



ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

LII Reunión Anual

Noviembre de 2017

ISSN 1852-0022

ISBN 978-987-28590-5-3

Educational Inequality and Intergenerational
Mobility in Latin America: A New Database

Neidhøfer, Guido
Serrano, Joaquín
Gasparini, Leonardo

Educational Inequality and Intergenerational Mobility in Latin America: A New Database

Guido Neidhöfer* Joaquín Serrano Leonardo Gasparini

26th July 2017

Abstract. The causes and consequences of the intergenerational persistence of inequality are a topic of great interest among various fields in economics. However, until now, issues of data availability have restricted a broader and cross-national perspective on the topic. Based on rich sets of harmonized household survey data, we contribute to filling this gap computing time series for several indexes of relative and absolute intergenerational education mobility for 18 Latin American countries over 50 years, and making them publicly available. We find that intergenerational mobility has been rising in Latin America, on average. This pattern seems to be driven by the high upward mobility of children from low-educated families; at the same time, there is substantial immobility at the top of the distribution. Significant cross-country differences are observed and are associated with income inequality, poverty, economic growth, public educational expenditures and assortative mating.

JEL D63, I24, J62, O15. **Keywords** Inequality, Intergenerational Mobility, Equality of Opportunity, Transition Probabilities, Assortative Mating, Education, Human Capital, Latin America.

***Contacts and affiliations:** Guido Neidhöfer, Freie Universität Berlin, Boltzmannstr. 20, 14195 Berlin (guido.neidhoefer@fu-berlin.de), corresponding author. Joaquín Serrano, CEDLAS Universidad Nacional de La Plata & CONICET (jserrano@cedlas.org). Leonardo Gasparini, CEDLAS Universidad Nacional de La Plata & CONICET (leonardo@depeco.econo.unlp.edu.ar). We would like to thank Stephan Klasen, Martin Nybom, Oliver Winkler, Christian Ambrosius, and Viktor Steiner for their fruitful remarks, as well as participants of conferences and seminars in Berlin, Buenos Aires, Göttingen, La Plata, Kiel, Santa Fe, Oxford, and New York. The panel data set can be downloaded at <https://sites.google.com/site/mobilitylatam/>

1 Introduction

Among the oldest and most argued topics in economics are the causes and consequences of economic inequality. On the one hand, the high levels of inequality experienced in most developed and developing countries during the last century have attracted special attention by researchers and policy makers. On the other hand, there are differing views on the beneficial, detrimental or neutral impact of economic inequality and it remains difficult to generate clear causal statements regarding the way inequality affects economic performance and vice versa. For instance, influential theoretical models and empirical analyses on the relationship between inequality and growth have thus far yielded opposite results ([Banerjee and Duflo, 2003](#)).¹ Hence, scholars have argued that answers to questions related to inequality require taking into account that the observed level of income inequality displays both the rewards obtained by individuals for their efforts as well as the returns to given circumstances that they cannot control, like the socioeconomic status of their parents ([Roemer, 2000](#)).² Extending distributional analysis to the degree of intergenerational mobility in a society has arisen as a possible solution ([Corak, 2013](#)).

The evaluation of intergenerational mobility allows us to address one important question: for a given level of inequality, how likely is it that families persist at the top or bottom of the distribution over the course of time? Analyzing the subject across multiple countries and periods further helps us determine which factors are associated with this likelihood. However, comparing estimates for different countries that are derived from different studies raises the question of whether the uncovered cross-country differences are real or due to differences in data and measurement ([Solon, 2002](#)). Therefore, in order to deepen our understanding of the factors associated with the intergenerational transmission of socioeconomic status, it is necessary to study the subject in a harmonized framework.

¹See [Furman and Stiglitz \(1998\)](#) for an overview of the consequences of inequality for growth. [Neves et al. \(2016\)](#) recently reviewed the empirical literature on the inequality-growth nexus and performed a Meta-Analysis. Their results point at non-significant results on average with a high amount of heterogeneity in effect sizes across countries.

²For instance, [Marrero and Rodríguez \(2013\)](#) show that across U.S. states, inequality of opportunity has a negative impact on growth, while the impact of income inequality based on merit and effort is positive.

Furthermore, while large data sets with multiple and comparable measures of economic inequality and even historical time series are available for a multitude of countries, this is not the case for intergenerational mobility. The subject has been extensively analyzed within countries, for instance, for the United States ([Chetty et al., 2014b,a](#)) and India ([Azam and Bhatt, 2015](#)), but research on this topic still suffers from the lack of comparable estimates across multiple countries and over longer periods of time. Our study (and the associated database that we provide) contributes to filling this gap by estimating trends of relative and absolute intergenerational mobility for educational attainment in Latin America using novel sets of harmonized household survey data.

We provide a panel of comparable summary indicators for intergenerational education mobility in 18 countries over more than 50 years that we make available for future research. The present study aims to introduce this new data set and provide a comprehensive analysis of the observed trends regarding intergenerational mobility in Latin America, as well as their association with macroeconomic and institutional characteristics. It extends and builds upon [Hertz et al. \(2007\)](#)'s influential cross-country analysis on educational mobility as well as the existing evidence on intergenerational mobility in Latin America, as recently reviewed by [Torche \(2014\)](#). First, we examine more countries over a longer time span and in a harmonized framework. Second, we provide more precise estimates that rely on several survey waves and a greater number of observations. Third, we obtain estimates from two independent sources for nine of the 18 countries in our sample. Fourth, we compute several indexes that fulfill different axioms and measure different dimensions of relative and absolute mobility. Fifth, we calculate estimates for father-son and mother-daughter pairs, as well as for the degree of assortative mating. Finally, we provide resulting panel data for use in future research.

The paper is structured as follows: Section 2 describes the data sources and harmonization procedure used to obtain our estimates. Section 3 explains the applied methodologies. Section 4 presents and summarizes our results: First, it describes the uncovered cross-country patterns, trends, heterogeneity by gender, and degrees of assortative mating. Then, it examines the associa-

tion between our intergenerational mobility estimates and economic performance and institutional characteristics. Section 5 concludes.

2 Data

2.1 Description of Data Sources

The sources of information used to obtain our estimates are derived from two sets of harmonized household survey data. We used the availability of information on the parental educational background of adult individuals as a selection criteria for our surveys, focusing on surveys that include retrospective questions about parental education in the questionnaire. To avoid a so-called *co-residency bias*, we did not use surveys in which information on parental characteristics could only be retrieved because parents and children resided in the same household.³

The first harmonized survey data set is derived from the annual opinion survey *Latinobarómetro*. *Latinobarómetro* records individual and household characteristics of a nationally representative sample of adult respondents in 18 Latin American countries since 1995, including questions about own and parental education since 1998.⁴ The annual survey uses a sample of 1000 to 1200 individuals per country, representing more than 600 million inhabitants. It is carried out by local firms under technical supervision of the *Latinobarómetro* Corporation, a private non-profit organization based in Santiago (Chile).⁵ For the present study, we use the survey waves that include retrospective questions on parental education (1998 to 2015). The second data set is retrieved through an ex-post harmonization of selected national household surveys that are mainly conducted by national statistical offices. All estimates based on both data sets (henceforth *Latinobarómetro* and

³For a recent analysis of co-residency bias in intergenerational mobility estimates, see [Emran et al. \(2016\)](#).

⁴The Dominican Republic was included for the first time in 2004. The representativeness of the survey has varied over time reaching 100% of the total population in all countries around the year 2000.

⁵The study receives financing from Latin American and non-Latin American governments, the private sector, and international organizations. Among others: IADB (Inter-American Development Bank), UNDP (United Nations Development Program), AECI (Agencia Española de Cooperación Internacional), SIDA (Swedish International Development Cooperation Agency), CIDA (Canadian International Development Agency), CAF (Corporación Andina de Fomento), OAS (Organization of American States), United States Office of Research, IDEA International, UK Data Archive.

National Household Surveys) are obtained by weighting each observation by the inverse probability of selection, normalizing the weights over the different survey waves. All the surveys used in our analysis are listed and described in Appendix A (Supplemental Material).

One advantage of Latinobarómetro is that it is harmonized ex-ante and is specifically developed to be used in cross-country studies. The other household surveys are not uniform across Latin American countries. Therefore, we made all possible efforts to make statistics comparable across countries and over time by using similar definitions of variables in each country and survey year, and by applying consistent methods of processing the data. In particular, the inclusion of retrospective questions is not a universal characteristic found in all household surveys. Thus, while with the sample retrieved from Latinobarómetro we estimated the indexes for 18 countries, with the National Household Surveys estimates for 9 countries could be obtained. The advantage of many of the National Household Surveys is that they offer a substantially higher number of observations. Furthermore, the survey structure allows us to estimate father-son, and mother-daughter associations while Latinobarómetro only includes information on the parent with the highest educational degree.

2.2 Restriction criteria

We draw the same sample for each country and survey. The sample comprises individuals born between 1940 and 1990 who were at least 23 years old when surveyed. The age limit ensures that individuals have a higher likelihood to have completed their educational career, thus avoiding biased estimates. Since parental education is retrieved through retrospective questions, whether the individual and her parents reside together in the same household is not relevant for inclusion in our sample. The main restriction criteria is therefore the availability of information on own and parental education. Our final samples, including all countries and cohorts, is comprised of 198,949 individuals from the Latinobarómetro survey and 1,179,217 individuals from the National Household Surveys.

