Formal employment and upward mobility of young adults: Experience of a developing country

Mauricio Quiñones[a,*], Christian Posso[b], Nicolas Mancera[c], Juan C. Duque[a], Carlos Medina[b,1]

[a]*Research in Spatial Economics - RiSE, EAFIT University*
[b]*Banco de la Republica, the Central Bank of Colombia - Medellin Branch*
[c]*M.Sc Applied Economics (C) EAFIT University*

Abstract

The cross-sectional estimates of inequality tend to describe an incomplete picture of individuals' income progress. Therefore, we study intragenerational income mobility to add a dynamic perspective to the inequality analysis by integrating mobility and inequality in one measure. This study uses two sources of administrative records to analyze income mobility for young adults in Colombia. We apply the concept of mobility as the equalization of longer-term incomes, explore associations, and quantify the dimensions' contributions. We find formally employed young adults experience income mobility with high inequality (disequalizing mobility). The results by gender and metropolitan area benefit females and the two main metropolitan areas. We find public education and more fiscal independence are positively associated with more equalizing mobility, and negatively associated with poverty. For dimensions, the largest effect on increasing disequalization is due to individual characteristics, while the effect of utilities and firms reduce disequalizing mobility.

Keywords: Intragenerational mobility, Inequality, Income mobility, Latin America, Regional Policy

*Corresponding author. Research in Spatial Economics - RiSE, EAFIT University. Address: Carrera 48 A 10 Sur 107, House 4, 2nd Floor, Medellin, Colombia.

Email addresses: mquinonesd@eafit.edu.co; maquido@gmail.com (Mauricio Quiñones), cpossosu@banrep.gov.co (Christian Posso), jnmanceraa@eafit.edu.co (Nicolas Mancera), jduquec1@eafit.edu.co (Juan C. Duque), cmedindu@banrep.gov.co (Carlos Medina)

[1]The opinions contained in this document are the sole responsibility of the authors and do not reflect the opinion of the Banco de Republica or its board of directors.

JEL Classification: D63, J62, N36, R23, R58

1. Introduction

Income inequality is identified as a structural phenomenon of the Latin America (LAT) region (Gasparini & Lustig, 2011). The LAT region has historically observed high levels of the Gini coefficient. In 2005, the LAT region was the most unequal region, followed by the regions of Sub-Saharan Africa and East Asia (Cuesta & Negre, 2016). Although, income inequality has decreased in recent years – the Gini coefficient for the region reduced from 0.50 in 2008 to 0.46 in 2017 (Azevedo, 2011) - LAT countries are still among those with high levels of inequality (see Figure 1).

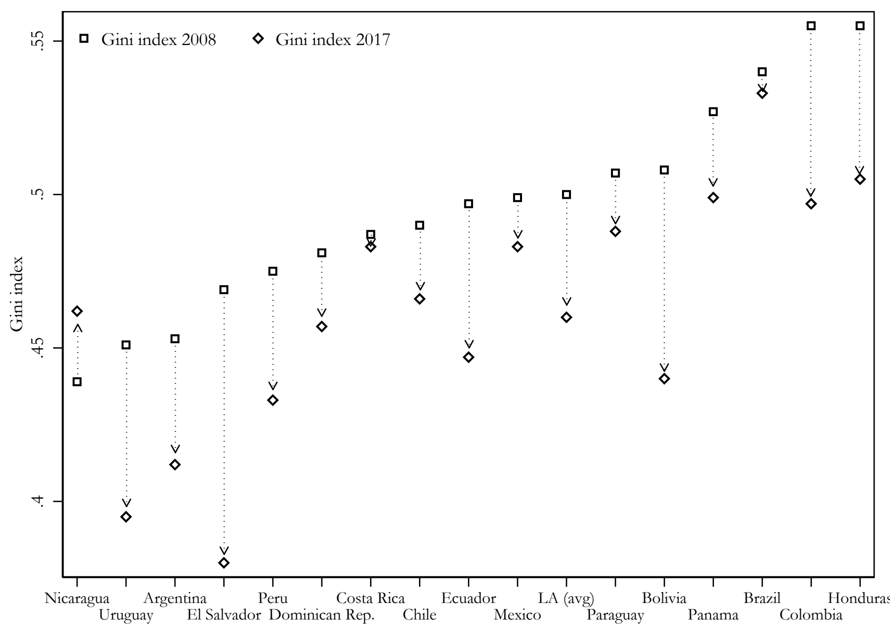

Figure 1: Gini changes, Latin American countries 2008-2017.
Source: Azevedo (2011). Note: For Chile, uses data from 2009; for Dominican Republic, uses data from 2016.

Although income inequality has been extensively studied, these provide an incomplete picture of longer-term economic well-being. On the other hand, income mobility measured as the degree of transmission of inequality across individuals (intragenerational mobility) or generations (intergenerational mobility) allows us to understand the dynamic perspective of economic development. However, income mobility analysis requires access to long-term panel data, which is not very common in the LAT region.[2]

The studies on intergenerational mobility have circumvented this challenge by creating pseudo-panels and focusing on measures to capture the transfer of inequality and opportunities from parents to children (Gasparini et al., 2014; Lambert et al., 2014; Luke & Munshi, 2011; Cuesta et al., 2011). The β (beta) coefficient is the most common measure used to assess intergenerational mobility based on education (Azevedo & Bouillon, 2010; Jantti & Jenkins, 2015). Essentially, in a linear regression model, β quantifies how much parents' education influences children's education. The two extreme cases are either β equals one, i.e., total immobility by dependence, or β equals zero, i.e., total mobility by independence. For instance, based on Neidhöfer, Serrano & Gasparini's estimations (2018) for a 1960 cohort, the top five countries with low levels of mobility are - with β in parenthesis – Nicaragua (0.83), Guatemala (0.80), Ecuador (0.70), Colombia (0.64), and Peru (0.61). The authors also find that for more recent cohorts there is an improvement; nonetheless, there are substantial disparities between and within countries.

Intragenerational mobility studies how individuals' income evolves along their life cycle. Unlike mobility between generations, intragenerational mobility is more feasible for constructing consistent panel data for income because it is collecting data for the same individual instead of for two different persons such as parents and children. Intragenerational mobility estimates also allow the proposal of policies to influence the current income distribution and the observation of the resulting effects in the medium run due to the boundary time span.

[2] Although the topic of opportunities links with that of mobility, this paper does not address it. For a complete discussion on the topic of opportunities see Roemer & Trannoy (2015).

This feature is specially useful for policymakers who can focus on boosting the evenness of the distribution in life-cycle income rather than focusing exclusively on having an annual impact on inequality (Maasoumi & Zandvakili, 1989).

In this paper, we contribute to the empirical literature on intragenerational income mobility by applying a class of measures of mobility that is grounded in the concept of mobility as the equalization of longer-term incomes. In particular, we quantify the E index, proposed for Fields (2010). Although there are several mobility measures in the literature based on different concepts (e.g. mobility as temporal independence using the β or positional movement using the Shorrocks/Prais index), utilizing the concept of the equalization of longer-term incomes makes it possible to simultaneously capture not only mobility but also changes in the distribution's equalization relative to base-year incomes (Fields, 2010). We ground our analysis in the E index because unlike other mobility indices, it captures if the mobility equalizes or disequalizes durign the period of study. In other words, it indicates if the mobility is more or less unequal relative to the base-year. Thus, it is possible to take advantage of the existence of a synthetic index to monitor the evolution of mobility and inequality at the same time to influence social policy in the medium run.

Our study aims to provide a comprehensive analysis of the observed trends regarding intragenerational mobility in Colombia, one of the countries with the highest levels of inequality in the LAT region and the world. We focus on formal employees between 25 and 30 years old between 2009 and 2018. Colombia's administrative records allow us to follow individuals on a monthly basis, which is a key characteristic for quantifying the E index. In addition, we focus our analysis on young adults because we aim to find the characteristics associated with successful mobility outcomes for individuals in the early stage of the labor life cycle (Weller, 2012).

We start by providing analyses for different consecutive cohorts and for the twenty-three main metropolitan areas for a comparative analysis (Birchenall, 2001). Then, we analyze the association between intragenerational mobility measured by the E index and some macroeconomic and microeconomic characteristics identified in the literature. From the macroeconomic perspective, using a panel regression on metropolitan areas, we explore potential associations with measures of employment, socioeconomic conditions (health, education and poverty) and fiscal independence. This analysis allows us to identify the direction of the association to hypothesize likely causal mechanisms, although this is not the focus of our study. From the microeconomic perspective, we introduce a decomposition method using the recentered influence function (RIF) regression (Firpo et al., 2009) to decompose the contribution of individual, household, living conditions, and firm characteristics on our mobility estimates. Other studies have used RIF regression to measure inequality in the LAT region (Fernández & Messina, 2018). This decomposition allows us to assess the effect of any time-invariant observable characteristic on our E index estimates.

We use Colombian administrative records from PILA[3] and process a panel data set with a final sample of over 1.3 million records of formally employed young adults in Colombia between 2009 and 2018. For each individual in our sample, we know his or her earnings at both the beginning and end of the period of analysis. However, PILA does not have socioeconomic variables beyond gender and age. To add more observable variables and decompose the E index estimates, we link the PILA data to data from SISBEN-3.[4] SISBEN-3 is a census of the low-income population that was collected in 2009, which allows the Colombian government to implement social safety net programs. This census provides detailed demographic information such as age, sex, municipality, living conditions, education level, family and household characteristics. While PILA provides information on earnings, SISBEN-3 provides information on individual and household information of low-income families. The case of Colombia can provide a good idea of the social mobility challenges that other LAT countries

[3]Administrative records for social security contributions or Planilla Integrada de Liquidacion de Aportes in Spanish.

[4]System for the selection of beneficiaries for social programs, third version, or Sistema de Seleccion de Beneficiarios para Programas Sociales, tercera version in Spanish.

face.

In most cases, mobility is not explicitly considered into governments' policy programs as inequality is, mainly because it is assumed as the logical outcome of other policies; therefore, research on the topic will offer more evidence to influence governments' decisions. Our findings describe an interesting picture based on the novel aspect of income mobility for a developing country. The negative values for Colombia's E index estimates indicate that the country's young adults experience a process of disequalizing income mobility, which means upward income mobility but with high inequality. We found that some degree of mobility is accompanied by rising inequality, as was found by Fields et al. (2007). In addition, the female population benefits from this mobility more than the male population. We also find that the E index is not significantly associated with the economic complexity index (Hidalgo & Hausmann, 2009) or a health measurement. Nonetheless, it is significantly associated with education, fiscal independence, and poverty. The expansion of public education and more fiscal independence are positively associated with more equalizing mobility. On the other hand, a reduction in poverty is associated with more equalizing mobility. As far as the decomposition with the RIF regression, the largest net effect on disequalization is due to individual characteristics. Conversely, the net effect of utilities and firms reduce disequalizing mobility.

Following this section, we delve into the concepts of income mobility and their applications for the LAT region and Colombia (section 2). In section 3, we present statistics of the LAT region and Colombia between 2008 and 2018 to contextualize their social and economic performances. We also describe the features of the data set and provide summary statistics of the population and periods of analysis. In section 4, we present the main results about the E index that include comparisons among metropolitan areas by cohorts and periods of analysis. We also present the results of exploring the associations of the E index with other relevant topics. In section 5, we apply the decomposition technique to identify the determinants. In section 6, we present the paper's main conclusions.

2. Literature Review

2.1. Mobility of what?: concepts and indices

Although the general public has an idea of what mobility means, a unified definition does not exist. Because different interpretations of mobility will lead to different conclusions, the standard practice for mobility analysis demands being explicit about which concept the research will use. Next, we introduce the most often applied six concepts of mobility in the literature, and we emphasize the concept suitable for our analysis. These six concepts are temporal independence, positional movement, share movement, non-directional income movement, directional income movement, and equalization of longer-term incomes (Fields, 2008).

Temporal independence focuses on quantifying the effect of parents' income on children's income. The most common indicators used to address it are the β (beta) coefficient or slope coefficient obtained from the least squared linear regression of the logarithm of income at time 2 on the logarithm of income at time 1 and the r or Pearson correlation index obtained between the logarithm of income in two periods. Both indicators have a functional relationship because r equals to β adjusted by the relationship between the standard deviations in the two periods (see Table 1). *Positional movement* evaluates individuals' positions based on quantile or rank definitions of the income distribution. Therefore, an individual will move if and only if he or she changes quantile or rank. A particular methodology used to assess positional movement is grounded on the construction of the transition matrix. The transition matrix represents changes in the income quantiles or ranks in two periods. To construct a transition matrix we divided income into quantiles or ranks for both periods. The transition matrix's rows present each income position in period one, and the transition matrix's columns present each income position in period two, where each matrix's cell is the probability of individuals changing from one position to another. Prais (1955), Shorrocks (1978b), and Bartholomew (1996) proposed some of the most used indices based on this methodological approach.

