality. Here, we have focused on the education Gini for Argentina to explore the inequality in educational attainment across Argentina's cities.

Using microdata from Argentina's Permanent Household Survey (EPH), our analysis covered the total of metropolitan areas for the period 2002 - 2007. We constructed the variable years of schooling instead of using the variable level of education given by INDEC; some precision is

lost while working with aggregate data because the reduced number of categories of the variable tends to underestimate the results. As mentioned previously, most of the studies on education Gini carried out worldwide are based on education level with the consequent restrictions. Therefore, we consider the variable years of education or schooling the more appropriate and accurate way to measure $G(s)$.

The results suggest that Buenos Aires City shows the best performance in education Gini. The city, which constitutes just 9.1 percent of Argentina's total population, has the most egalitarian distribution of education of the whole country. A remarkable difference remains between Buenos Aires City and the rest of the regions if average years of schooling and education Gini coefficient are taken into account. Much effort directed toward the rest of the country will be needed to bring it up to the level of Buenos Aires City. The worst levels of educational inequality are in Posadas in the northeast (NEA), Tucuman city in the northwest (NOA), Viedma in Patagonia, and San Luis in Cuyo. Posadas not only holds the worst position in NEA but also is one of the cities that increased education inequality in the period under study. Posadas, Mar del Plata, Rosario, and Formosa increased the education Gini, thus showing the worst performance in lessening education inequality.

We believe that the analysis presented here provides a useful framework for exploring inequality in educational attainment within Argentina. Further research could embrace, for instance, the analysis of the demographic structure across cities to explore in more detail to what extent the average year of schooling is influenced by a given demographic structure. In addition, future research could complement the analysis by exploring the educational progress across generations. We hope this study may help to guide future policy-making.

5 References

- Chakravarty, R. S. (1990): *Ethical Social Index Numbers*, New York: Springer-Verlag.
- Deaton, A. (1997): *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*, the World Bank and Johns Hopkins University Press.
- Gasparini, L. (2002): On the measurement of unfairness, An application to high school attendance in Argentina, *Social Choice and Welfare*, 19, pp. 795-810.
- Gasparini, L. (2007): *Monitoring the Socio-Economic Conditions in Argentina 1992-2006*, www.depeco.econo.unlp.edu.ar/cedlas/monitoreo/pdfs/argentina.pdf
- Gasparini, L. and Sosa Escudero, W. (2000): A note on the Statistical Significance of Changes in Inequality. *Económica* 46 (1), La Plata, pp. 111-122.
- Mills, J. and Zandvakili, S. (1997): Statistical Inference Via Bootstrapping for Measures of Inequality, *Journal of Applied Econometrics*, Vol. 12, No. 2, pp. 133-150
- Qian, X., and Smyth, R. (2008): Measuring Regional Inequality of Education in China: Widening coast-inland gap or widening rural-urban gap?, *Journal of International Development* 20, pp. 132-144.
- Thomas, V., Wang Y., and Fan, X. (2001): *Measuring Education Inequality: Gini coefficients of education*, Policy Research Working Paper 2525, The World Bank: Washington, D.C.
- Thomas, V., Wang Y., and Fan, X. (2002): *A New Dataset on Inequality in Education: Gini and Theil Indices of Schooling for 140 Countries, 1960-2000*, mimeo.
- Zhang, J. and Tianyou, L. (2002): International Inequality and Convergence in Educational Attainment, 1960-1990, *Review of Development Economics* 6(3): 383-392.