The amount of information about parental educational background that is missing is relatively small for Latinobarómetro—on average about 12% of all individuals in the survey with available information on own education. For some of the National Household Surveys the number is much higher, ranging from 2 % in Guatemala to 61 % in Peru and 83 % in Brazil. In order to prove if selectivity issues bias our intergenerational mobility estimates, we compare the average years of education of all individuals in the household survey with a sample of individuals for whom we have information regarding parental educational background. Differences are negligible in both data sets, counting at most 0.4 years of schooling, and in most countries not statistically significant. Furthermore, no clear pattern hints at a specific direction of a possible selectivity bias (e.g. for Peru, the average of the sample used to compute our estimates is 0.2 years lower than the unrestricted sample, while for Brazil the mean of our sample is 0.4 years higher).

2.3 Measurement of educational attainment

In Latinobarómetro the information recorded regarding parental education refers only to the parent with highest education among the two. In the National Household Surveys, the education of both parents, mother and father, is provided. In that case, we use the parent with the highest educational degree, as is most commonly done in the literature ([Black and Devereux, 2011](#)), to obtain our baseline estimates.

In order to improve the comparability of the completed years of education, which is our main result variable, we use the same coding used by Latinobarómetro to process the National Household Survey. That is, we truncate the years of education at the university level because the degree of heterogeneity is greater at that level. Thus, completed years of education range from 0 to 15. Furthermore, Latinobarómetro uses the same variable to measure the education of individuals and their parents. Most other surveys record years of formal education for individuals who are interviewed, but are not as precise for data regarding parents. In those cases, we impute the years of education

required to complete the obtained degree and follow the same scheme used in the Latinobarómetro survey.⁶

Figure 1 shows the mean and coefficient of variation of completed years of education in our samples, comparing the statistics obtained from Latinobarómetro and the National Household Surveys. The cohorts always refer to the children's generation. It is evident that in most countries the two harmonized survey sets yield very similar statistics in trends and levels. Throughout the cohorts, educational attainment of individuals in Latin America increased steadily, while there is certain heterogeneity in the levels of schooling among countries. In the youngest cohort, we find Guatemala, Honduras, and Nicaragua, with around six years of education on average; on the other end of the spectrum we find Argentina, Chile, and Colombia, with around 12 years.

In order to give an idea of how educational attainment is related to economic well-being, Figure 2 shows the mean income levels for six broad educational categories and the returns to education – measured by the ratio of incomes achieved by high and low educated people – for two different cohorts in each Latin American country. This analysis helps to read our intergenerational education mobility estimates and put the results in the right context.⁷ We see that, although substantial differences between countries exist, higher educational degrees are clearly associated with higher level of income. Furthermore, despite the educational expansions experienced in all countries, returns to education are rather similar for people of different ages. Thus, apart from the intrinsic value of educational mobility as one of the drivers of human development, our measures are also meaningful indicators for intergenerational mobility of (material) well-being.

3 Estimated Mobility Indexes

Pioneering works by [Becker and Tomes \(1979\)](#) and [Solon \(1992\)](#) conceptualize the mechanisms and transmission channels that explain the observed degree of persistence between the economic

⁶Detailed information on the codification of educational attainment for parents and children in each country is available in the Supplemental Material.

⁷As shown by [Blanden \(2013\)](#), there is a meaningful correlation between estimates of intergenerational income mobility and educational mobility across countries.

outcomes of parents and children. However, especially in cross country comparisons, different indexes measuring intergenerational mobility may yield very different pictures. Researchers should therefore adopt the measurement which fulfill the needs of the dimension they aim to analyze and the questions they seek to answer.⁸

In the context of educational mobility, some questions might need absolute mobility measures, as would be the case to capture educational expansions (structural mobility). Others might need to neglect this dimension and focus on positional changes of families within the distribution (exchange mobility). In this study, and with the creation of the associated database, we try to offer an exhaustive panorama of absolute and relative indexes and show the overall picture of intergenerational mobility in Latin America from different angles.

Future research using our estimates should use the indexes which fit the requirements of the research question regarding two key aspects: i) what is the intuition behind the phenomena that has to be analyzed, and ii) which axioms have to be fulfilled. In what follows, we describe the computed indexes. The key variables are always referring to educational outcomes of parents (y^p) and children (y^c) measured either in completed years of education or the obtainment of a certain educational degree. The indexes are estimated for each cohort j and country k separately.⁹

3.1 Slope coefficient and intergenerational correlations

The most widely used mobility index in the intergenerational mobility literature is the slope coefficient from a linear regression of children's on parents' outcomes.¹⁰ Here, we regress the years of

⁸For conceptual and methodological reviews on intergenerational mobility, see [Black and Devereux \(2011\)](#); [Jäntti and Jenkins \(2013\)](#); [Piketty \(2000\)](#).

⁹[Neidhöfer \(2016\)](#) develops a method to transform the educational outcomes of parents and children in a way that makes them more appropriate as a proxy measure for socioeconomic status and more comparable across time (see also [Neidhöfer and Stockhausen, 2016](#)). Here, this correction is not necessary since the analysis is performed for each cohort separately. Proper methods are applied to standardize the estimated coefficients ex-post, as explained below.

¹⁰The specification of the model displayed here simplifies to one child per family.

education of the child from family i belonging to cohort j in country k on the years of education of his parent with the highest educational attainment among the two:

$$y_{ijk}^c = \alpha_{jk} + \beta_{jk} \cdot y_{ijk}^p + \gamma_{jk} X_{ijk} + \varepsilon_{ijk}. \quad (1)$$

In this equation, α is a constant, X is a vector of control variables for age and sex and ε is the error term. The slope coefficient can also be standardized to take differences in the distributions of children's and parents' outcomes into account:

$$r_{jk} = \beta_{jk} \frac{\sigma_{jk}^p}{\sigma_{jk}^c}. \quad (2)$$

If no control variables are included in the regression, the standardization yields an index equal to Pearson's correlation coefficient.

β and r are measures for positional mobility that capture both dimensions, structural mobility as well as exchange mobility, and reflect the degree of regression to the population mean between two generations. Its wider use in the literature has the advantage of comparability between these and other estimates for the same or other countries. Hereby, r "corrects" β by the changes in inequality in the marginal distributions of the outcome of interest. Scholars still argue about which of the two is more suitable for cross-country (and cross-cohort) comparisons (see [Jäntti and Jenkins, 2013](#)). Therefore, it seems important to report both.

An index which fully controls for the marginal distributions – and not only for the changes in inequality – and captures the pure positional change aspect of mobility, is Spearman's rank correlation coefficient:

$$\rho_{jk} = \frac{\text{cov}(\text{rank}_{jk}^c, \text{rank}_{jk}^p)}{\sigma_{jk, \text{rank}}^c \sigma_{jk, \text{rank}}^p}. \quad (3)$$

Whether these corrections are necessary or not depends on the research question. As stated before, the intergenerational transmission of inequality could be an important dimension and it

may get lost if one measures mobility by (2) and (3). However, if exchange mobility is the only important aspect to be accounted for, (1) might not be the suitable index to rely on.

The outcome that is most often available for two subsequent generations and is also comparable across countries is educational attainment measured in completed years of education. The indexes thus have one important feature in common: they give a broad and intuitive picture of the overall educational persistence experienced by a certain cohort in a given country.¹¹

3.2 Transition probabilities

Another insightful measure in terms of intergenerational mobility is the probability of children facing different circumstances, measured by parental educational background, to afford a certain minimum level of education. We compute two different indicators:

The *probability of bottom upward mobility*

$$BUM_{jk} = Prob(y_{ijk}^c \geq s | y_{ijk}^p < s), \quad (4)$$

and the *probability of upper class persistence*

$$UCP_{jk} = Prob(y_{ijk}^c \geq s | y_{ijk}^p \geq s). \quad (5)$$

The indicators yield the probabilities of children to achieve at least a secondary educational degree (s) conditional on their parents' education. Parent's education is hereby measured by two different types: i) low parental education, i.e. less than completed secondary education. ii) high

¹¹These measures assume a linear and monotonic relationship of years of education from one generation to the next. Although this method is usually applied in the literature, the validity of the linearity assumption has been questioned since the slope might vary with rising parental education. So far, linear and non-linear measures has been found to be correlated across countries (see [Blanden, 2013](#)), but future research on this topic should investigate this issue in more detail. For completeness, in the Supplemental Material we include an analysis of the correlation between the educational level of parents and children measured in categories using a bivariate ordered probit model. Equation (1) might be also estimated on the logarithm of the outcome of interest, i.e. years of education, hence assuming a log-linear relationship. In this case, the slope coefficient is an elasticity measuring marginal changes in children's education associated with marginal changes in their parent's education. The intuitive difference between the educational persistence explained above and the intergenerational education elasticity (not discussed in this paper but included in the database) lies mainly in the functional form assumed to underlie the intergenerational transmission of education and social status.

parental education, i.e. at least a secondary school degree. In terms of social mobility and equality of opportunity these probabilities measure upward mobility for people at the bottom of the distribution and class persistence at the top, respectively.