Share movement assesses whether an individual's income rises or falls relative to the income distribution's mean. In this concept, an individual with an unchanged income or unchanged position may experience mobility. *Non-directional income movement or flux* focuses only on the fluctuations of individuals' income over all distributions. Conversely, *directional income movement* focuses on both the fluctuations of individuals' income and the direction of these fluctuations (Fields & Ok, 1999). Table 1 summarizes the main indices for these concepts.

Table 1: Mobility concepts and some indices

Mobility Concept	**Indices**
Temporal independence	For periods t_1 and t_2 and Y income, (1) $logY_2 = \alpha + \beta * logY_1 + u_2$ or beta coefficient (2) r or Pearson correlation index Relationship $r = \beta\,(\sigma_1/\sigma_2)$
Positional movement	(3) Mean absolute quantile change $= 1/n \sum_{j=1} \lvert Q_j^2 - Q_j^1 \rvert$ Based on a transition matrix P_{KxK}, with generic element p_{ij} (4) Shorrocks/Prais $= \frac{(K-trace(P))}{(K-1)}$ (5) Bartholomew $= \frac{\sum_i \sum_j p_{ij} \lvert i-j \rvert}{(K-1)}$
Share movement	(6) Mean absolute Value of share changes $= 1/n \sum_{j=1}^{n} \lvert Y_j^2/\mu_2 - Y_j^1/\mu_1 \rvert$
Non-directional income movement	(7) $M_1 = 1/n \sum_{j=1}^{n} \lvert Y_j^2 - Y_j^1 \rvert$ (8) $M_2 = 1/n \sum_{j=1}^{n} \lvert logY_j^2 - logY_j^1 \rvert$
Directional income movement	(9) $M_3 = 1/n \sum_{j=1}^{n} \left(logY_j^2 - logY_j^1\right)$
Equalization of longer-term incomes	(10) Shorrocks $= 1 - \frac{I\left(\sum_{t=1}^{T} Y_i^t\right)}{\sum_{t=1}^{T} w^t I(Y^t)}$ (11) Chakravarty $= \frac{1-I\left(aggr.Y^t\right)}{1-I(Y^1)} - 1$ (12) Fields $= 1 - \frac{I\left(avg.Y^t\right)}{I(Y^1)}$

Source: Own construction

2.1.1. Mobility as the equalization of longer-term incomes

The concept of mobility as the equalization of longer-term incomes assesses a pattern of income that will increase or decrease inequality the longer the period extends (Fields, 2008). Slemrod (1992) emphasized that cross-sectional analysis may offer an incomplete picture of inequality due to temporary fluctuations in incomes, therefore, a comprehensive picture of inequality and mobility, implies an analysis across periods. Maasoumi & Zandvakili (1986) also developed their work with *lifetime equity* as the core notion because of its advantages compared to analyzing inequality at just one period at a time. Moreover, Atkinson et al. (2012) supported this perspective, arguing that *"the lifetime sum of earnings relative to that in a single period"* (cited by Fields (2010), p. 410)) is the key feature to proposing policies to foster mobility.

Empirical approaches of the concept have two directions. One approach is led by Shorrocks (1978a), who created an index by comparing the inequality of longer-term incomes relative to a weighted average of the inequalities of single-year incomes. Another approach led by Fields (2010) and Chakravarty et al. (1985) as representatives, it compares the inequality of longer-term incomes relative to a single base-year. Nonetheless, Fields (2010) and Chakravarty et al. (1985) differed in the sensitivity of each index to the social judgments captured by the inequality function (see Table 1). Long-term income is quantified by averaging individuals' income over time; the process of averaging reduces income variability by smoothing each individual's income trend. Therefore, an individual's income trend will have less dispersion than a marginal distribution in a specific period.

Taking from Fields (2010), Y_i^t represents the income of individual i in period t. Then, l_i quantifies the long-term economic well-being, and s_i quantifies the short-term economic well-being of individual i. Thus, it is possible to define n-vectors $l \equiv (l_1...l_n)$ and $s \equiv (s_1...s_n)$ in the population as a whole. $I(l)$ and $I(s)$ denote the same inequality functions applied to l and s respectively. The *equalization function*, denoted by $E_{l,s} = E(I(l), I(s))$, measures how much the long-term distribution is compared to the short-term distribution. Fields (2010) made four decisions about the final structure of the index. First, he constrained $I(l)$ and $I(s)$ as a relative structure. Second, the short-term was made the base-year; third, the long-term was made the average of the base-year and final-year incomes. Finally, he posited a Lorenz-consistent inequality measure such as the

Gini coefficient or the Theil index. He also established properties of the index, such as normalization, equalization, disequalization, greater equalization and greater disequalization, which mainly consist of using distributional properties to identify the dominance between these long- and short-term distributions. Therefore, the general expression of the new class of measures for the concept is as follows,

$$E \equiv 1 - \frac{I\left(a\right)}{I\left(Y^1\right)} \tag{1}$$

where: E denotes Fields' class of measures, a is the vector of average incomes, Y^1 is the vector of base-year incomes and *I(.)* is a Lorenz-consistent inequality measure. The measure E is an index of equalization with the following features:

- A positive value indicates that average incomes a are more equally distributed than a base-year incomes Y^1;
- A negative value indicates that average incomes a are less equally distributed than a base-year incomes Y^1; and,
- A zero value indicates that a and Y^1 are distributed equally unequally.

$$Gini = \frac{1}{2n^2\overline{Y}} \sum_{i=1}^{n} \sum_{j=1}^{n} |Y_i - Y_j| \tag{2}$$

$$Theil = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i}{\overline{Y}} \log\left(\frac{Y_i}{\overline{Y}}\right) \tag{3}$$

We use the Gini and Theil inequality functions. Estimations utilizing the Gini function were used in the United States Fields et al. (2002) and for France Buchinsky et al. (2003). We extend the application to the Theil function as a robustness check for sensitivity to our results.

There are no unified criteria for choosing the length of the period to analyze mobility; in fact, it is mostly determined by data availability. Studies for some countries define time spans as short as two years, e.g., France, or as long as five years, e.g., the United States –. For our study, we analyze a moving estimation of six cohorts. Each cohort includes the base-year and the follow-up year four years later, starting in 2009-2013 and ending in 2014-2018.

2.2. Empirical studies on mobility in Latin America

The lack of suitable data structure is a constraint that has shaped the empirical literature of mobility in the LAT region. Most of the studies we reference have used cross-sectional household surveys and focused on education - either as years of schooling or education categories. Cohort analysis or pseudo-panel analysis is the main methodological approach used to build the data sets, although some studies have used retrospective questions on parents' educational level. Studies mainly cover the 1990s; however, more recent publications describe the existence of panel data sets with short time span. Despite their commonalities, outcomes for LAT countries present relevant differences.

Dahan & Gaviria (2001), Behrman et al. (2001), and Gaviria (2007) quantified intergenerational mobility for LAT countries and included the USA in their estimations. They quantified β coefficient and education correlation. Their results show that the USA is more mobile than any LAT country but high heterogeneity are also found among LAT countries. Dahan and Gaviria's (2001) top five countries based on their mobility index are Costa Rica, Peru, Uruguay, Paraguay, and Chile. In the bottom part of the ranking are found El Salvador, Mexico, Colombia, Ecuador, and Nicaragua. Behrman et al.'s (2001) study, which is based on four LAT countries, ranked Mexico as being more mobile, followed by Peru, Colombia, and Brazil. Gaviria (2007) aggregated LAT countries in one category and compared their mobility with that of the USA; however, the author presented the difference in years of schooling of parents and children and showed less difference in Chile, Brazil, Argentina, El Salvador, and Uruguay in the first five places, respectively. All these authors also agree on the high levels of immobility of the LAT region and its rigidity to changes over time.

Neidhöfer et al. (2018) contributed to the creation of a data set of several mobility indices –including β coefficient- based on education for all between 1940 and 1990. Their detailed analysis concluded that older cohorts describe patterns with less mobility than younger cohorts in the LAT region. For instance, the cohort born in 1940-1943 has a β coefficient of 0.6 and the cohort born in 1984-1987 has a β coefficient of 0.41. This finding is in line with those of Azevedo & Bouillon (2010), who identified the expansion of access to higher education as a key factor for upper mobility. Within the LAT region, Neidhöfer et al. (2018) found that the mobility transition from low to high ranks the Dominican Republic in first place followed by Nicaragua, Venezuela, Costa Rica, and Uruguay; in contrast, the last places of the ranking, are found Chile, Peru, Ecuador, Paraguay, and Guatemala.

Unlike previous references, Fields et al. (2007) addressed intragenerational mobility for Argentina, Mexico, and Venezuela, this study is nearly the only one that uses this perspective of mobility at the country level for the LAT region. Based on true panel data for earnings – the Permanent Household Survey (PHS) 1996-2003 Argentina, the National Urban Employment Survey (NUES) Mexico 1988-2001, and the Household Sample Survey (HSS) 1994-1999 Venezuela- the authors modified the econometric specification to quantify the β and add a vector of individuals' features. The findings describe a trend of increasing mobility and reduction of inequality in periods of negative growth mainly due to earnings falling at the top end of the income distribution. In contrast, their estimations of inequality with cross-sectional data indicate an upward trend of inequality for Argentina and Mexico, but an inverted V-shape for Venezuela. They emphasized the quite different picture described for both results based on different data structures pointing out that the analysis over time is more robust.

However, the most often used index also draws criticisms. Some researchers have fundamental concerns regarding beta calculations' usefulness for other methodological approaches such as quantile regression because its pure scalar nature reduces the reliability for its conclusions. Nevertheless, it remains a relevant measure for tracking mobility evolution because all researchers calculate it (Jantti & Jenkins, 2015).

2.2.1. Studies focused on Colombia and why this study focuses on young-adults

Colombian studies on mobility follow the same approaches of LAT studies. They study intergenerational mobility to estimate the β coefficient and implement transition matrices. The data sets used are cross-sectional from Colombian Living Standards Measurement Surveys (LSMS) and the National Household Survey (NHS) for different years. Low mobility and regional heterogeneity are common conclusions among publications. This conclusion is shared by Gaviria (2002), Angulo et al. (2012), and Nina et al. (2003) whose analyses are based on the β and a transition matrix; the first two studies focus on the entire country, while the third focuses only on the country's capital, Bogota. These studies agree with Colombia's low social mobility in the international context, but within the country they estimate high mobility for individuals with lower levels of education.

In contrast, Bonilla's (2010) research focused on regional differences estimating mobility indices from temporal independence and positional movement concepts. The results based on indices of positional movement do not follow economic intuition based on the socioeconomic performance of regions. For instance, using the Shorrock/Prais index Monteria, Neiva, and Valledupar are more mobile than Bogota, Cartagena, and Cali, which is counterintuitive.[5] Bonilla's paper highlights how sensitive conclusions are to different concepts.

Cartagena (2003) and García et al. (2015) differentiated from previous literature regarding their methodological approaches. Although both studies analyzed education, Cartagena (2003) used a probability model to calculate the children's likelihood of obtaining higher educational attainment than their parents. On the other hand, the novel research of García et al. (2015) built a panel data set after merging two governmental data sets; one data set contains information on low-income subsidy recipients, and the other contains the records

[5]For our data, we calculated the Shorrocks/Prais and Bartholomew indices obtaining city rankings going in the opposite direction as the rankings for positive socioeconomic performance.

of the higher educational outcomes. The merged data set allowed the authors to analyze more than 1.9 million records to assess upward mobility through their proposed Educational Vulnerability Index (EVI). Using mothers' years of education for those who only received an elementary school education and comparing this with their children's education, the authors found low upward social mobility in Colombia. Simultaneously, the authors observed that the gap in access to higher education is widening between the low-income population and the middle- and high- income population, with the municipalities located on the Pacific coast being the most affected.

We focus our analysis on individuals between 25 and 30 years old. We set the lower bound at 25 because the literature indicates that all main educational decisions have been made by this age, which is relevant for a mobility cohort definition. In addition, we set the upper bound at 30 because the transition from college to consolidation into the job market is described as particularly difficult for this share of the LAT population. Gontero and Weller's (2015) study analyzed a group aged between 15 and 29 years that demonstrated a mismatch because either the population did not complete college within the age range, or after graduating, they aim to find a job with attributes such as fair payment or social protection, which takes months or even years to find. In fact, there exists a share of the population that neither study nor have a job; this population is a matter of special policy focus on the LAT region (Weller, 2012).

The literature also demonstrates young adults as a relevant population of study for mobility, which is associated with factors such as mortality rate (De Grande et al., 2015), housing decisions (Coulter, 2018), attitudes towards work (Schuck & Shore, 2019), subjective well-being (Schuck & Steiber, 2018) and consequences of being a young adult out of the job market (Zwysen, 2015). From a policy standpoint for the LAT region, Levy & Schady (2013) and Weller (2012) suggested implementing training programs to foster skills, increase the likelihood of being absorbed by the job market and support sustained productivity growth.