3.3 Absolute and directional mobility

The measures described above cover the relative and absolute dimensions of intergenerational mobility, understood as the movement of families within the distribution over time. However, they do not give comparable information about the size of those movements. Two more indexes – initially developed by Fields (1996) and mostly applied to measure individual income movements in an intragenerational context – are therefore computed to measure the per capita movements in years of education:

$$M1_{jk} = \frac{1}{N_{jk}} \sum_{i=1}^{N_{jk}} |y_{ijk}^c - y_{ijk}^p|. \quad (6)$$

$$M2_{jk} = \frac{1}{N_{jk}} \sum_{i=1}^{N_{jk}} (y_{ijk}^c - y_{ijk}^p), \quad (7)$$

$M1$ shows the average difference between the two generations within the same families, regardless of the direction of the change. Upward and downward movements are summed up to one summary measure. In contrast, $M2$ measures the average directional change between two generations. High values of $M2$ can, for example, be a sign of educational expansion. Together, $M1$ and $M2$ also give insightful information on the degree of downward movements: The smaller is the difference between the two, the lower is the amount, or average degree, of downward mobility.

4 Results: Intergenerational Mobility in Latin America

4.1 Cross-Country Patterns

Before reporting the intergenerational mobility trends through the summary measures described in Section 3, we describe the cross-country differences in mobility patterns for the entire sample.

First, Figure 3 illustrates absolute (or structural) mobility patterns, and, then, Figure 4 illustrates relative (or exchange) mobility; both using Latinobarómetro as data source. Tables 1, 2 and 3 show descriptive statistics of the summary measures described in Section 3 for each country and the Latin American average using both data sources.

Figure 3 ranks countries in Latin America according to the percentage of people who have more education than their parents, measured in completed years of schooling. We see that more than 50% of people born between 1940 and 1990 in all countries in the region have achieved higher educational attainment than their parents. Venezuela and Paraguay lead the group of countries with high absolute mobility, while Guatemala, Nicaragua, and Honduras are at the bottom end of the ranking. Although this evidence is illustrative of the differences between countries in terms of mobility, it is far from complete because it does not take into account the position of individuals in the distribution and the size of the change between generations.

Figure 4 is more informative about the movement of families within the distribution. In the upper part, a transition matrix for Latin America is displayed. Here, individuals and their parents are ranked according to their relative educational position, measured in standard deviations from the country's average years of education, and grouped in three different classes: high, middle, and low levels of education. The cells of the transition matrix contain the percentage of individuals in the children's generation associated with the respective parental educational class. Complete intergenerational mobility is displayed by equal entries in each cell of a transition matrix. As has been argued in past, under certain circumstances complete mobility can be understood as equality of opportunity.¹²

We see that the Latin American reality is far from achieving complete mobility. Focusing on the three most meaningful cells of the transition matrix – the ones that display persistence at the top and at the bottom of the distribution, as well as the degree of bottom-up mobility – Latin America appears to be a region with low intergenerational mobility, on average. Almost 60% of children with high and low education, respectively, have parents in the same educational class. Moreover,

¹²For an exhaustive discussion of conceptual differences between intergenerational mobility and equality of opportunity, see [Roemer \(2004\)](#).

only 14% of the individuals in the high education class come from low-education families. The lower part of Figure 4 ranks the countries by this last indicator for bottom-up mobility. We see that the share ranges from less than 10% in Chile to about 20 % in Nicaragua and Dominican Republic. To give a benchmark for these estimates, we compute transition matrices for the U.S. and Germany using the same sample restriction criteria and comparable household surveys (PSID and SOEP, respectively). It turns out, that in these two countries persistence at the bottom is higher than the Latin American average (USA 61.5 %, Germany 56.5 %). In contrast, persistence at the top is lower (USA 51.2 %, Germany 55.8 %) and bottom-up mobility higher (USA 21.5 %, Germany 17.8 %) than in most Latin American countries.

It is worth noting that the country rankings change considerably depending on the adopted concept of mobility (relative or absolute). For example, it is particularly striking that Nicaragua is both one of the countries with the highest relative mobility and the lowest absolute mobility. What explains this seemingly controversial finding is that Nicaragua is one of the countries with the lowest and most unequally distributed educational attainments on average. Hence, while the opportunities of children from low educated families to improve their educational level are high, the chances that this improvement translates into a considerable jump within the distribution are quite modest. This finding confirms the importance of i) evaluate intergenerational mobility adopting multiple measures and ii) to measure the mobility of people born in different year spans separately.

4.2 Trends

Figures 5, 7 and 9 show the trends and geography of intergenerational mobility in Latin America measured by the seven indexes explained in Section 3 with the Latinobarómetro survey. Figures 6, 8 and 10 show the corresponding averages for the nine countries where we have National Household Surveys available to perform the analysis. Since the trends and levels obtained with the National Household Surveys basically mirror the results obtained with Latinobarómetro for all the estimated indexes, we will restrict the descriptive analysis in this section mainly to the results obtained with Latinobarómetro. Furthermore, we exclude point estimates obtained from less than 200 individual

observations. Charts for each country with both surveys are included in Appendix C (Supplemental Material).

Figure 5 and 6 show intergenerational mobility measured by the regression coefficient (β), the standardized coefficient (r) and the Spearman's rank correlation coefficient (ρ). Aggregate results for Latin America are constructed as the unweighted average of the 18 or 9 countries analyzed, depending on whether Latinobarómetro or National Household Surveys were used. β changes substantially and significantly over the observed period. For people born in the 1940s, an additional year of parental education is associated with an average increase of about 0.6 years of education, while for people born in the 1980s the same measure is around 0.4.¹³ Comparing these trends with the ones observed for other countries, we see that, while Latin America has historically been perceived as one of the regions with the least social mobility worldwide, the educational mobility of the youngest cohorts is on similar levels as developed countries like the U.S. and Germany (see Hertz et al., 2007; Neidhöfer and Stockhausen, 2016). The map shows that this increase was recorded for almost all Latin American countries. In contrast, r and ρ are relatively stable around 0.5 over the entire period. This shows that the type of mobility experienced in Latin America has mainly been structural. However, in the two countries where the rise in intergenerational mobility has been the strongest, Dominican Republic and Venezuela, both structural as well as exchange mobility increased significantly. Guatemala and Honduras are the only countries where structural as well as exchange mobility did not rise over the observation period.

Figures 7 and 8 illustrate the extent and differences across cohorts of the probability of upward mobility for people at the bottom of the distribution, as well as the probability of class persistence at the top. On average, the predicted probability of upper class persistence is high and oscillates around 0.7. By contrast, the predicted probability that individuals who were born in the 1980s to low-educated parents attain a secondary school degree is more than twice as high as the same

¹³The results for the older cohorts are consistent with past estimates, e.g. by Hertz et al. (2007). Because of surviving bias associated with own and parental education the sample of older individuals that participate in household surveys might be selective. Hence, intergenerational persistence estimates of the cohorts 1940 to 1950 might be upwardly biased by differential mortality rates among low and highly educated people. Furthermore, the strength of this bias might depend on cross-country characteristics like the extensiveness and quality of the health system.

probability for individuals born in the 1940s. However, not all countries show the same pattern. Although in most of the countries bottom-up mobility increased – up to a 300 % increase in Brazil and Mexico – it is on low levels and almost unchanged over time in Central American countries, like Guatemala, Honduras and Nicaragua.¹⁴ Very high bottom-up mobility rates in the youngest cohorts (higher than 0.5) are observed in Argentina, Mexico, Peru, and Venezuela. One striking finding is that in Nicaragua, the youngest cohorts of individuals show a surprisingly low probability of attaining a secondary school degree. This applies even to people with a high parental educational background. One possible explanation for this finding could be the violent wars suffered by the country from 1978 to 1990, which affected the people born in this age interval.

Figures 9 and 10 show absolute and directional mobility trends. These measures show the magnitude and pattern of the change between the educational attainment of parents and children on average. As is evident, since the outcome measure – completed years of education – is bounded, rising parental education also reduces the margins and possibilities for the children to experience an improvement. This fact explains the inverted U-shape pattern of the time series for these two indexes. In the sixties, the distance between parents' and children's education reaches a maximum and later decreases as parents' education rises. Interestingly, the gap between M1 and M2 does not change significantly across cohorts, showing that downward mobility is almost stable around one year of schooling on average.

4.3 Heterogeneity by Gender and Assortative Mating

In this part of the analysis, we first disentangle our estimates by father-son and mother-daughter lineages. These estimates provide an overview of how social, cultural or institutional factors may influence the educational mobility of men and women differently. For instance, families might dedicate more resources to the education of male offspring, either because the returns on sons' education are expected to be higher, or because of traditional gender roles. For this last reason, imitation might cause the educational attainments of children to be related more strongly to the

¹⁴The spatial dimension of this phenomenon is a striking finding that might deserve special attention in future studies.

education of the parent with the same sex (see e.g. [Schneebaum et al., 2015](#)). Then, we relate our intergenerational mobility estimates to the degree of assortative mating, i.e. the likelihood of people with similar socioeconomic status to marry each another. This analysis is particularly interesting since there seems to be a fundamental interrelation between the two concepts; e.g. because higher spouse correlations are argued to cause a stronger heritability of unobserved and observable endowments. However, few studies have been able to empirically prove this relationship so far (e.g. [Chadwick and Solon, 2002](#); [Ermisch et al., 2006](#); [Guell et al., 2015](#)). We can perform this evaluation for nine countries where we have information on both the father's and the mother's educational attainment.

As shown in [Figure 11](#), the estimates for father-son and mother-daughter pairs show the same trend and are rather similar for younger cohorts. Coinciding with the expansion of educational attainment among women, the mobility of daughters also rises considerably and approaches the mobility levels experienced by sons, on average. Generally, the patterns confirm the picture of rising intergenerational mobility in Latin America driven by high upward mobility from the bottom and with substantial immobility at the top of the distribution.