We contribute to the empirical literature on intragenerational income mo-

Table 2: Summary of Empirical Literature in Latin America

Authors	**Temporal Perspective: Inter- or Intra-generational**	**Index and Variable of analysis**	**Geographical Scope**	**Time Scope**
Dahan & Gaviria (2001)	Inter	New index based on correlation of schooling between siblings Education	16 LAT countries and USA	Household Surveys in the '90s
Behrman et al. (2001)	Inter	β coefficient Education	Brazil, Colombia, Peru, Mexico and USA	Household Surveys in the '90s
Gaviria (2007)	Inter	β coefficient and Transition Matrix Education	16 LAT countries and USA	Latinobarometro[6] 1996 and 2000
Neidhöfer et al. (2018)	Inter	β coefficient Education	18 LAT countries	Household Surveys and Latinobarometro for several years in the '90s
Fields et al. (2007)	*Intra*	Modified β coefficient Based on *true panel* *Earnings*	Argentina Mexico Venezuela	PHS 1996-2003 NUES 1988-2001 HSS 1994-1999
Gaviria (2002)	Inter	β coefficient Education	Colombia	Household Surveys for several years in the '90s
Angulo et al. (2012)	Inter	β coefficient Education	Colombia Mexico Chile	Household Surveys 2003 2010 2006
Nina et al. (2003)	Inter	Transition Matrix Shorrocks Index Education	Bogota (Colombia)	Household Surveys 1978 and 1998
Bonilla (2010)	Inter	Seven indices from temporal independence and positional movement Education	Colombia	Household Surveys 2008
Cartagena (2003)	Inter	Probabilistic Model Education	Colombia	Household Surveys 1997
García et al. (2015)	Inter	Educational Vulnerability Index Education	Colombia	Panel Data Built based on governmental data sets

Source: Authors' own construction

bility by utilizing the concept of the equalization of longer-term incomes and the index derived from it, which is named the E index. We quantify the E index for Colombian young adults and twenty-three main metropolitan areas for comparative analysis. We process a panel data set containing over 1.3 million records of the formally employed population aged between 25 and 30 years in Colombia between 2008 and 2018.

3. Setting and data

3.1. Macroeconomic context

In this section we present some macroeconomic indicators of LAT countries to contextualize the relative position of Colombia in the region. Then, we present time series of selected indicators for Colombia only. There is evidence that supports the downward trend of inequality and poverty since 2002 in the LAT region, as well as some hints about the trend's deceleration (Gasparini et al., 2016). Figure 2 presents a comparison of the lagged economic growth with inequality (left) and poverty with unemployment (right) between 2008 and 2017. The figure shows a reduction of the Gini coefficient and poverty headcount for all countries except for poverty for Honduras. Conversely, lagged GDP per capita growth and unemployment outcomes are heterogeneous among the countries. Brazil and Argentina lead the adverse macroeconomic outcomes with some of the highest negative variations in GDP growth. As far as unemployment, Brazil and Costa Rica present a substantial increase in the period. In contrast, El Salvador shows a small positive GDP growth and reduction of unemployment.

Colombia's outcomes for inequality and unemployment are among the highest improvements compared to those of other countries. Colombia claims the largest reduction of unemployment and the second largest reduction of inequality - after El Salvador. Despite these encouraging outcomes, Figure 2 also shows Colombia having the second highest unemployment rate of the region and inequality above the 2017 LAT average (CEPAL, 2019).[7] In this period, conditions

[7] Colombia's Gini coefficient is also higher than the 2018 OECD average (average

based on inequality and poverty outcomes have slightly improved, while changes in GDP and unemployment continue to describe a decade of challenges.

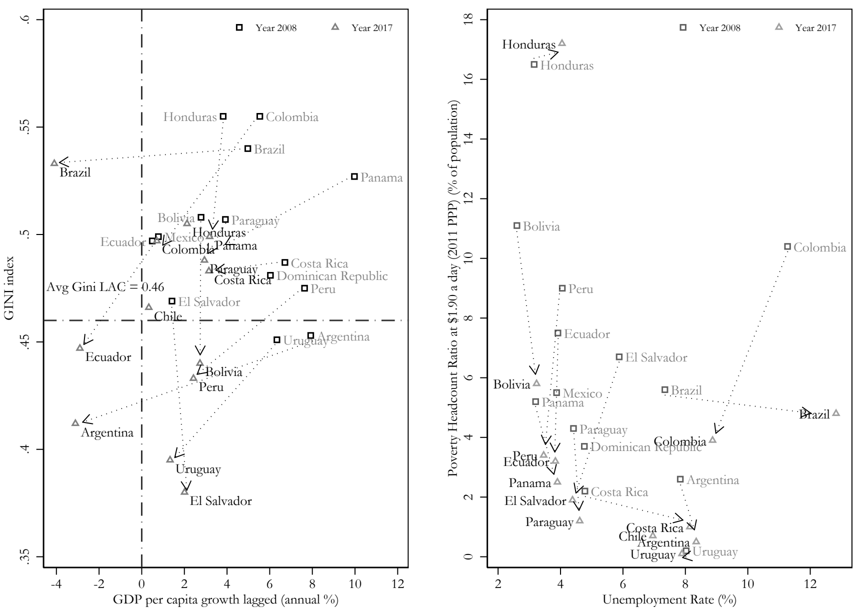

Figure 2: Inequality, Economic Growth, Poverty, and Unemployment, LAT countries 2008-2017. Source: Azevedo (2011)

In the next section, we describe the PILA data set of formal employees in Colombia, therefore, it is relevant to contextualize the countries' trends of formal employment. Figure 3 presents the time series of the proportion of wage and salaried workers over the total employees population between 2007 and 2018. This time series allows us to identify three groups in terms of high, intermediate, and low percentages. The group of high percentages ranges from 64.2% to 77.8%, the intermediate group ranges from 49.3% to 60.4%, and the low group ranges from 37.2% to 49%. Countries with a high percentage of formal

Gini=0.317) (OECD, 2019). Colombia was ratified as an OECD member in April 2020. See http://www.oecd.org/latin-america/countries/colombia/#d.en.345234 employment are among the most advanced economies in South and Central America, namely, Brazil, Argentina, Chile, Uruguay, Mexico, and Costa Rica. In particular, Colombia's employment formality rate fluctuates approximately 48%, indicating a low level of formality. We identified Colombia's trend as the *ad-hoc* upper threshold to differentiate the low group from the intermediate group.

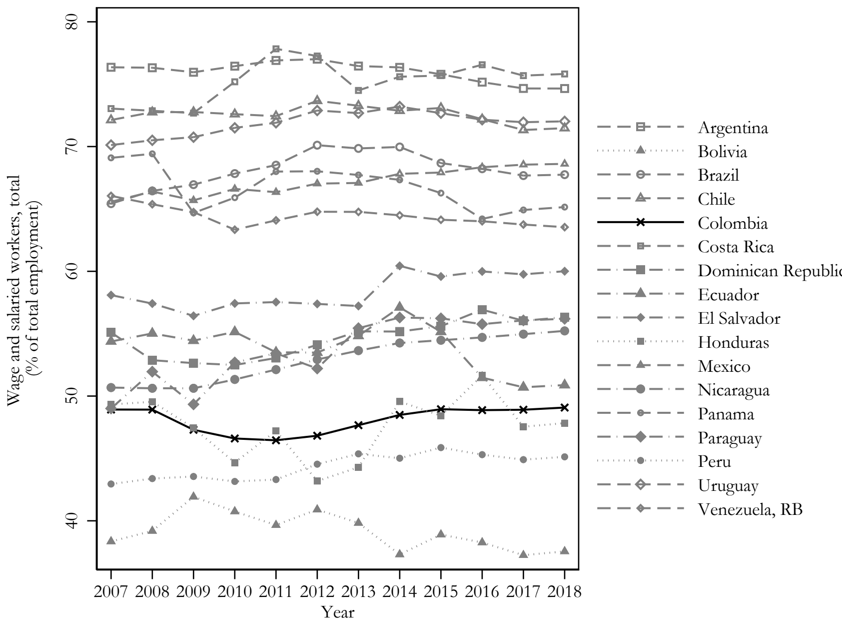

Figure 3: Trends of wage and salaried workers in LAT countries 2007-2018. Source: Azevedo (2011)

The literature identifies structural challenges that hinder regional development such as the institutional framework, chronicle informality and high levels of inequality (Gasparini & Tornarolli, 2009; Gasparini et al., 2016). Nonetheless, the previous contextualization indicates that despite negative outcomes on economic growth, countries' performances on inequality, poverty, and labor market conditions were not substantially affected in the period of analysis. However, recent evidence suggests that the decreasing trend in inequality can be jeopardized by the low performance in regard to economic growth (Gasparini et al., 2018; Tornarolli et al., 2014; CEPAL, 2019).

3.2. The Colombian formally employed population: PILA, a panel data set

In our analyses, we calculate the E index using information from the social security administrative records known as PILA, which is collected by the Colombian Ministry of Health and Social Security (MHSS). Formally certified firms and individuals engaged in economic activities should mandatorily report labor information to the MHSS. The reported information includes employees' sex, age, and monthly earnings; the total number of firms' employees; industry classification; and the city location. The MHSS started collecting data in the second half of the year 2008; we have access to PILA records since the starting date until the first half of the year 2018. We focus on the main twenty-three metropolitan areas, which correspond to seventy-two cities classified based on population size and greater contributions to the country's GDP. We emphasize the importance of PILA as a source of information because it has a panel data structure and includes the entire formally employed population.

We use monthly formal earnings for our calculations.[8] For each individual, we calculated the average earnings for twelve months from 2009 to 2018. We took the location of each employee as the city where he or she appeared the greatest number of times among the twelve months in the base-year. Therefore, the income mobility index estimates are conditioned on employee location in the base-year. This decision is supported based on the literature on inequality of opportunities, and individuals' conditions in the early stages of their lives explain the lifetime earnings path (Ferreira & Meléndez, 2012; Chetty et al., 2014; Roemer & Trannoy, 2015). Finally, we selected the population for analysis based on the following criteria. First, because social security payments can be

[8] We excluded the year 2008 from the analysis since during the first half of that year, PILA had just began to be collected by the MHSS and the quality of the data is thus not the best for our long-term comparison.

made daily, we kept records for individuals who made payments for at least 20 days per month. Second, we performed an outliers analysis; as a result, we excluded individuals with earnings below the 1^{st} percentile and above the 99.5^{th} percentile. Finally, we focused on individuals between the ages of 25 and 30 years old. After applying these criteria, our final sample had more than 1.3 million records.

3.2.1. Data description by Colombian metropolitan areas

In this section, we contextualize Colombia's main metropolitan areas based on population, geographical extension, relative weight to the country's GDP and population density. In the Appendix A.1, Figure A1 presents a map of Colombia with the metropolitan areas' location and total observations per area as obtained from PILA. The twenty-three metropolitan areas correspond to 72 cities out of total of 1,122. The metropolitan areas are Bogota D.C., Medellin, Cali, Barranquilla, Bucaramanga, Cartagena, Pereira, Cucuta, Ibague, Manizales, Villavicencio, Santa Marta, Neiva, Valledupar, Monteria, Armenia, Pasto, Popayan, Tunja, Sincelejo, Riohacha, Florencia, and Quibdo.

Historically, Bogota, Colombia's capital city, has led the best performances of social and economic indicators, followed by Medellin, Cali, Barranquilla, and Bucaramanga (Pérez et al., 2014). Therefore, the expectation of important mobility outcomes relies on these areas. Table Appendix A1 shows that the twenty-three metropolitan areas concentrate approximately 66.1% of the Colombian GDP and 56.3% of the country's population. With a population density as widespread as ranging from 4,003.8 for Barranquilla to 37.7 for Quibdo, we observed a positive relationship between the relative weight to the country's GDP and population density (correlation is 0.52), which is in line with the hypotheses of agglomeration economies (Ferreyra & Roberts, 2018).

After applying the criteria to select the sample for the analysis, we obtained more than 1.3 million records with the following distribution. The total number of observations for Bogota – obs= 565,432; ranked in the first place – is more than twice that of the total observations for Medellin – obs= 225,101; ranked

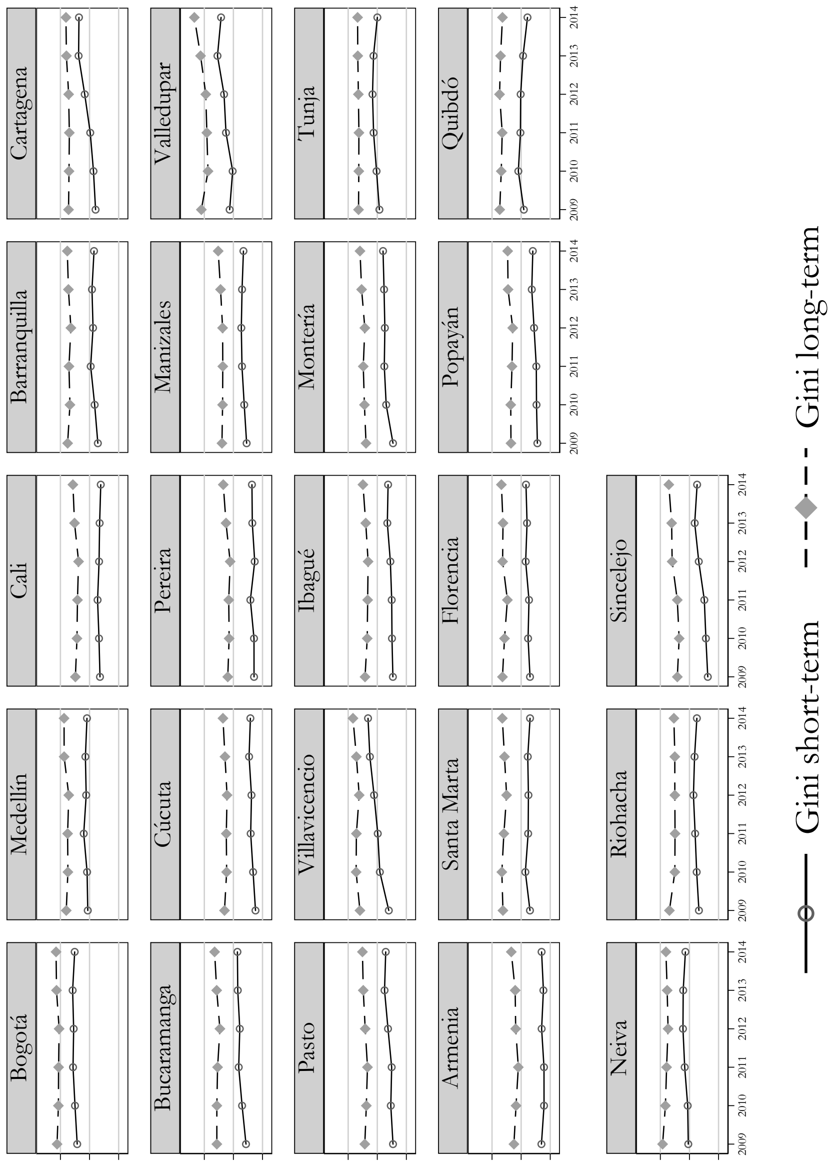

Figure 4: Trends comparison of the Gini coefficient in the short- and long-term by metropolitan area

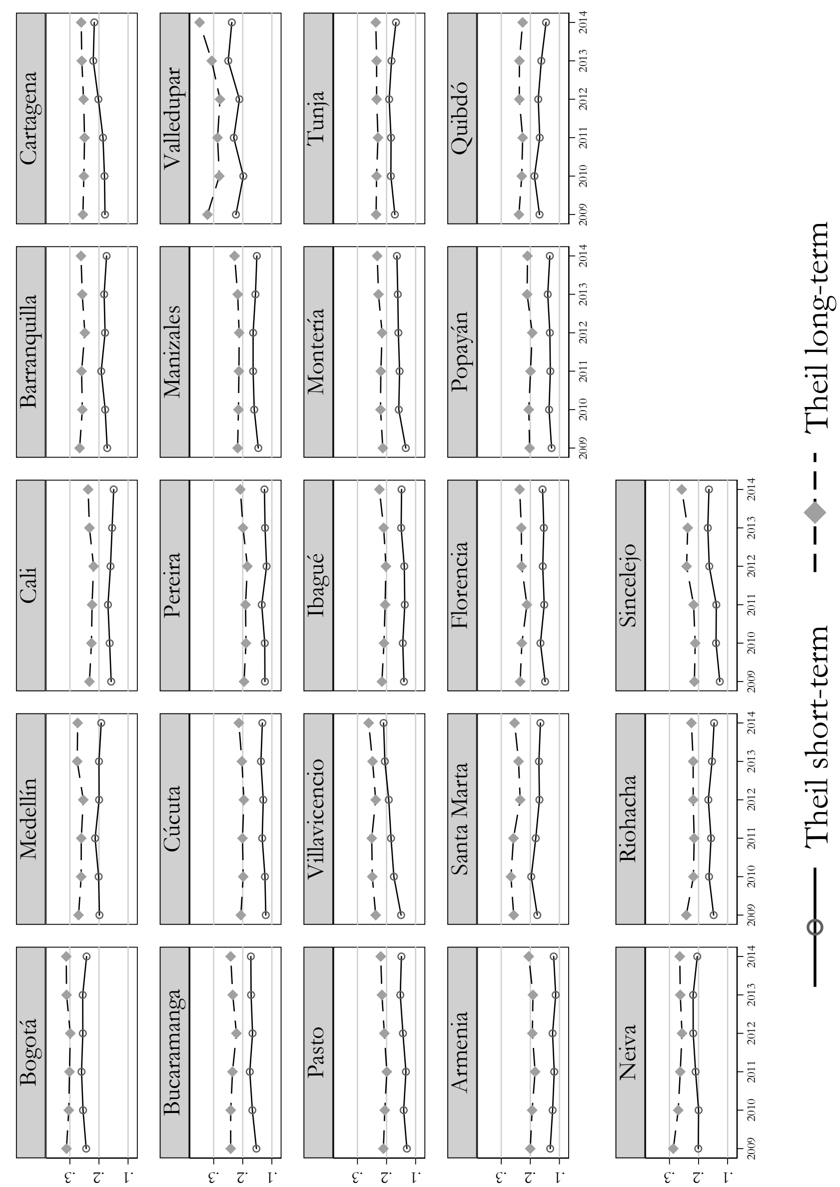

Figure 5: Trends comparison of the Theil coefficient in the short- and long-term by metropolitan area

in the second place. Medellin and Cali are ranked as having between 100,000 and 500,000 observations. Barranquilla and Bucaramanga are ranked as having between 50,000 and 100,000 observations. The remaining eighteen areas have less than 50,000 observations each. The final number of observations retains the heterogeneity from the census population.

Figure 4 and Figure 5 compare the short- and long-term trends by metropolitan area with both the Gini and Theil functions, respectively. For both figures, the long-term trend is always above the short-term trend in all metropolitan areas. Bogota, Medellin, Cartagena, Villavicencio, and Tunja depict the highest levels of inequality with both the Gini and Theil functions. Conversely, Armenia, Cucuta, Pereira, Sincelejo, and Popayan depict the lowest levels of inequality. Regarding the inequality trends, the short-term estimates of most of the areas are described by an inverted U-shape, excluding Cartagena, Valledupar, Villavicencio, Monteria, and Sincelejo, which are depicted by an upward inequality trend. The downward trend of the final year is most observable for Tunja, Santa Marta, and Quibdo. The conclusions hold for both the Gini and Theil functions; the main difference between the two functions relies on the magnitude, with Theil estimates being lower than Gini estimates.

4. Results: Mobility as the equalization of longer-term incomes, the Colombian case and its metropolitan areas, 2009 - 2018

In this section, we discuss the degree of equalizing mobility based on the estimations of the E index with both the Gini and Theil functions for the country by both gender and metropolitan area. Figure 6 depicts the trend for the country, Figure 7 depicts the trend by gender, and Figure 8 depicts the trend by metropolitan area. In the Appendix A.1, Figure A2 and Figure A2, we map the results by metropolitan areas with the Gini and Theil functions, respectively. To understand the changes over time, in each figure we study mobile cohorts in 4-year windows using the same restrictions mentioned in section 3.2.

The E negative values for Colombia's index estimates indicate that the coun-

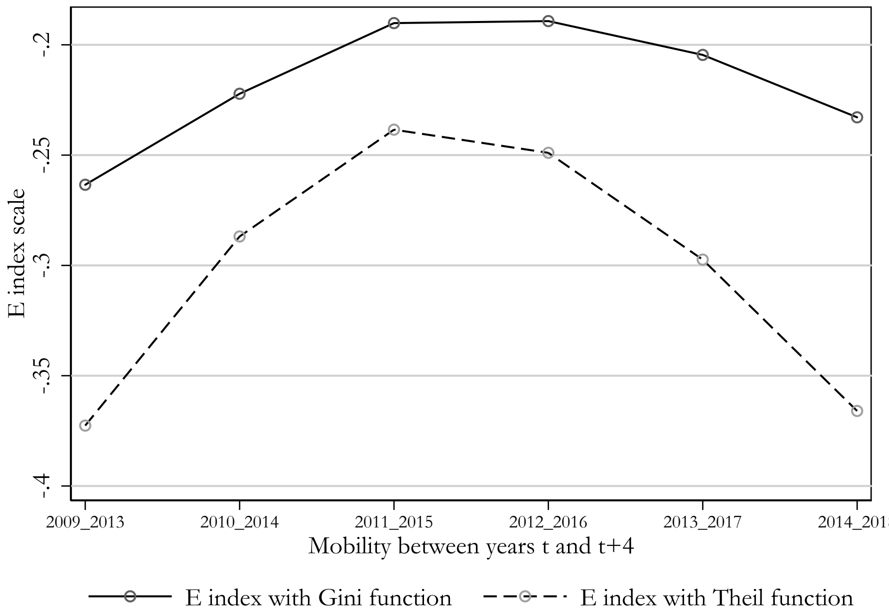

Figure 6: Trend of the E for Colombia with the Gini and Theil functions

try's young adults experience a process of disequalizing income mobility. That is, there is income mobility for formally employed young adults while, simultaneously, there is also increasing inequality within this population.[9] Figure 6 depicts an inverted U-shape, which means that young adults mobility was less disequalizing up to the 2011-2015 period; then, it remained virtually stagnate for the period of 2012-2016. After this period, income mobility has become more disequalizing, which means that inequality has worsened. The level of disequalizing mobility is almost the same as it was at the beginning of the period, which describes a setback to the improvement of inequality.

The inverted U-shaped trend also holds for both females and males; how-

[9]Countries such as the United States also show negative values of the E for some periods (Fields et al., 2002). In contrast, France always shows positive values for the analyzed periods (Buchinsky et al., 2003).

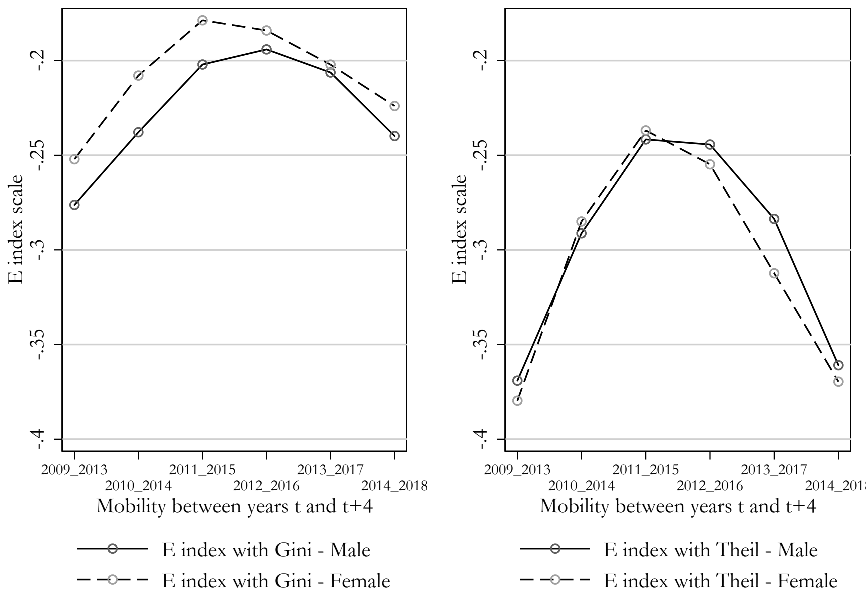

Figure 7: Trend of the E by Gender with the Gini and Theil functions

ever, some interesting differences arise between the two groups (Figure 7). The female young adult population experiences less disequalizing mobility compared to male young adults according to the Gini estimates. Nonetheless, the differences between the groups show a decrease from 2012-2016 until the end of the period. This is not the case using the Theil estimates because, depending on the period, one population has less disequalizing mobility than another. In fact, the downward section of the trend indicates that the female population experiences more disequalizing mobility than the male population.