Taking into account the high degree of assortative mating in Latin American countries, these findings are not particularly surprising: when the education of both parents is similar, the education of only one of the two is a valid proxy for the education of the other. Our findings show that assortative mating in Latin America, measured by the correlation of father's and mother's educational attainment, is constantly high (around 0.7, with countries ranging between 0.6 and 0.8; see [Figure 12](#)). Interestingly, most countries show a slight but decreasing trend. Indeed, past research found an inverse relationship between assortative mating and intergenerational mobility ([Guell et al., 2015](#)).

We test the relationship between assortative mating and intergenerational mobility using our database, regressing the seven estimated mobility indexes on the estimated degree of spouse correlation in the parent's generation controlling for cross country heterogeneity by fixed effects. As shown in [Table 4](#), the degree of spouse correlation is positively and significantly associated with educational persistence (measured by the regression coefficient, the correlation coefficient and the

rank correlation) and negatively associated with the index of bottom upward mobility. The relationship with the index for upper class persistence and the measures of directional and absolute mobility point at the same picture – higher spouse correlation associated with lower intergenerational mobility – but are not statistically significant. Hence, our findings confirm a clear association between assortative mating and intergenerational mobility.

4.4 Intergenerational Mobility, Institutions and Economic Performance

The aim of this part of the analysis is to show the association between intergenerational mobility and macroeconomic and institutional characteristics. The first descriptive part of this analysis is focused on the regression coefficient as an indicator of intergenerational persistence because this indicator comprises both structural as well as exchange mobility. In the second part, all estimated indexes are evaluated separately in models that control for cross-country heterogeneity. In order to make use of all the available data, we take the average of our educational persistence estimates from three broader cohorts (people born 1940-54, 1955-69 and 1970-84) and associate them with data at the country level averaged over three time periods (1990-99, 2000-09 and 2010-14), respectively. The criteria on how to match the two data sets is thereby completely determined by the time periods for which information is available. Hence, these correlations cannot be interpreted as causal effects. Yet, they might be seen as a first step to understand potential underlying mechanisms.

Figure 13 and Figure 14 show scatter-plots, linear fits and the related correlation coefficients. We find that higher degrees of intergenerational mobility are associated with: i) High levels of household income per capita and GDP per capita. ii) Lower levels of income inequality and poverty. iii) Lower returns to education, as measured by the ratio of hourly wages of people with high and low education. iv) Higher amounts of public expenditure in education and, in particular, the share of expenditure devoted to primary education. These findings confirm the predictions of influential theoretical models (Becker and Tomes, 1979) and the patterns uncovered in empirical findings in the past.¹⁵

¹⁵For instance, the negative relationship between inequality and intergenerational mobility has been shown to hold within the U.S. (Chetty et al., 2014a) and China (Fan et al., 2015), as well as across and within Latin American

Finally, we regress a series of macroeconomic outcomes separately using our intergenerational mobility estimates controlling for country fixed effects. The association of our estimates and the macroeconomic outcomes is performed as explained above. Figure 15 shows the estimated standardized coefficients and their respective confidence interval. We see that the two mobility indicators that capture the structural mobility component, educational mobility (β) and bottom upward mobility (BUM), are positively associated with economic growth and well-being, and negatively with poverty and inequality. The standardized persistence estimates (r) have a qualitatively similar association with the above mentioned macroeconomic outcomes, however they are not statistically significant. A possible interpretation of these findings is that what positively influences economic performance is not the amount of exchange mobility – the rise of some families that is necessarily accompanied by the fall of other families – but the opportunities for children from the lower bottom of the distribution to improve their human capital as compared to their parents. Furthermore, since absolute ($M1$) and directional mobility ($M2$) – i.e. the magnitude of the change from one generation to the next – show no meaningful association, the strength of the structural mobility component seems to be less influential than the marginal improvement of human capital in itself. Last, the probability of upper class persistence (UCP) shows the same pattern of conditional correlation with economic performance as the BUM . This is mainly due to the fact that the two measures are highly correlated: changes in the probability of attaining a secondary education degree, like reforms that raise levels of compulsory education, are likely to affect all individuals regardless of their parental background. Including both as independent variables in the regressions, the coefficients of BUM are significantly different from zero, while the coefficients of UCP are not.

These preliminary analyses using our database open up interesting avenues for future research. Especially because of the temporal structure of the associations, the potential mechanisms behind the statistical relationships shown here must be understood as either implying a steady-state rela-

countries (Neidhöfer, 2016). Güell et al. (2015) find that intergenerational mobility within Italy is positively correlated with economic performance. It is argued that one of the primary mechanisms that cause this relationship is investment in children’s human capital in the presence of credit constraints. Rising private and public investments in the human capital of poor children, driven by economic growth, anti-poverty programs or public educational expenditures, thus leads to higher intergenerational mobility. For a survey of the theoretical explanations of the underlying mechanisms, see (Neidhöfer, 2016).

tionship or as an indication of the effect of intergenerational mobility on economic performance and institutions. For instance, a mechanism driving the latter might operate through preferences for redistribution that have been recently shown to be positively associated with perceptions about social mobility (see [Alesina et al., 2017](#)). A more suitable way to analyze the driving forces of social intergenerational mobility would be to relate a cohort's level of mobility with indicators of its initial conditions, as in [Neidhöfer \(2016\)](#). The exact identification of causal channels goes beyond the scope of this work. Nevertheless, the dataset created here makes it possible for these aspects to be analyzed in greater detail in the future.

5 Conclusions

In this paper, we introduced a new panel data set of intergenerational mobility estimates for Latin America and provided a comprehensive descriptive analysis of observed trends and patterns. We found that intergenerational mobility of educational attainment has been on the rise in Latin America, driven by the educational expansions of the last decades that have particularly benefited children from the bottom of the distribution. In contrast, the educational persistence at the top of the distribution has remained consistently high and has not changed substantially. Furthermore, we found intergenerational mobility to be positively associated with economic growth and progressive public expenditure in education, and negatively associated with income inequality, poverty, returns to education, and the degree of assortative mating. The positive relationship between intergenerational mobility and economic performance was also found in estimations controlling for cross-country heterogeneity by fixed effects.

The strength of our analysis is that it provides highly comparable estimates of educational mobility for people born over a span of over 50 years and in multiple countries, extending the influential work by [Hertz et al. \(2007\)](#). In the future, these estimates can be used to analyze the characteristics that influence or are influenced by the degree of intergenerational mobility of socioeconomic status. For instance, in the context of developing countries, key aspects include: the intergenerational

transmission of poverty, the impact of educational expansions and social programs on equality of opportunity, and the role played by institutions.

In our view, the data set is useful for at least one important reason: equality of opportunity and social mobility seem to be common goals for policy makers, as well as among egalitarians and utilitarians. Hence, our panel provides an essential tool for discussions and future research on the topic, at both the cross country and within country levels.

References

- ALESINA, A., S. STANTCHEVA AND E. TESO, “Intergenerational Mobility and Preferences for Redistribution,” , 2017.
- AZAM, M. AND V. BHATT, “Like Father, Like Son? Intergenerational Educational Mobility in India,” *Demography* 52 (2015), 1929–1959.
- BANERJEE, A. V. AND E. DUFLO, “Inequality and Growth: What Can the Data Say?,” *Journal of Economic Growth* 8 (2003), 267–299.
- BECKER, G. S. AND N. TOMES, “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility,” *Journal of Political Economy* 87 (1979), 1153–1189.
- BLACK, S. E. AND P. J. DEVEREUX, “Recent developments in intergenerational mobility,” volume 4 of *Handbook of Labor Economics*, chapter 12 (Elsevier B.V., 2011), 1487–1541.
- BLANDEN, J., “Cross-Country Rankings in Intergenerational Mobility: a Comparison of Approaches From Economics and Sociology,” *Journal of Economic Surveys* 27 (2013), 38–73.
- CHADWICK, L. AND G. SOLON, “Intergenerational Income Mobility Among Daughters,” *American Economic Review* 92 (mar 2002), 335–344.

- CHETTY, R., N. HENDREN, P. KLINE AND E. SAEZ, “Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States,” *The Quarterly Journal of Economics* 129 (nov 2014a), 1553–1623.
- CHETTY, R., N. HENDREN, P. KLINE, E. SAEZ AND N. TURNER, “Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility,” *American Economic Review Papers and Proceedings* 104 (2014b), 141–147.
- CORAK, M., “Income Inequality, Equality of Opportunity, and Intergenerational Mobility,” *Journal of Economic Perspectives* 27 (aug 2013), 79–102.
- EMRAN, M., W. GREENE AND F. SHILPI, “When Measure Matters: Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries,” , 2016.
- ERMISCH, J., M. FRANCESCONI AND T. SIEDLER, “Intergenerational Mobility and Marital Sorting”,,” *The Economic Journal*, 116 (2006), 659–679.
- FAN, Y., J. YI AND J. ZHANG, “The Great Gatsby Curve in China : Cross-Sectional Inequality and Intergenerational Mobility,” , 2015.
- FIELDS, G. S., “The Meaning and Measurement of Income Mobility *,” 377 (1996), 349–377.
- FURMAN, J. AND J. E. STIGLITZ, “Economic consequences of income inequality,” *Income Inequality: Issues and Policy Options, Jackson Hole, Wyoming, Federal Reserve Bank of Kansas City* (1998).
- GÜELL, M., M. PELLIZZARI, G. PICA AND J. V. RODRIGUEZ MORA, “Correlating Social Mobility and Economic Outcomes,” Centre for Economic Policy Research, 2015.
- GUELL, M., J. V. RODRIGUEZ MORA AND C. I. TELMER, “The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating,” *The Review of Economic Studies* 82 (apr 2015), 693–735.

- HERTZ, T., T. JAYASUNDERA, P. PIRAINO, S. SELCUK, N. SMITH AND A. VERASHCHAGINA, “The inheritance of educational inequality: International comparisons and fifty-year trends,” *The BE Journal of Economic Analysis & Policy* 7 (2007).
- JÄNTTI, M. AND S. P. JENKINS, “Income Mobility,” (2013).
- MARRERO, G. A. AND J. G. RODRÍGUEZ, “Inequality of opportunity and growth,” *Journal of Development Economics* 104 (sep 2013), 107–122.
- NEIDHÖFER, G., “Intergenerational Mobility and the Rise and Fall of Inequality : Lessons from Latin America,” CEDLAS, 2016.
- NEIDHÖFER, G. AND M. STOCKHAUSEN, “Dynastic Inequality Compared Multigenerational Mobility in the US , the UK , and Germany,” Freie Universität Berlin, 2016.
- NEVES, P. C., Ó. AFONSO AND S. T. SILVA, “A Meta-Analytic Reassessment of the Effects of Inequality on Growth,” *World Development* 78 (2016), 386–400.
- PIKETTY, T., “Theories of persistent inequality and intergenerational mobility,” *Handbook of income distribution* 1 (2000), 429–476.
- ROEMER, J. E., *Equality of Opportunity* (Harvard University Press, 2000).
- , “Equal opportunity and intergenerational mobility: going beyond intergenerational income transition matrices,” in M. Corak, ed., *Generational Income Mobility in North America and Europe* (Cambridge University Press, 2004).
- SCHNEEBAUM, A., B. RUMPLMAIER AND W. ALTZINGER, “Gender and migration background in intergenerational educational mobility,” *Education Economics* (2015), 1–23.
- SOLON, G., “Intergenerational income mobility in the United States,” *The American Economic Review* (1992).

———, “Cross-country differences in intergenerational earnings mobility,” *The Journal of Economic Perspectives* 16 (2002), 59–66.

TORCHE, F., “Intergenerational Mobility and Inequality: The Latin American Case,” *Annual Review of Sociology* 40 (2014), 619–642.

6 Tables

Table 1: Descriptive Statistics: Regression and Correlation Coefficients.

Panel A – Source: Latinobarometro, own estimates.									
	Regression coeff.				Correlation coeff.				
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.	
Argentina	0.44	0.16	0.32	0.54	0.51	0.06	0.46	0.56	
Bolivia	0.54	0.14	0.40	0.64	0.55	0.04	0.51	0.60	
Brazil	0.56	0.21	0.38	0.74	0.50	0.08	0.44	0.59	
Chile	0.49	0.11	0.42	0.56	0.62	0.10	0.54	0.79	
Colombia	0.54	0.16	0.38	0.72	0.54	0.07	0.50	0.63	
Costa Rica	0.41	0.12	0.34	0.49	0.42	0.07	0.36	0.47	
Dominican Rep.	0.44	0.27	0.33	0.65	0.42	0.17	0.34	0.57	
Ecuador	0.54	0.10	0.47	0.63	0.53	0.06	0.48	0.58	
El Salvador	0.62	0.19	0.43	0.81	0.56	0.09	0.48	0.63	
Guatemala	0.58	0.08	0.49	0.65	0.51	0.07	0.45	0.56	
Honduras	0.54	0.09	0.44	0.63	0.47	0.10	0.40	0.54	
Mexico	0.38	0.21	0.29	0.53	0.40	0.12	0.35	0.48	
Nicaragua	0.43	0.14	0.32	0.56	0.42	0.11	0.36	0.50	
Panama	0.49	0.12	0.42	0.59	0.51	0.06	0.43	0.56	
Paraguay	0.55	0.14	0.40	0.70	0.52	0.08	0.43	0.60	
Peru	0.51	0.20	0.39	0.70	0.56	0.05	0.51	0.64	
Uruguay	0.48	0.12	0.41	0.58	0.49	0.06	0.42	0.53	
Venezuela	0.39	0.21	0.31	0.60	0.42	0.11	0.36	0.52	
Latin America	0.50	0.15	0.39	0.63	0.50	0.08	0.44	0.57	

Panel B – Source: National Household Surveys, own estimates.									
	Regression coeff.				Correlation coeff.				
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.	
Brazil	0.59	0.27	0.37	0.84	0.51	0.08	0.44	0.58	
Chile	0.40	0.26	0.26	0.57	0.51	0.09	0.43	0.59	
Colombia	0.60	0.18	0.42	0.76	0.52	0.07	0.49	0.62	
Ecuador	0.61	0.13	0.51	0.73	0.59	0.05	0.55	0.64	
Guatemala	0.80	0.10	0.66	0.92	0.63	0.04	0.60	0.67	
Mexico	0.46	0.20	0.35	0.63	0.53	0.09	0.48	0.66	
Nicaragua	0.65	0.18	0.50	0.80	0.53	0.11	0.44	0.59	
Panama	0.56	0.16	0.45	0.73	0.59	0.06	0.54	0.67	
Peru	0.55	0.30	0.32	0.80	0.54	0.11	0.45	0.64	
Latin America	0.58	0.20	0.43	0.75	0.55	0.08	0.49	0.63	

Notes: Mean, coefficient of variation (C.V.), minimum and maximum values of the complete time series for the respective country.

Table 2: Descriptive Statistics: Upper Class Persistence and Bottom Upward Mobility.

Panel A – *Source*: Latinobarometro, own estimates.

	Upper class persistence				Bottom-Up Mobility			
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.
Argentina	0.84	0.06	0.71	0.91	0.38	0.25	0.21	0.58
Bolivia	0.81	0.09	0.69	0.90	0.26	0.43	0.12	0.46
Brazil	0.76	0.11	0.55	0.84	0.27	0.44	0.11	0.48
Chile	0.85	0.05	0.79	0.94	0.37	0.17	0.28	0.49
Colombia	0.78	0.09	0.65	0.88	0.28	0.36	0.11	0.42
Costa Rica	0.65	0.12	0.50	0.74	0.22	0.23	0.13	0.30
Dominican Rep.	0.52	0.24	0.32	0.71	0.25	0.34	0.10	0.37
Ecuador	0.78	0.15	0.54	0.88	0.31	0.36	0.12	0.43
El Salvador	0.81	0.11	0.61	0.90	0.19	0.35	0.08	0.28
Guatemala	0.67	0.11	0.57	0.77	0.14	0.26	0.09	0.20
Honduras	0.71	0.12	0.58	0.86	0.14	0.18	0.11	0.18
Mexico	0.63	0.20	0.45	0.91	0.36	0.42	0.15	0.66
Nicaragua	0.62	0.16	0.45	0.79	0.16	0.29	0.06	0.21
Panama	0.78	0.06	0.70	0.89	0.36	0.20	0.23	0.42
Paraguay	0.80	0.07	0.69	0.91	0.25	0.32	0.16	0.40
Peru	0.86	0.07	0.73	0.93	0.42	0.24	0.24	0.56
Uruguay	0.70	0.07	0.62	0.79	0.23	0.12	0.17	0.28
Venezuela	0.61	0.34	0.25	0.84	0.35	0.34	0.15	0.54
Latin America	0.73	0.12	0.58	0.85	0.27	0.29	0.15	0.40

Panel B – *Source*: National Household Surveys, own estimates.

	Upper class persistence				Bottom-Up Mobility			
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.
Brazil	0.85	0.07	0.71	0.92	0.36	0.39	0.15	0.55
Chile	0.82	0.10	0.66	0.92	0.45	0.40	0.17	0.71
Colombia	0.83	0.08	0.71	0.91	0.34	0.43	0.12	0.56
Ecuador	0.77	0.12	0.53	0.86	0.25	0.43	0.06	0.41
Guatemala	0.79	0.11	0.61	0.87	0.12	0.44	0.04	0.21
Mexico	0.78	0.11	0.63	0.94	0.24	0.31	0.09	0.35
Nicaragua	0.58	0.27	0.31	0.80	0.13	0.39	0.05	0.19
Panama	0.79	0.05	0.71	0.83	0.30	0.27	0.16	0.40
Peru	0.88	0.03	0.82	0.92	0.41	0.27	0.19	0.57
Latin America	0.79	0.10	0.63	0.89	0.29	0.37	0.11	0.44

Notes: Mean, coefficient of variation (C.V.), minimum and maximum values of the complete time series for the respective country.

Table 3: Descriptive Statistics: Absolute and Directional Mobility.

Panel A – *Source:* Latinobarometro, own estimates.