Income mobility patterns across metropolitan areas unveil a large heterogeneity within the country. We analyze this heterogeneity in two dimensions: type of trends and differences in magnitude. In Figure 8, we observed three types of trends; inverted U-shaped, upward, and fluctuating. In the inverted U-shaped trend type we classify the main Colombian cities including Bogota, Medellin, Cali, Barranquilla, together with Bucaramanga, Cucuta, Ibague, Tunja, and

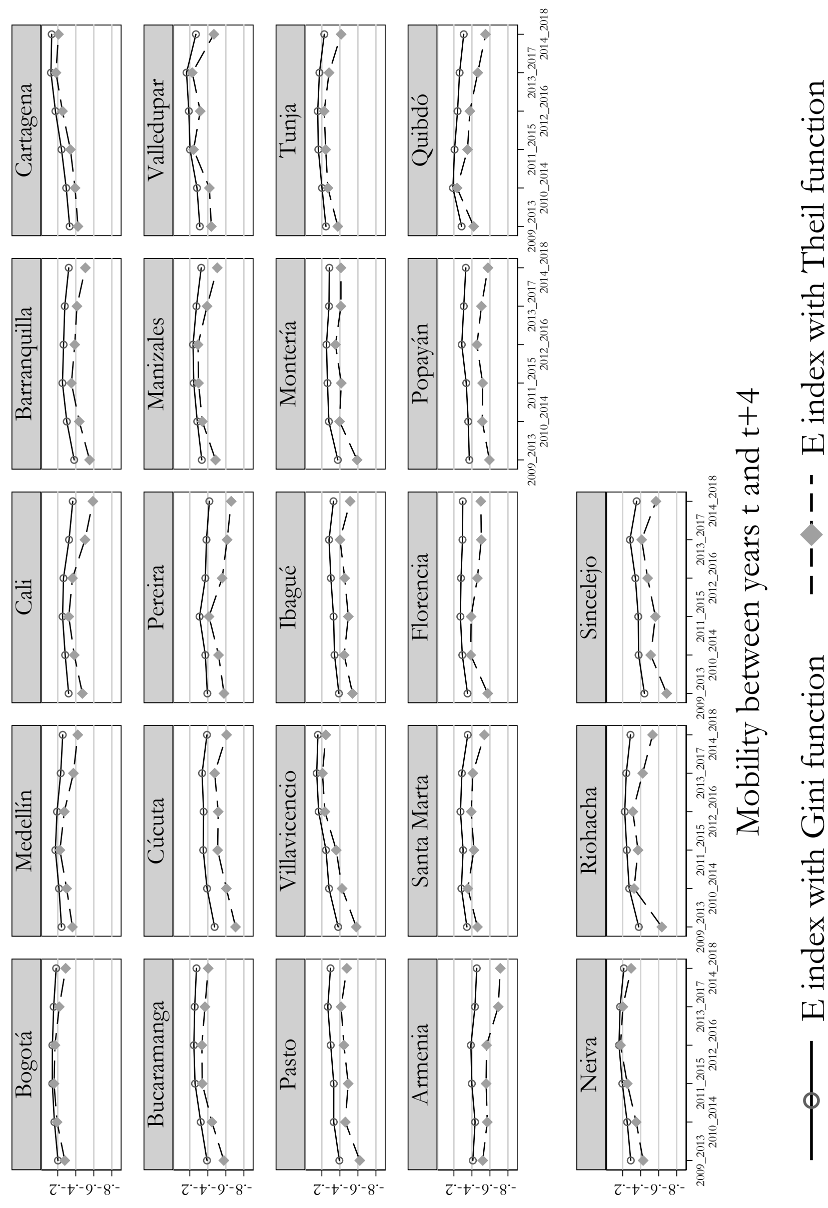

Figure 8: The E trends by metropolitan area with the Gini and Theil functions

Neiva. In the upward trend type, we classify Cartagena and Villavicencio. In the fluctuating trend type, we classify the remaining metropolitan areas. Nonetheless, in this last group, we highlight the downward trend in the final section of the time series for Quibdo and Pereira. Although trends using the Gini and Theil calculations are similar, the Theil calculations emphasize more patterns within the time series.

Regarding the differences in magnitude, in Figure 9, we present a ranking based on the average value for the time series. As a result, the metropolitan areas that are ranked better for reducing the disequalizing mobility for young adults are Bogota, Tunja, Neiva, Cartagena, Medellin, and Villavicencio; at the other end of the spectrum, we find Armenia, Cucuta, Pereira, Sincelejo, Popayan, and Ibague. The rankings do not substantially differ after comparing both inequality functions.

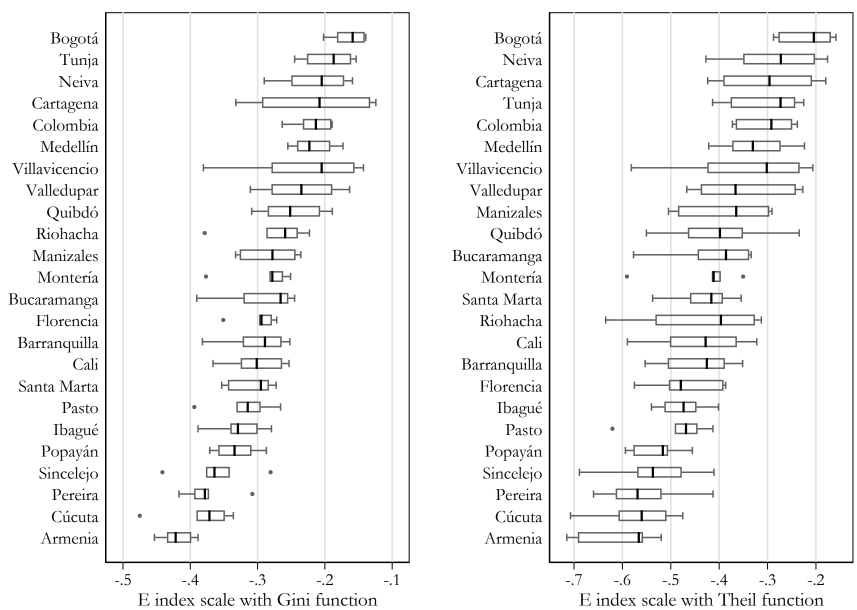

Figure 9: Ranking based on the average E for the study period

Overall, formally employed young adults experience disequalizing mobility.

The female population benefits from this mobility more than the male population, based on Gini estimates, although this conclusions is not as clear with Theil estimates. We observe less disequalizing mobility in Bogota, Tunja, Neiva, Cartagena, Medellin, and Villavicencio but some of these metropolitan areas present high levels of cross-sectional (short-term) inequality. Therefore, we find that some degree of mobility is accompanied by rising inequality, as was found by Fields et al. (2007). Based on our findings, the distinction between intra- and intergenerational mobility is relevant because the literature on mobility using education in Colombia shows important improvements, but it does not address how its upward trend affects inequality.

4.1. Mobility as the equalization of longer-term incomes and its association with industry employment, fiscal features, and socioeconomic performance

In this section we show the association between intragenerational mobility measured by the E index and some macroeconomic and institutional characteristics identified in the literature. Lora (forthcoming) and O'Clery et al. (forthcoming) adapted complexity theory to analyze cities; the theory was originally developed to analyze countries (Hidalgo & Hausmann, 2009). The authors ground their theory in the linkage between industrial diversity and the complementarities developed among sectors and showed that these interactions drive cities to move towards the production of more complex/value-added final products. The emergence of more complex industries will then change the labor market structure, which will tend to increase formal employment. We hypothesize that cities with higher complexity and more formal employment foster mobility (Gasparini et al., 2011).

The literature on geographical disparities identifies some relevant dimensions that might be associated with mobility as the equalization of longer-term incomes. One of these dimensions is tax structure. In Colombia, there is evidence that the tax structure creates a negative association between the GDP per capita and tax transfers from the central government to local governments (Rocha García & Vivas Benitez, 1998). In addition, this negative association is also related to the low level of local autonomy in regard to public expenditure (Barón & Meisel-Roca, 2003; Meisel-Roca, 2007). Furthermore, Meisel-Roca and Galvis (2016) analyzed regional intergenerational mobility using education. The authors found a center-periphery pattern of mobility in the country, where the center region is more mobile. The authors also concluded that an individual's location (origin) is a key determinant of upward mobility between parents and children (Chetty et al., 2014).

Based on education as the key driver of mobility, Meisel-Roca's (2014) policy proposal stated the importance of expanding education expenditures through local governments. Galvis & Meisel-Roca (2012) found persistent low performance in historical vulnerable regions in Colombia. Other studies have explored regional disparities by analyzing additional variables such as poverty and inequality (Poveda & Martínez, 2011; Cortés & Vargas, 2012; Hahn-De-Castro & Meisel-Roca, 2018).

Therefore, we explore the association through regressions of a series of separately variables on our E index estimates. For each metropolitan area, we approach industry employment through the economic complexity index (ECI), [10] education expenditures through the number of students in public system per 100,000 inhabitants, health expenditures through the number of low-income individuals covered with public health insurance per 100,000 inhabitants, fiscal independence through metropolitan areas' tax income as proportion of their total budget, and poverty through poverty headcount. Our estimations follow the specification of equation 4:

$$E_{i,t+s} = \alpha_0 + \alpha_1 * Feature_{i,t} + \beta_i + \epsilon_{i,t} \tag{4}$$

where $E_{i,t+s}$ is computed as shown in equation 1 for periods $t+s$, $Feature_{i,t}$

[10]The ECI measures the knowledge intensity of a city by considering the knowledge intensity of each sector through the number of employees working on in each sector. Thus, negative ECI values indicate less complexity, while positive values indicate more complexity. See more details in the Appendix A.2

represents the selected variables for each metropolitan area i in year t, β_i captures the metropolitan area fixed effect, and $\epsilon_{i,t}$ is the error term. Because several issues challenge the causal interpretation of α_1 such as omitted variables or endogeneity due to the relationship between income inequality and mobility in our index estimations, we limit parameter interpretation as a simple association and we focus on the direction of the relationship between the index and the feature studied. However, our objective aims to explore equalizing/disequalizing mobility and the directions of the potential underlying mechanisms.

Table 3 presents the estimates of α_1 with the selected variables for the E index with both the Gini and Theil functions displayed in different subtables. We also include the standardized coefficient estimates. The E index with the Gini function is not significantly associated with the economic complexity index or health. Nonetheless, it is significantly associated with education, fiscal independence, and poverty. The expansion of public education and more fiscal independence are positively associated with more equalizing mobility. On the other hand, a reduction in poverty is associated with more equalizing mobility. The E index with the Theil function is positively and significantly associated with fiscal independence only. Thus, we find that metropolitan areas with better capacity to collect their own resources is the robust identified feature.

Findings from this section uncover relationships not previously explored, thereby opening discussion for future research on causal mechanisms. Nevertheless, an interesting reflection based on Meisel-Roca and Galvis' 2016 region ranking, which describes a very close match with our metropolitan areas results, suggests that locations attract individuals with more years of education who will take advantage of better jobs and higher earnings. As a consequence, it is highly likely that inequality within metropolitan areas increases and the gap among them widens. Thus, metropolitan areas performances based on formal employed population describes mobility with increasing inequality. This outcome is compatible with fiscal independence as currently is.

Table 3: Equalizing mobility with industry employment, fiscal features and socioeconomic performance

E index	Industry Employment (1)	Education (2)	Health (2)	Fiscal Independence (2)	Poverty (3)
Gini Function					
α_1	0.005	0.0017	5.3e-07	0.0047	-0.0038
Standard Error	(0.0245)	(0.0006)	(6.2e07)	(0.0015)	(0.0011)
Standardized Coefficient	0.065	0.571	0.130	0.808	-0.520
Area Met. F.E.	Yes	Yes	Yes	Yes	Yes
Observations	138	115	138	138	128
Theil Function					
α_1	0.014	0.0029	2.1e-06	0.0085	-0.0018
Standard Error	(0.0445)	(0.0018)	(1.2e-06)	(0.0025)	(0.0021)
Standardized Coefficient	0.108	0.588	0.317	0.890	-0.148
Area Met. F.E.	Yes	Yes	Yes	Yes	Yes
Observations	138	115	138	138	128

Author's own calculations based on (1) PILA and (2) Municipalities Statistical Lab, Universidad de los Andes - CEDE

5. Decomposing Upward Mobility

Similar to the previous section in which we explored the association between intragenerational mobility and macroeconomic variables, the aim of this part of the analysis is to show the association between intragenerational mobility and microeconomic factors associated with individual, family, household, and firms characteristics. To do so, we introduce a simple decomposition method for the E index based on the recentered influence function (RIF) regression (Firpo et al., 2009).

Policy assessment is relevant to identifying effects on outcome distributions, e.g., earnings, due to changes in one or more observations that compose the distribution. This means, that more information on distribution behavior helps to understand how small changes in one or more observations affect the statistic of that distribution. RIF regression (Firpo et al., 2009) is one of the techniques contributes to this process. Because RIF allows the decomposition of any statistics of the unconditional distribution of an outcome, we applied its properties to our long-term Gini estimates (Becchetti et al., 2014; Gradín, 2018) and we extended them to our long-term Theil estimates on the determinants by individual, household, living conditions, and firms characteristics.

To implement RIF, we link PILA individual administrative records to SISBEN-3. SISBEN-3 is an administrative cross-sectional data set from 2009 that provides detailed demographic information such as age, sex, municipality, living conditions, education level, and family characteristics. While PILA provides information on earnings, SISBEN-3 provides information on individual and household information of low-income families. Thus, we are able to construct an individual-level year panel to study income mobility and its determinants. Due to data restrictions, we merge the individuals of SISBEN-3 year 2009 with those in PILA year 2010. As a result, our sample contains young adults in 2010 with their socioeconomic information from 2009. The sample for this exercise is 379,874 observations.