	Absolute mobility (M1)				Directional mobility (M2)			
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.
Argentina	3.4	0.08	2.7	3.6	2.8	0.14	1.9	3.2
Bolivia	4.3	0.12	3.3	4.8	3.6	0.17	2.5	4.3
Brazil	4.0	0.14	2.9	4.5	3.3	0.17	2.2	3.9
Chile	3.4	0.12	2.7	3.9	2.8	0.16	1.8	3.2
Colombia	4.0	0.11	2.9	4.5	3.1	0.15	2.2	3.7
Costa Rica	3.9	0.09	3.5	4.5	2.8	0.25	1.6	3.8
Dominican Rep.	4.4	0.14	3.3	5.0	3.3	0.19	2.4	4.1
Ecuador	3.8	0.11	3.2	4.4	3.1	0.17	2.2	3.9
El Salvador	4.0	0.14	3.0	4.6	3.4	0.12	2.5	3.9
Guatemala	3.2	0.10	2.6	3.6	2.0	0.16	1.5	2.5
Honduras	3.5	0.09	3.2	3.9	2.7	0.17	2.0	3.3
Mexico	4.3	0.08	3.6	4.8	3.1	0.11	2.5	3.6
Nicaragua	3.9	0.15	2.8	4.7	2.7	0.23	1.7	3.6
Panama	4.2	0.09	3.4	4.8	3.5	0.19	2.1	4.3
Paraguay	3.8	0.06	3.4	4.3	3.2	0.11	2.8	4.0
Peru	4.1	0.10	3.5	4.6	3.3	0.18	2.5	4.0
Uruguay	3.2	0.11	2.6	3.6	2.3	0.26	1.3	2.9
Venezuela	4.4	0.11	3.7	5.2	3.8	0.16	2.7	4.5
Latin America	3.9	0.11	3.1	4.4	3.0	0.17	2.1	3.7

Panel B – *Source:* National Household Surveys, own estimates.

	Absolute mobility (M1)				Directional mobility (M2)			
	Mean	C.V.	Min.	Max.	Mean	C.V.	Min.	Max.
Brazil	4.6	0.17	3.2	5.6	4.1	0.20	2.4	5.1
Chile	3.9	0.13	3.0	4.4	3.1	0.20	1.9	3.7
Colombia	4.0	0.15	2.7	4.5	3.1	0.25	1.5	3.8
Ecuador	3.4	0.14	2.5	3.8	2.4	0.35	0.6	3.1
Guatemala	2.7	0.26	1.6	3.7	1.9	0.45	0.6	3.1
Mexico	4.3	0.11	3.4	5.0	3.6	0.17	2.9	4.6
Nicaragua	3.2	0.21	2.1	4.0	2.2	0.34	0.9	2.9
Panama	3.4	0.11	3.0	4.1	2.5	0.23	1.5	3.4
Peru	4.5	0.13	3.1	5.0	3.8	0.20	2.0	4.5
Latin America	3.8	0.16	2.7	4.5	3.0	0.26	1.6	3.8

Notes: Mean, coefficient of variation (C.V.), minimum and maximum values of the complete time series for the respective country.

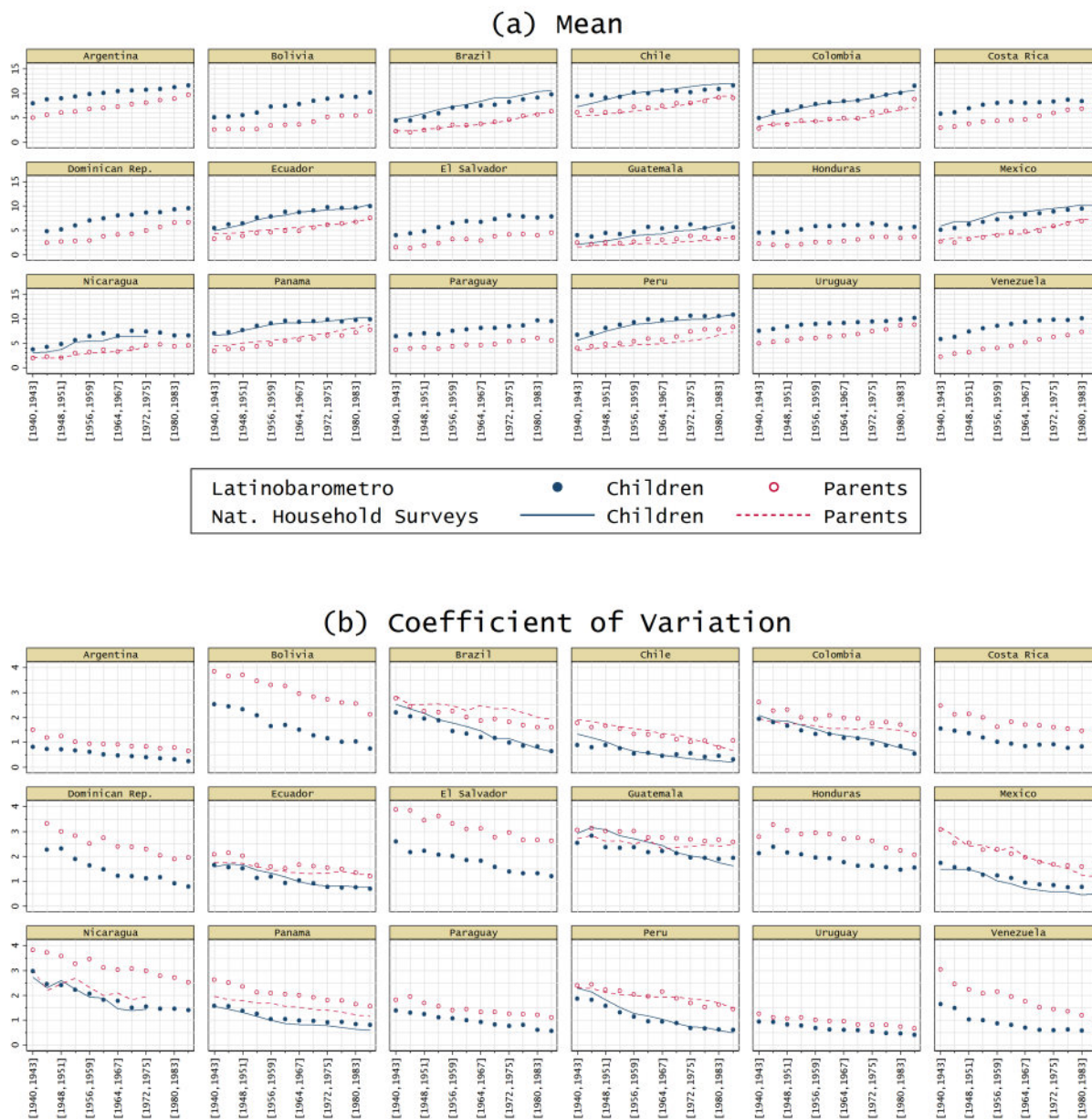
Table 4: Assortative mating and intergenerational mobility – Linear Regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	β	ρ	r	BUM	UCP	$M1$	$M2$
Spouse correlation (parents)	0.921*** (0.3416)	0.444*** (0.1479)	0.131** (0.0650)	-1.028*** (0.3710)	0.178 (0.2549)	-0.648 (1.7139)	-0.830 (1.8928)
Country F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	105	105	105	105	105	105	105

Notes: Table shows the coefficients of the computed spouse correlation index in linear regressions using the single mobility indexes as dependent variables. All regressions include country dummies. Robust standard errors in parentheses. Statistical significance level * 0.1 ** 0.05 *** 0.01. *Source:* National Household Surveys 1994-2015, own estimates.

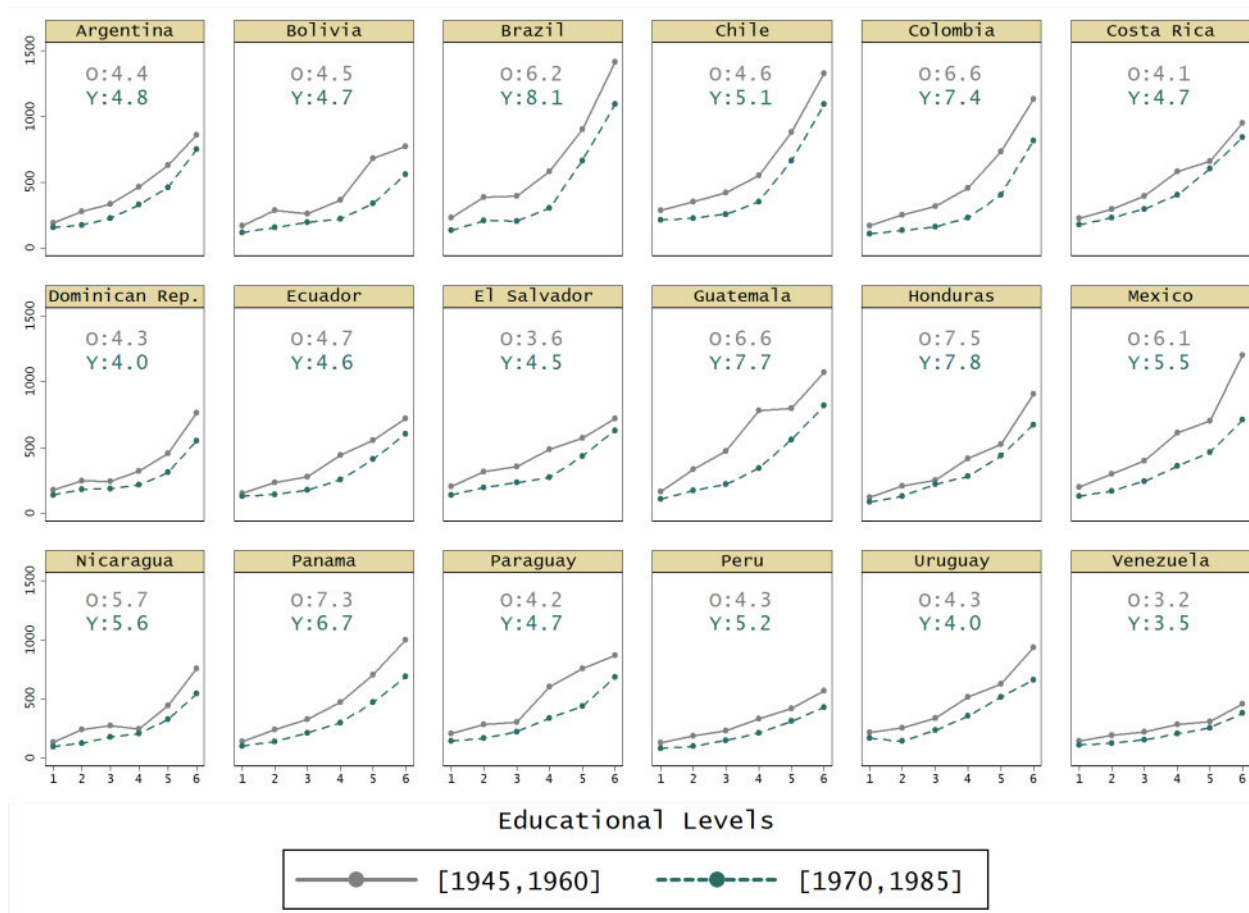
7 Figures

Figure 1: Completed years of education. Sample means and coefficients of variation by cohorts.



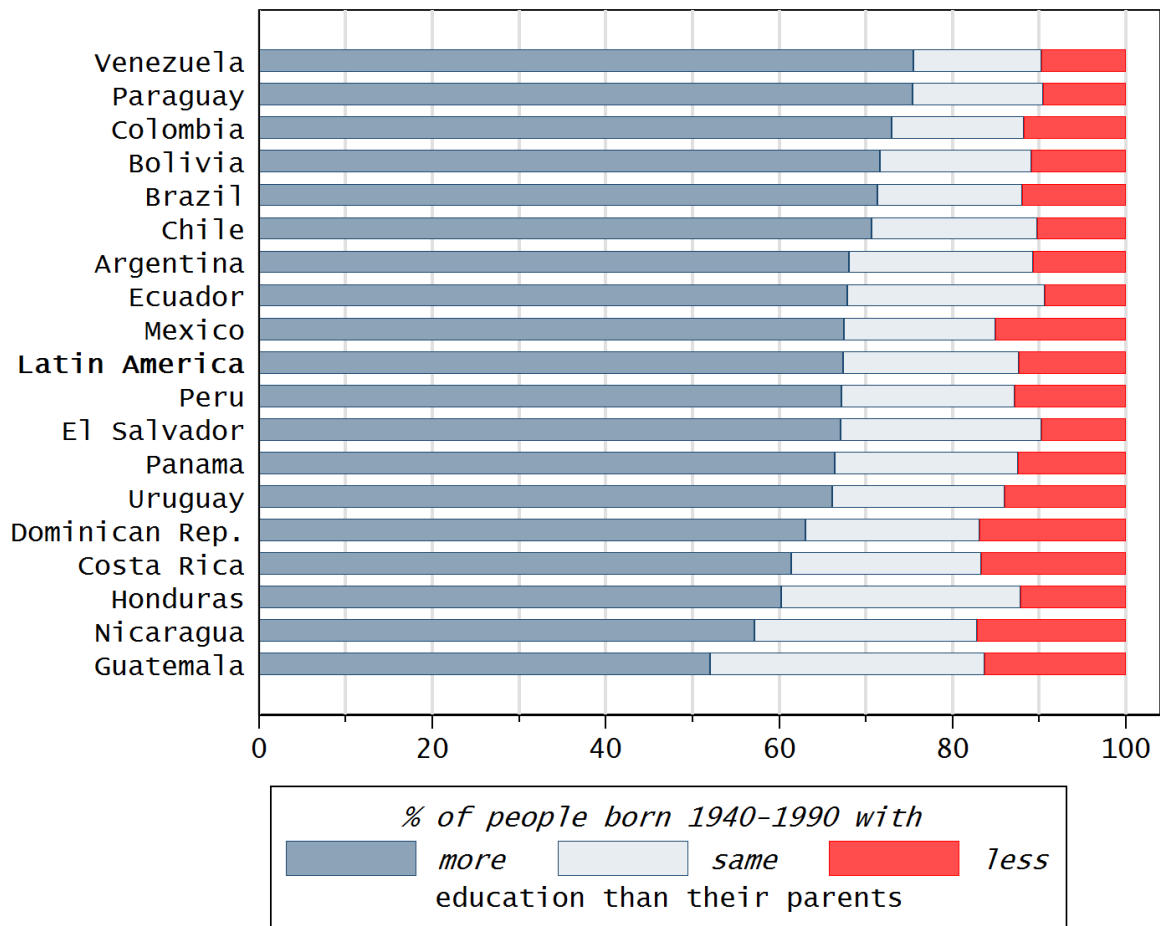
Notes: Cohorts refer to the year of birth of the children. *Source:* Latinobarometro 1998-2015, National Household Surveys 1994-2015.

Figure 2: Education as indicator for well-being: average income by educational level.



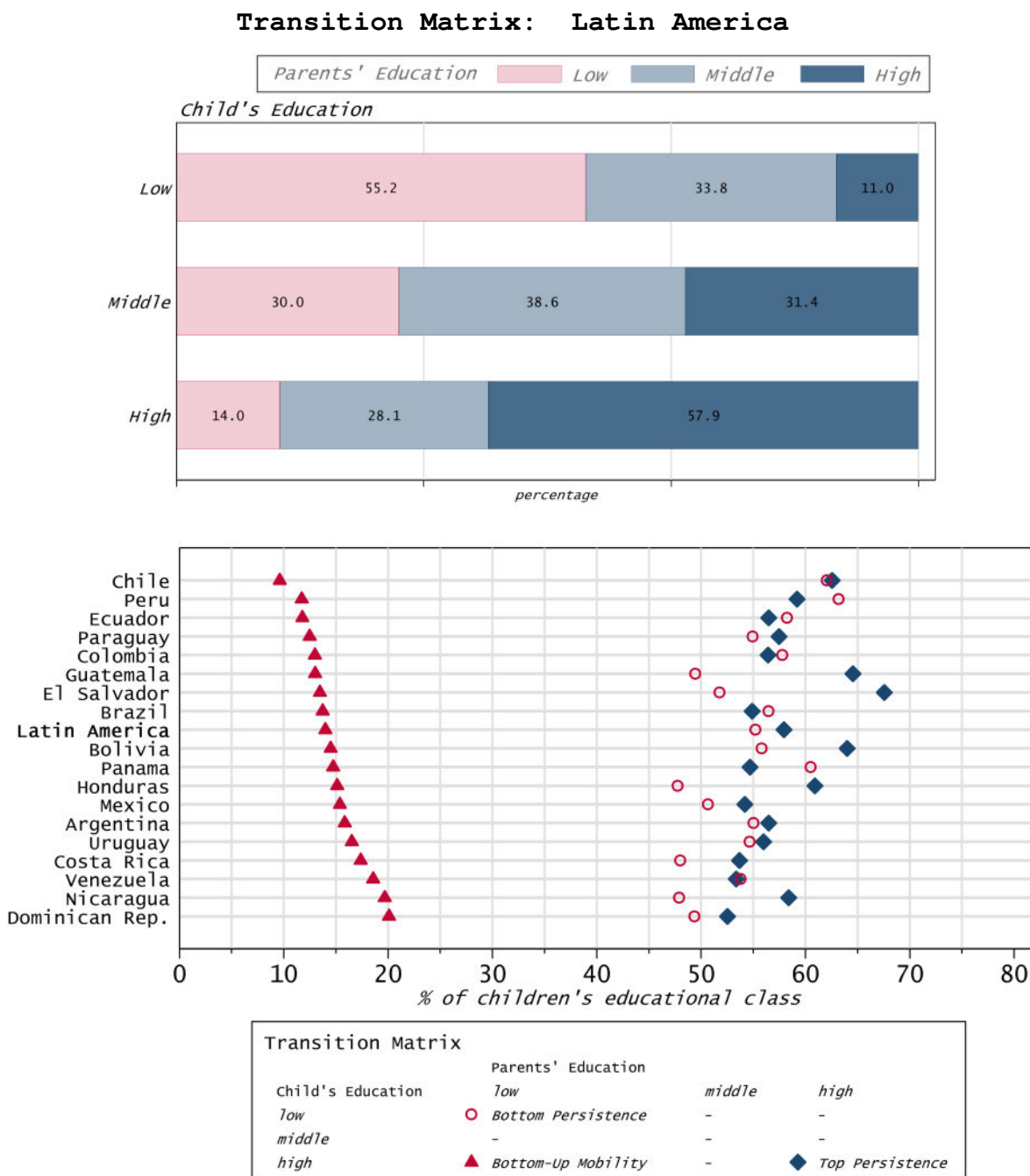
Notes: Average household per capita income (constant 2005 PPP international USD). Educational levels: 1 without education or primary incomplete; 2 primary complete; 3 secondary incomplete; 4 secondary complete; 5 tertiary incomplete; 6 tertiary complete. Numbers show the ratio of the monetary returns to education for people with a completed tertiary degree (category 6) and without education or with incomplete primary education (category 1). O: Older Cohort. Y: Younger Cohort. Example on how to read this numbers: In Argentina, individuals with completed tertiary degree born between 1945 and 1960 have a 4.4 times higher average household per capita income than their peers without education or with incomplete primary education. *Source:* SEDLAC circa 2005, own estimates.