5.1. Recentered Influence Function and inequality functions

Following Firpo et al. (2009), the general expression of the linear approximation for RIF is as follows:[11]

$$RIF(y; v, F_Y) = v(F_Y) + IF(y; v, F_Y) \tag{5}$$

where $IF(y; v, F_Y)$ is the influence function and $v(F_Y)$ is a distributional statistic of interest. Under the assumption that the conditional expectation of the $RIF(y; v, F_Y)$ can be written as a linear function of explanatory variables X where the β coefficient can be estimated by OLS, the RIF regression is defined by Firpo et al. (2009), as shown in equation 5. Therefore, any statistic - including Gini and Theil - can be expressed as a general function of a covariates group

[11] See more details in appendix Appendix A.3

X, expressing the β as the marginal impact on the statistic function on small changes in the average values of its covariates, as shown in equation 6:

$$E[RIF(y; v, F_Y)|X] = X'\beta \tag{6}$$

Substituting equation 6 into equation 1, we have the following:

$$E = 1 - \left[\frac{\gamma_0^{v(F_a)} + \overline{X}'\beta^{v(F_a)}}{v(F_{Y^1})}\right] \tag{7}$$

Then, we rearrange the terms as follows:

$$E = 1 - \frac{\gamma_0^{v(F_a)}}{v(F_{Y^1})} - \frac{\overline{X}'\beta^{v(F_a)}}{v(F_{Y^1})} \tag{8}$$

This final expression allows us to decompose the total effect of each covariate on the Gini or Theil statistic associated with the E index. In our empirical application, we divided the covariates into groups of individual, household, living conditions, and firm characteristics, measured at or before the baseline year of 2010, to identify how the initial conditions of individuals determine their income mobility between 2010 and 2018.

5.2. RIF results

Table 4 shows the coefficients associated with the RIF regression for the E index with Gini and Theil statistics. This is equivalent to the numerator of the expression in brackets from equation 7.[12] In both models, we include metropolitan area fixed effects to control for common factors associated with an individual's location at the baseline which are key determinants of upward mobility (Chetty et al., 2014). A negative coefficient indicates the contribution to the reduction of disequalization, while a positive coefficient indicates the opposite.

[12] In Table Appendix A2 and Table Appendix A3 in the appendix Appendix A.3, we present estimates for mean, median, percentile 25, and 75 as benchmark application of the technique. For instance, the results for the mean are equivalent to OLS, and coefficients interpretation are consistent with standard results in the literature.

For Gini estimates, the gender coefficient is negative and significant, indicating that the female population contributes to reducing disequalizing mobility. Similarly, having secondary and technological education (T&T) contributes to reducing disequalizing mobility as well. Other types of job contracts, such as independent and cooperative contracts, also reduce disequalizing mobility. Compared with the agricultural sector, working in the trade, hotels, and food, real estate, community and personal services, and other sectors, produces the same effect on our mobility estimates. Conversely, a higher age and higher level of education increase disequalizing mobility. Being married and owning a home also produce the same effect on mobility. Owning a washing machine, a shower heater, and a microwave oven are also associated with more disequalization. Compared with the agricultural sector, working in the electricity, gas, and water, and finance sectors increase disequalizing mobility.

The differences between the Gini and Theil estimates are not substantial. In general, both agree with the significance of the coefficients as well as the direction of their effects. However, one important discrepancy between both estimations is that of gender. For Theil estimates this individual feature is not significant. Nonetheless, it is consistent with the results in Figure 7, where the female population describes better results with the Gini function and almost no differences between genders with the Theil function. Another interesting result arises from the education variable coefficients. These coefficients describe an scenario in which individuals with high human capital take advantage of more rewarding labor markets, while individuals with low human capital obtain more homogeneous outcomes. In terms of the earnings distribution, this process creates more inequality - distance - between individuals.

To provide a better picture of these results, we group the combined effects of characteristics associated with individual, household, living conditions, and firm characteristics. To group the effects, we replace $\overline{X}$ from equation 8 with the average of the characteristic and β with the estimated coefficient. Then we add each subset of characteristics (see Figure 10). The net effect of utilities and firms reduces disequalizing mobility. Although the remaining group of features

Table 4: RIF Estimates for the E index with Gini and Theil Statistics

Variable	Gini		Theil	
	Coef.	Std. Err.	Coef.	Std. Err.
Gender	-0.0031	(0.0012)	0.0004	(0.0020)
Age	0.0025	(0.0003)	0.0046	(0.0005)
Secondary	-0.0350	(0.0015)	-0.0516	(0.0026)
T&T	-0.0288	(0.0023)	-0.0695	(0.0037)
Undergraduate or higher	0.1905	(0.0032)	0.2103	(0.0056)
Married	0.0143	(0.0011)	0.0166	(0.0019)
Home property	0.0129	(0.0012)	0.0155	(0.0020)
Overcrowding	-0.0005	(0.0013)	0.0011	(0.0020)
Wall material	-0.0063	(0.0018)	-0.0084	(0.0029)
Floor Material	0.0101	(0.0011)	0.0103	(0.0018)
Sewer service	-0.0073	(0.0018)	-0.0080	(0.0029)
Gas service	0.0003	(0.0014)	-0.0010	(0.0023)
Refrigerator	-0.0087	(0.0012)	-0.0126	(0.0020)
Washing Machine	0.0075	(0.0012)	0.0072	(0.0019)
Cable	0.0017	(0.0011)	0.0027	(0.0018)
Shower Heater	0.0221	(0.0021)	0.0284	(0.0036)
Microwave Oven	0.0363	(0.0024)	0.0461	(0.0041)
Independent	-0.0165	(0.0016)	-0.0215	(0.0025)
Cooperative	-0.0148	(0.0015)	-0.0151	(0.0023)
Manufacturing	-0.0071	(0.0030)	-0.0277	(0.0053)
Electricity; gas and water	0.0291	(0.0111)	0.0147	(0.0183)
Construction	-0.0026	(0.0037)	-0.0225	(0.0061)
Trade; hotels and food	-0.0330	(0.0027)	-0.0527	(0.0046)
Transportation and warehousing	0.0005	(0.0040)	-0.0060	(0.0061)
Finance	0.0319	(0.0041)	0.0104	(0.0071)
Real estate	-0.0138	(0.0024)	-0.0269	(0.0043)
Community; social and personal services	-0.0104	(0.0024)	-0.0279	(0.0043)
Other sector	-0.0091	(0.0029)	-0.0231	(0.0050)
Constant	0.2826	(0.0094)	0.1273	(0.0152)
Observations	379,874		379,874	

Standard errors in parentheses

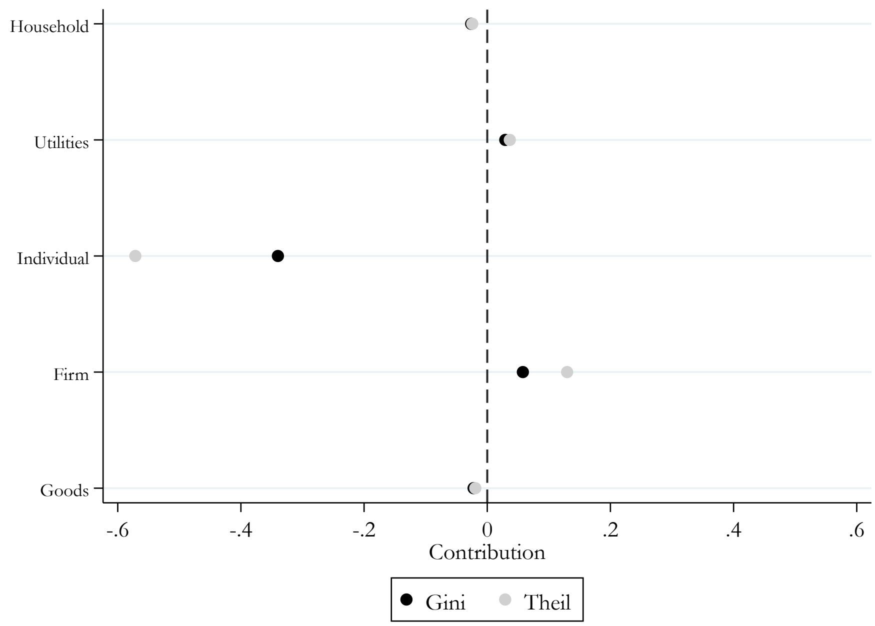

Figure 10: Contribution in E index by groups with Gini and Theil functions

has a negative net effect, household and durable goods are very close to having a null net effect. The individual characteristics have the largest negative effect led by the effect of higher education.

6. Conclusions

In this paper we analyze intragenerational income mobility based on the concept of income mobility as the equalization of longer-term incomes. We estimate the proposed index of Fields (2010), the E index, for formally employed young adults in Colombia and its main twenty-three metropolitan areas between 2009 and 2018. Our comparative analysis's main strengths rely on presenting a setback in terms of disequalizing mobility for young adults and its heterogeneous effect across the metropolitan areas. In addition, we capture the components that contribute to improving mobility for a subsample of our data. Thus, we empirically quantify the income mobility as one of the ways to transfer inequality along an individual‘s productive life cycle.

We find that formally employed young adults experience disequalizing mobility, and it is observed less disequalizing mobility in Bogota, Tunja, Neiva, Cartagena, Medellin, and Villavicencio. We also find the E index with the Gini function is not significantly associated with the economic complexity index or health. Nonetheless, it is significantly associated with education, fiscal independence, and poverty. Thus, the expansion of public education and more fiscal independence are positively associated with more equalizing mobility. On the other hand, a reduction in poverty is associated with more equalizing mobility. The E index with the Theil function is positively and significantly associated with fiscal independence only. As far as the decomposition with the RIF regression, the largest effect on increasing disequalization is due to individual characteristics. Conversely, the net effect of utilities and firms reduces disequalizing mobility.

Comparing our findings with the current studies on the topic we contribute to the literature because we implement an index with administrative records for a developing country, we link a concept with a methodological approach, and we expand the exploration to the metropolitan area level. In addition, we present a robust estimation after using two different inequality functions. The positive correlation between income and education (Neidhöfer et al., 2018) posits a challenge to closing the gap between lagged and advanced cities because individuals with more years of education will take advantage of better jobs; thus, higher earnings will increase inequality. We will delve into this hypothesis in future research. We emphasize the importance of PILA as a source of information because it has a panel data structure, which is not very common in LAT countries, and it includes the entire formally employed population.

Our findings are encouraging; however, we are aware of the methodological and theoretical scope of this study. Methodologically, our data set collects information on formal employees, thereby excluding informal employees from our analysis. Nonetheless, we believe the results of the formal sector are reasonable signals of the conditions of the economy at large. Theoretically, as discussed by

Fields (2010), since the family of index measures depends on the selection of the inequality function, we use the Gini and Theil functions as robustness checks, however, additional research will need to test the results sensibility based on other different inequality functions.

Acknowledgments

The authors would like to thank all the participants to the Banco de la Republica seminar and the PEAK Urban Workshop at Universidad EAFIT for their helpful comments in February 2020.

Funding

This article was completed with support from the PEAK Urban Program, supported by UKRI's Global Challenge Research Fund, Grant Ref: ES/P011055/1.

References

Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en Colombia, . URL: `http://econpapers.repec.org/paper/col000089/010323.htm`.

Atkinson, A. B., Bourguignon, F., & Morrison, C. (2012). *Empirical studies of earnings mobility*. London: Harwood Academic Publishers.

Azevedo, J. P. (2011). WBOPENDATA: Stata module to access World Bank databases. Statistical Software Components S457234, Boston College Department of Economics. Revised 02 Jul 2020.

Azevedo, V., & Bouillon, C. (2010). Intergenerational Social Mobility in Latin America: A Review of Existing Evidence. *RAE Revista de Análisis Económico*, *25*, 7–42.

Barón, J. D., & Meisel-Roca, A. (2003). La descentralización y las disparidades económicas regionales en colombia en la década de 1990, .

Bartholomew, D. J. (1996). The statistical approach to social measurement .

Becchetti, L., Massari, R., & Naticchioni, P. (2014). The drivers of happiness inequality: suggestions for promoting social cohesion. *Oxford Economic Papers*, *66*, 419–442.

Behrman, J. R., Gaviria, A., Székely, M., Birdsall, N., & Galiani, S. (2001). Intergenerational mobility in latin america [with comments]. *Economia*, *2*, 1–44.

Birchenall, J. A. (2001). Income distribution, human capital and economic growth in colombia. *Journal of Development Economics*, *66*, 271 – 287. URL: http://www.sciencedirect.com/science/article/pii/S0304387801001626. doi:https://doi.org/10.1016/S0304-3878(01)00162-6.

Bonilla, L. (2010). Movilidad inter-generacional en educación en las ciudades y regiones de Colombia. *Documentos de Trabajo Sobre Economía Regional y Urbana; No. 130*, .

Buchinsky, M., Fields, G., Fougère, D., & Kramarz, F. (2003). Francs or ranks? earnings mobility in france, 1967-1999, .

Cartagena, K. (2003). Educacion y movilidad intergeneracional en Colombia, 1929-1996. *Estudios Economicos de Desarrollo Internacional*, *3*. URL: https://ideas.repec.org/a/eaa/eedein/v3y2003i3_7.html.

CEPAL (2019). *Social Panorama of Latin America 2018*. ECLAC.

Chakravarty, S. R., Dutta, B., & Weymark, J. A. (1985). Ethical indices of income mobility. *Social Choice and Welfare*, *2*, 1–21.

Chetty, N., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States, . *129*. doi:10.1093/qje/qju022.

Cortés, D., & Vargas, J. F. (2012). Inequidad regional en Colombia. In A. Montenegro, & M. Meléndez (Eds.), *Equidad movilidad social diagnosticos y propuestas para la transformacion de la sociedad colombiana* (pp. 695–768). Universidad de los Andes volume 2012. URL: https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/Equidad_y_movilidad_social_web2.pdf.

Coulter, R. (2018). Parental background and housing outcomes in young adulthood. *Housing Studies*, *33*, 201–223.

Cowell, F. A., & Flachaire, E. (2007). Income distribution and inequality measurement: The problem of extreme values. *Journal of Econometrics*, *141*, 1044–1072.

Cuesta, J., & Negre, M. (2016). *Poverty and shared prosperity 2016: Taking on inequality*. Technical Report Washington, D.C. : World Bank Group. URL: http://documents.worldbank.org/curated/en/242251476706821424/Poverty-and-shared-prosperity-2016-taking-on-inequality. doi:10.1596/978-1-4648-0958-3.

Cuesta, J., Ñopo, H., & Pizzolitto, G. (2011). Using pseudo-panels to measure income mobility in latin america. *Review of Income and Wealth*, *57*, 224–246.

Dahan, M., & Gaviria, A. (2001). Sibling Correlations and Intergenerational Mobility in Latin America. *Economic Development and Cultural Change*, *49*, 537–554. URL: http://www.jstor.org/stable/10.1086/452514. doi:10.1086/452514.

De Grande, H., Vandenheede, H., & Deboosere, P. (2015). Educational inequalities in the transition to adulthood in belgium: The impact of intergenerational mobility on young-adult mortality in 2001-2009. *PloS one*, *10*.

Fernández, M., & Messina, J. (2018). Skill premium, labor supply, and changes in the structure of wages in latin america. *Journal of Development Economics*, *135*, 555 – 573. URL: http:

//www.sciencedirect.com/science/article/pii/S0304387818304905. doi:https://doi.org/10.1016/j.jdeveco.2018.08.012.

Ferreira, F. H. G., & Meléndez, M. (2012). *Desigualdad de resultados y oportunidades en Colombia: 1997-2010*. Technical Report.

Ferreyra, M. M., & Roberts, M. (2018). *Raising the Bar for Productive Cities in Latin America and the Caribbean*. Latin America and Caribbean Studies. The World Bank. URL: https://openknowledge.worldbank.org/handle/10986/29279.

Fields, G. S. (2008). Income mobility. *Cornell University Digital Commons ILR Collection*, . URL: http://digitalcommons.ilr.cornell.edu/articles?utm_source=digitalcommons.ilr.cornell.edu%2Farticles%2F453&utm_medium=PDF&utm_campaign=PDFCoverPages.

Fields, G. S. (2010). Does income mobility equalize longer-term incomes? New measures of an old concept. *The Journal of Economic Inequality*, *8*, 409–427.

Fields, G. S., Hernández, R. D., Freije, S., Puerta, M. L. S., Arias, O., & Assunção, J. (2007). Intragenerational income mobility in Latin America [with Comments]. *Economia*, *7*, 101–154.

Fields, G. S., Leary, J. B., & Ok, E. A. (2002). Stochastic dominance in mobility analysis. *Economics Letters*, *75*, 333–339.

Fields, G. S., & Ok, E. A. (1999). Measuring Movement of Incomes. *Economica*, *66*, 455–471. doi:10.1111/1468-0335.00183.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*, 953–973.

Firpo, S., Fortin, N. M., & Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics*, *6*, 28.

Galvis, L. A., & Meisel-Roca, A. (2012). *Convergencia y trampas espaciales de pobreza en Colombia: Evidencia reciente*. CEER.

Galvis-Aponte, L. A., & Meisel-Roca, A. (2016). Aspectos regionales de la movilidad social y la igualdad de oportunidades en colombia. *Revista de Economía del Rosario*, *17*, 257–297.

García, S., Rodriguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la cuna: La movilidad social a través de la educación en los municipios de Colombia, .

Gasparini, L., Bracco, J., Galeano, L., & Pistorio, M. (2018). Desigualdad en países en desarrollo: ajustando las expectativas" (2018). *Documentos de Trabajo del CEDLAS*, .

Gasparini, L., Cicowiez, M., & Sosa-Escudero (2014). *Pobreza y desigualdad en America Latina : Conceptos, herramientas y aplicaciones*. Editorial Temas, La Plata.

Gasparini, L., Cruces, G., & Tornarolli, L. (2016). Chronicle of a Deceleration Foretold Income inequality in Latin America in the 2010s. *Revista de economía mundial*, (pp. 25–45).

Gasparini, L., Cruces, G., Tornarolli, L., & Mejía, D. (2011). Recent trends in income inequality in latin america [with comments]. *Economía*, *11*, 147–201. URL: http://www.jstor.org/stable/41343452.

Gasparini, L., & Lustig, N. (2011). *The rise and fall of income inequality in Latin America*. CEDLAS, Working Papers CEDLAS, Universidad Nacional de La Plata.

Gasparini, L., & Tornarolli, L. (2009). Labor informality in latin america and the caribbean: Patterns and trends from household survey microdata. *Revista Desarrollo y Sociedad*, (pp. 13–80).

Gaviria, A. (2002). *Los que suben y los que bajan: educación y movilidad social en Colombia*. Fedesarrollo.

Gaviria, A. (2007). Social Mobility and Preferences for Redistribution in Latin America. *Economia*, *8*, 55–88.

Gontero, S., & Weller, J. (2015). Estudias o trabajas? el largo camino hacia la independencia económica de los jóvenes de américa latina. *Serie Macroeconomica del Desarrollo*, *169*.

Gradín, C. (2018). *Explaining cross-state earnings inequality differentials in India: An RIF decomposition approach*. Technical Report WIDER Working Paper.

Hahn-De-Castro, L. W., & Meisel-Roca, A. (2018). La desigualdad económica entre las regiones de colombia, 1926-2016. *Cuadernos de Historia Económica; No. 47*, .

Hampel, F. (1974). The influence curve and its role in robust estimation. *Journal of the american statistical association*, *69*, 383–393.

Hampel, F., Ronchetti, E., Rousseeuw, P., & Stahel, W. (2005). Robust statistics. hoboken.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the national academy of sciences*, *106*, 10570–10575.

Jantti, M., & Jenkins, S. P. (2015). Income Mobility. In A. B. Atkinson, & F. Bourguignon (Eds.), *Handbook of Income Distribution* chapter 10. (pp. 807–935). Elsevier Science B.V. (Vol. 2a ed.).

Lambert, S., Ravallion, M., & van de Walle, D. (2014). Intergenerational mobility and interpersonal inequality in an african economy. *Journal of Development Economics*, *110*, 327 – 344. URL: http:

//www.sciencedirect.com/science/article/pii/S0304387814000716. doi:https://doi.org/10.1016/j.jdeveco.2014.05.007. Land and Property Rights.

Levy, S., & Schady, N. (2013). Latin america's social policy challenge: Education, social insurance, redistribution. *Journal of Economic Perspectives*, *27*, 193–218.

Lora, E. (forthcoming). Forecasting formal employment in cities. *Desarrollo y Sociedad*, .

Luke, N., & Munshi, K. (2011). Women as agents of change: Female income and mobility in india. *Journal of Development Economics*, *94*, 1 – 17. URL: http://www.sciencedirect.com/science/article/pii/S0304387810000040. doi:https://doi.org/10.1016/j.jdeveco.2010.01.002.

Maasoumi, E., & Zandvakili, S. (1986). A class of generalized measures of mobility with applications. *Economics Letters*, *22*, 97–102.

Maasoumi, E., & Zandvakili, S. (1989). Mobility profiles and time aggregates of individual incomes. *Research on Economic Inequality*, *1*, 195–218.

Meisel-Roca, A. (2007). *¿ Por Qué Se Necesita Una Política Económica Regional En Colombia?*. Technical Report Banco de la Republica de Colombia.

Neidhöfer, G., Serrano, J., & Gasparini, L. (2018). Educational inequality and intergenerational mobility in Latin America: A new database. *Journal of Development Economics*, *134*, 329–349.

Nina, E., Grillo, S., & Malaver, C. (2003). Movilidad social y transmisión de la pobreza en Bogotá. *Economía y desarrollo*, *2*, 119–156.

O'Clery, N., Chaparro, J. C., Gómez-Liévano, A., & Lora, E. (forthcoming). How skill diversity leads to formal employment creation in cities. *Economic Development and Cultural Change*, .

OECD (2019). GINI index Income inequality (indicator). URL: https://data.oecd.org/inequality/income-inequality.htm. doi:doi:10.1787/459aa7f1-en.

Pérez, G. J., Aguilera, M., Otero, A., Sánchez, A., & Acosta, K. (2014). *Economía de las grandes ciudades en Colombia: seis estudios de caso* volume 1. Banco de la República, Colombia.

Poveda, A. C., & Martínez, C. I. P. (2011). Las tendencias de la pobreza y la desigualdad: una evidencia para los departamentos de colombia. *Ensayos Revista de Economía (Ensayos Journal of Economics)*, *30*.

Prais, S. J. (1955). Measuring social mobility. *Journal of the Royal Statistical Society. Series A (General)*, *118*, 56–66.

Rocha García, R., & Vivas Benitez, A. (1998). Crecimiento regional en colombia:¿ persiste la desigualdad? *Revista de economía del Rosario*, *1*, 67–108.

Roemer, J., & Trannoy, A. (2015). Equality of Opportunity. In A. B. Atkinson, & F. Bourguignon (Eds.), *Handbook of Income Distribution* chapter 4. (pp. 217–300). Elsevier Science B.V. (Vol. 2a ed.).

Schuck, B., & Shore, J. (2019). How intergenerational mobility shapes attitudes toward work and welfare. *The ANNALS of the American Academy of Political and Social Science*, *682*, 139–154.

Schuck, B., & Steiber, N. (2018). Does intergenerational educational mobility shape the well-being of young europeans? evidence from the european social survey. *Social indicators research*, *139*, 1237–1255.

Shorrocks, A. F. (1978a). Income inequality and income mobility. *Journal of Economic Theory*, *19*, 376–393.

Shorrocks, A. F. (1978b). The measurement of mobility. *Econometrica: Journal of the Econometric Society*, (pp. 1013–1024).

Slemrod, J. (1992). Taxation and inequality: a time-exposure perspective. *Tax policy and the economy*, *6*, 105–127.

Tornarolli, L., Battistón, D., Gasparini, L., & Gluzmann, P. (2014). *Exploring trends in labor informality in Latin America, 1990-2010*. Technical Report Documento de Trabajo.

Weller, J. (2012). Crecimiento, empleo y distribución de ingresos en américa latina. *Serie Macroeconomica del Desarrollo*, *122*.

Zwysen, W. (2015). The effects of father's worklessness on young adults in the uk. *IZA Journal of European Labor Studies*, *4*, 2.

Appendix A. Appendix

Appendix A.1. First Appendix

Metropolitan Area	Matching observations 2014 - 2018
Bogotá	565,432
Medellín	225,101
Cali	135,128
Barranquilla	74,662
Cartagena	45,954
Bucaramanga	66,366
Cúcuta	26,988
Pereira	32,935
Manizales	21,587
Valledupar	16,921
Pasto	14,305
Villavicencio	26,412
Ibagué	22,258
Montería	14,841
Tunja	11,051
Armenia	13,618
Santa Marta	17,785
Florencia	4,957
Popayán	13,169
Quibdó	3,272
Neiva	17,991
Riohacha	6,618
Sincelejo	8,271
Total	1,385,622

Figure A1: Location and total observations of twenty-three metropolitan areas in Colombia

Table Appendix A1: Characterization of Colombia's Metropolitan Areas

Name of Metropolitan Area	**Total Population Metropolitan Area (Data 2018)**	**Total Geographical Area Metropolitan Area (Km^2)**	**Participation in Country's GDP (Data 2016)**	**Density Pop/Km^2**
Bogota	9,389,321	3,498	29.96	2,684.2
Medellin	3,821,890	1,166	10.29	3,277.8
Cali	3,026,542	2,745	6.15	1,102.6
Barranquilla	2,049,949	512	4.07	4,003.8
Cartagena	1,013,389	559	2.40	1,812.9
Bucaramanga	1,132,339	1,417	2.38	799.1
Cucuta	863,463	2,036	1.17	424.1
Pereira	704,944	820	1.28	859.7
Manizales	454,849	957	0.85	475.3
Valledupar	564,733	8,073	0.84	70.0
Pasto	445,409	1,131	0.54	393.8
Villavicencio	495,227	1,328	1.02	372.9
Ibague	558,805	1,439	0.97	388.3
Monteria	447,668	3,043	0.51	147.1
Tunja	191,924	118	0.39	1,626.5
Armenia	298,199	115	0.52	2,593.0
Santa Marta	491,535	2,369	0.68	207.5
Florencia	175,407	2,292	0.20	76.5
Popayan	280,054	464	0.44	603.6
Quibdo	115,907	3,075	0.18	37.7
Neiva	344,026	1,468	0.66	234.4
Riohacha	268,712	3,171	0.27	84.7
Sincelejo	279,031	292	0.34	955.6

Own calculations based on information from the Colombian National Department of Statistics (DANE in Spanish).

Source: https://www.dane.gov.co/

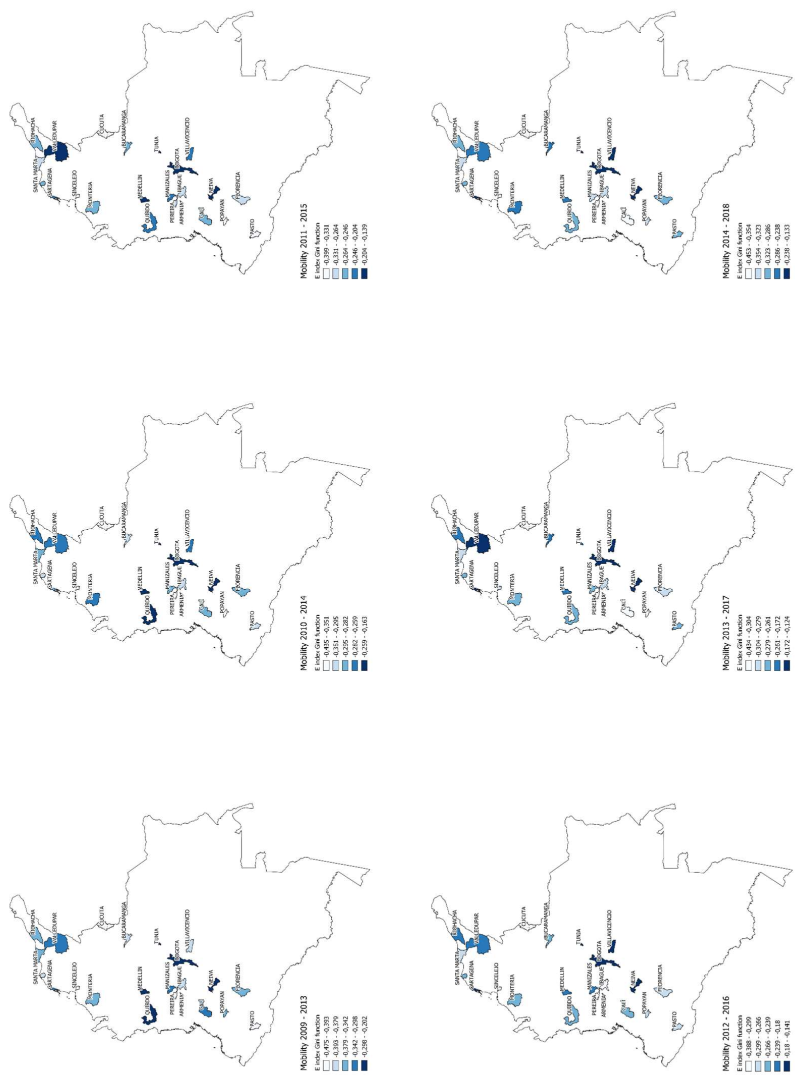

Figure A2: The E trends by metropolitan area with the Gini function

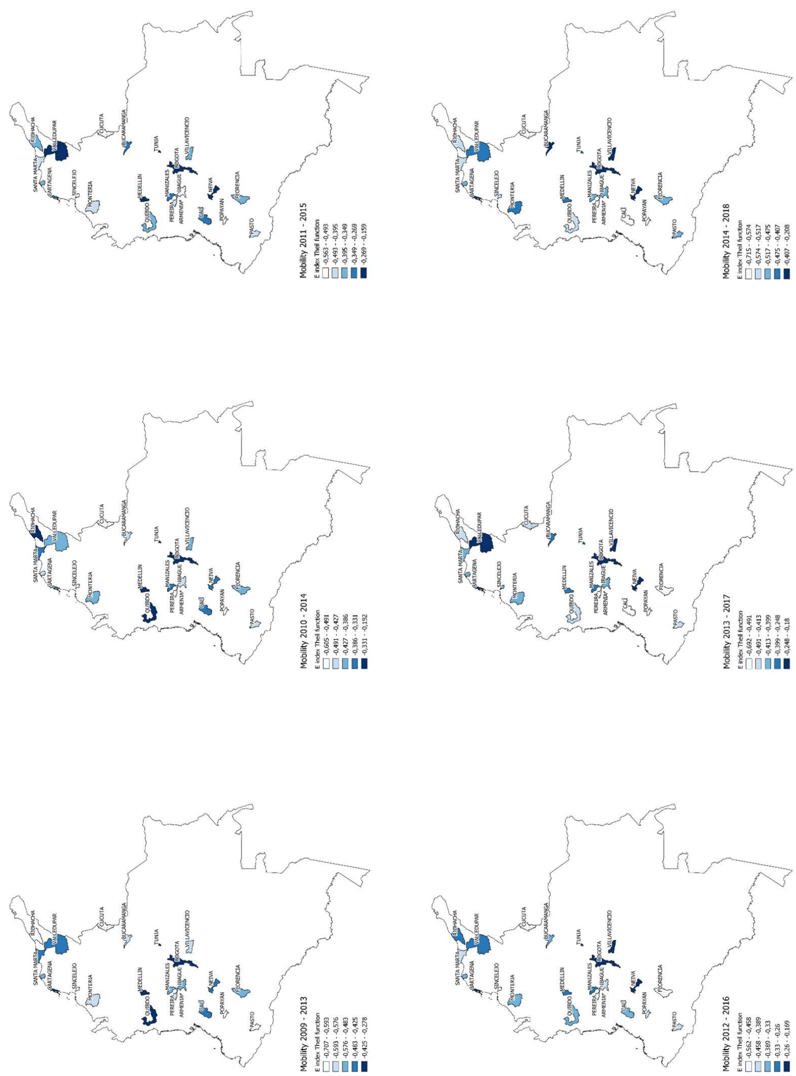

Figure A3: The E trends by metropolitan area with the Theil function

Appendix A.2. Second appendix

The Economic Complexity Index (ECI) is defined as follows:

$$ECI = \frac{\left(\vec{K} - <\vec{K}>\right)}{\left(stdev\vec{K}\right)} \tag{A.1}$$

where $<>$ represents an average, stdev stands for the standard deviation, and $\vec{K}$ = Eigenvector of $\overline{M}cc$, associated with the second largest eigenvalue. For more details see Hausmann et al. (2014), p. 24.

Appendix A.3. Third appendix

Equation 5 is derived from the following procedure. Following Firpo et al. (2009), the influence function is constructed as follows:

$$F_Y(y) = \int F_{Y|X}(y|X = x)dF_X(x) \tag{A.2}$$

where $F_Y(y)$ is the unconditional distribution function of Y, and covariates X. Then, assuming another distribution G_Y of the same class of F_Y they construct the mixing distribution F_{Y,tG_Y} which is t away from F_Y in the direction of the probability distribution G_Y. The influence function of the functional $v(F_Y)$ proposed by Hampel (1974); Hampel et al. (2005) is defined as follows:

$$IF(Y; v, F_Y) = \lim_{t \to 0} \frac{v(F_{Y,tGY}) - v(F_Y)}{t} = \lim_{t \to 0} \frac{v[(1-t)F_Y + \epsilon\delta_x] - v(F_Y)}{t} \tag{A.3}$$

where δ_x is a probability measure that puts mass 1 at the point $x \in R$.

Recalling equation 1 we denote a as the vector of average incomes, therefore, we call $v^G(F_a)$ for Gini and its RIF expression is as follows (Firpo et al., 2018),

$$RIF(a; v^G, F_a) = 1 + \frac{2}{\mu_a^2}R_a - \frac{2}{\mu_a}(a(1 - F_a(a))) \tag{A.4}$$

where $R_a = \int_0^1 GL_a(p)dp$ and $GL_a(p) = \int_{-\infty}^{q_a(p)} adF_a(a)$ or the generalized Lorenz curve.

We call $v^T(F_a)$ for Theil and its RIF expression is as follows (Cowell & Flachaire, 2007),

$$RIF(a; v^T, F_a) = (v^T, F_a) + \frac{1}{\mu_a}(a \ln a - w) - \frac{w + \mu_a}{\mu_a^2}(a - \mu_a) \tag{A.5}$$

where $(v^T, F_a) = \frac{w}{\mu_a} - \log \mu_a$ and $w = \int a \ln a dF_a(a)4$

Table Appendix A2: RIF estimates for the E index with mean and median functions

Variable	Mean		Median	
	Coef.	Std. Err.	Coef.	Std. Err.
Gender	35660.1	(2045.6)	15673.1	(948.5)
Age	7941.8	(540.5)	490.7	(267.3)
Secondary	19868.0	(2655.1)	44075.5	(1461.3)
T&T	153425.1	(4175.8)	124282.6	(2118.4)
Undergraduate or higher	587185.4	(5294.6)	189293.6	(1906.1)
Married	46776.4	(1972.8)	15145.5	(970.0)
Home property	40677.5	(2076.5)	13976.3	(990.5)
Overcrowding	-23585.2	(2274.5)	-13050.6	(1365.2)
Wall material	-1138.7	(3037.0)	4849.6	(1565.6)
Floor Material	37906.7	(1949.0)	14626.8	(1030.6)
Sewer service	-5666.3	(3177.3)	4535.2	(1695.6)
Gas service	3799.6	(2321.0)	3933.0	(1150.6)
Refrigerator	14487.5	(2186.2)	11649.5	(1251.5)
Washing Machine	35786.1	(2105.5)	14487.1	(1079.0)
Cable	13050.2	(2013.2)	5892.5	(1041.1)
Shower Heater	51053.9	(3461.7)	11767.1	(1408.4)
Microwave Oven	76338.1	(3993.4)	16076.1	(1516.6)
Independent	-164091.7	(2883.5)	-86259.6	(1505.4)
Cooperative	-98508.6	(2663.1)	-46372.1	(1514.9)
Manufacturing	18274.8	(4876.3)	31047.5	(2225.5)
Electricity; gas and water	56084.2	(18237.5)	17772.7	(7684.3)
Construction	-24717.3	(6135.0)	167.0	(3032.3)
Trade; hotels and food	-94717.7	(4476.2)	-29837.1	(2248.3)
Transportation and warehousing	825.4	(6405.2)	-5334.8	(2763.4)
Finance	128093.2	(6620.3)	49646.2	(2571.9)
Real estate	-37566.1	(3971.7)	-7378.0	(1818.3)
Community; social and personal services	-27197.8	(3937.1)	-4380.6	(1792.0)
Other sector	-61992.5	(4700.02)	-25458.5	(2193.25)
Constant	372479.3	(15869.2)	493896.9	(7915.9)
Observations	379,874		379,874	

Standard errors in parentheses

Table Appendix A3: RIF Estimates for the E index with p25 and p75 functions

Variable	p25		p75	
	Coef.	Std. Err.	Coef.	Std. Err.
Gender	42922.2	(1386.3)	26505.1	(2382.7)
Age	4644.5	(389.1)	3521.2	(659.8)
Secondary	46714.8	(2288.9)	55270.0	(3136.5)
T&T	95565.9	(3091.9)	302842.2	(5523.0)
Undergraduate or higher	141664.1	(2829.7)	653406.3	(5103.6)
Married	17580.4	(1414.7)	48571.8	(2391.7)
Home property	9857.8	(1430.3)	47836.8	(2496.3)
Overcrowding	-14274.8	(2059.5)	-36662.7	(3043.2)
Wall material	10221.7	(2321.9)	940.4	(3724.8)
Floor Material	11347.1	(1487.4)	43934.8	(2508.2)
Sewer service	6730.0	(2549.4)	-4537.6	(3971.4)
Gas service	-706.6	(1671.)	7919.1	(2806.7)
Refrigerator	22764.8	(1874.7)	24499.9	(2867.9)
Washing Machine	11547.4	(1541.5)	47053.0	(2689.4)
Cable	7072.3	(1505.3)	8519.9	(2558.9)
Shower Heater	7285.6	(1935.0)	53279.5	(3831.9)
Microwave Oven	7127.6	(2087.4)	72294.5	(4252.7)
Independent	-112503.1	(2413.5)	-176058.9	(3527.8)
Cooperative	-43270.1	(2273.4)	-111419.6	(3519.5)
Manufacturing	-9079.9	(3116.4)	83607.7	(5677.3)
Electricity; gas and water	-30040.6	(11173.5)	125904.1	(20160.5)
Construction	-48005.6	(4449.1)	28956.9	(7427.8)
Trade; hotels and food	-33378.5	(3230.4)	-67374.0	(5393.5)
Transportation and warehousing	-4191.9	(3828.9)	22526.5	(6992.2)
Finance	13693.8	(3564.9)	206734.0	(7029.0)
Real estate	-19109.7	(2586.5)	-18385.1	(4394.2)
Community; social and personal services	-26084.3	(2557.1)	7404.3	(4370.5)
Other sector	-47560.9	(3164.6)	-30847.0	(5343.7)
Constant	264429.6	(11569.6)	498193.9	(19388.0)
Observations	379,874		379,874	

Standard errors in parentheses