Figure 3: Absolute educational mobility in Latin America.



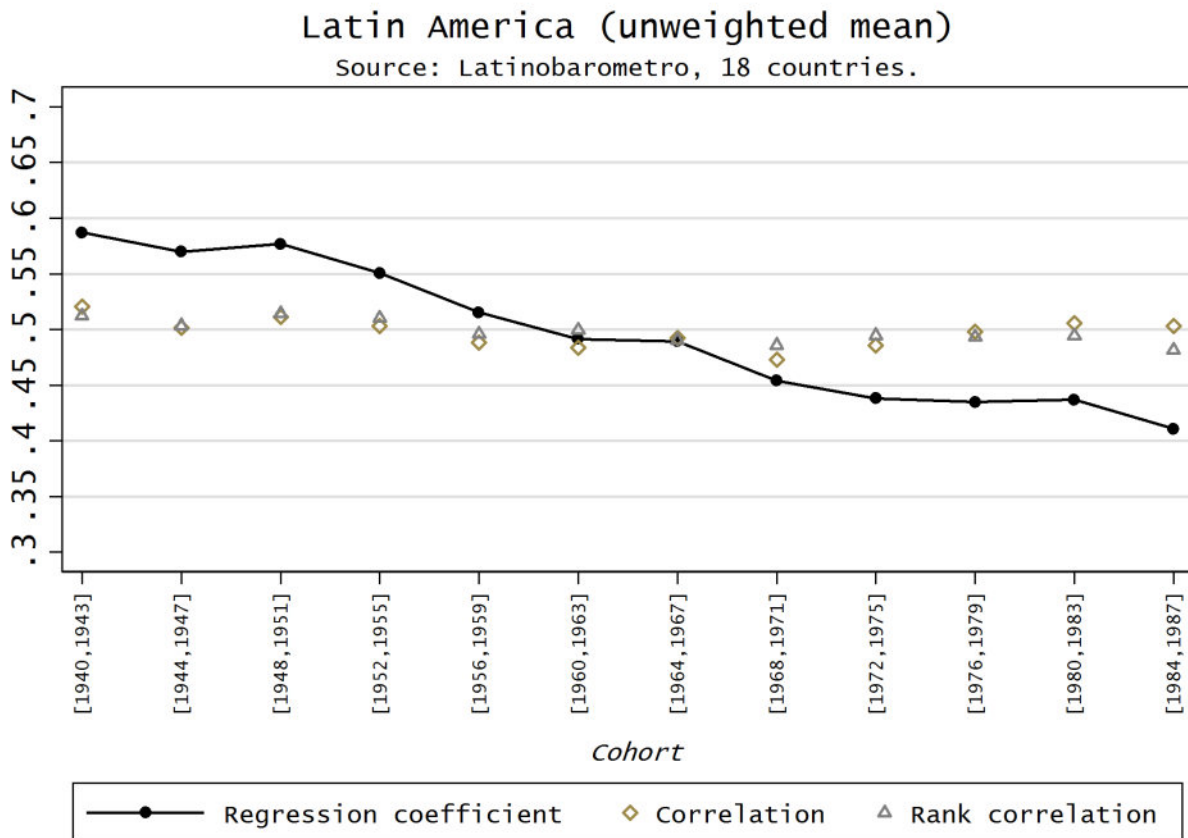
Notes: Education measured in completed years of education. *Source:* Latinobarometro 1998-2015, own estimates.

Figure 4: Educational persistence in Latin America: Insights from transition matrices (People born 1940-1990).

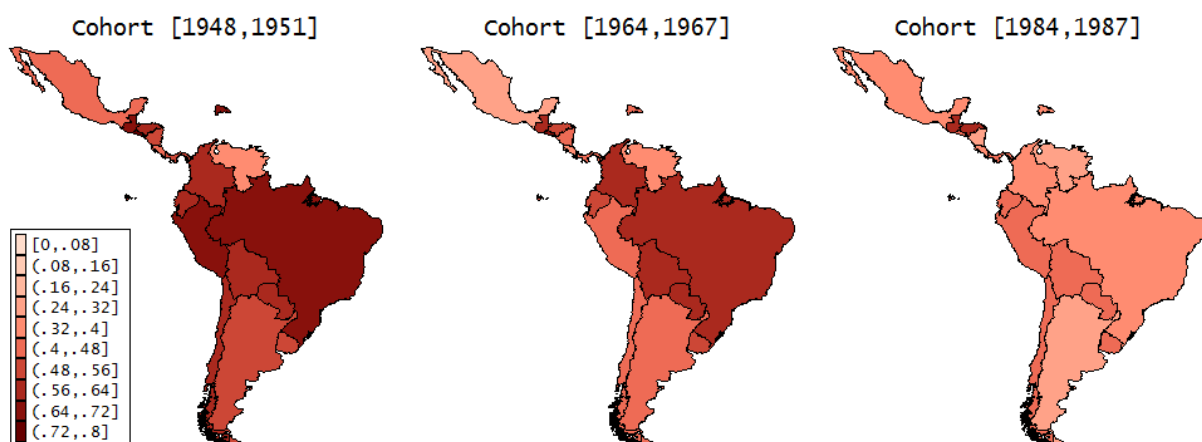


Notes: The points show the percentage of individuals in three different cells of the transition matrix. *Bottom persistence:* Individuals with low education and low parental education. *Bottom-Up Mobility:* Individuals with high education and low parental education. *Top persistence:* Individuals with high education and high parental education. Educational classes (low, middle, high) refer to three quantiles of the within-country and within-cohort distributions. Benchmarks USA (PSID, own estimates) / Germany (SOEP, own estimates): *Bottom persistence* 61.5 % / 56.5 %, *Top persistence* 51.2 % / 55.8 %, *Bottom-up mobility* 21.5 % / 17.8 %. Source: Latinobarometro 1998-2015, own estimates.

Figure 5: Educational persistence in Latin America: Regression and correlation coefficients.

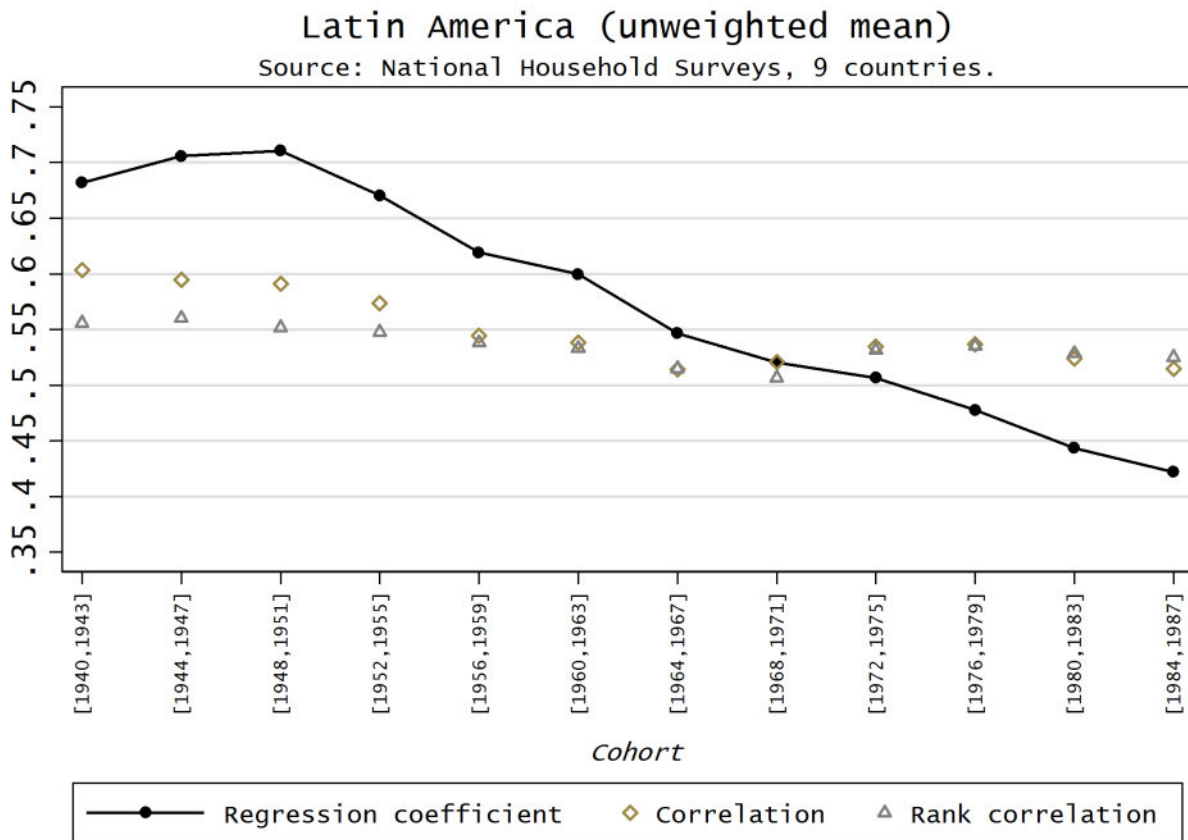


Regression coefficient: Geography and Trends for Latin America



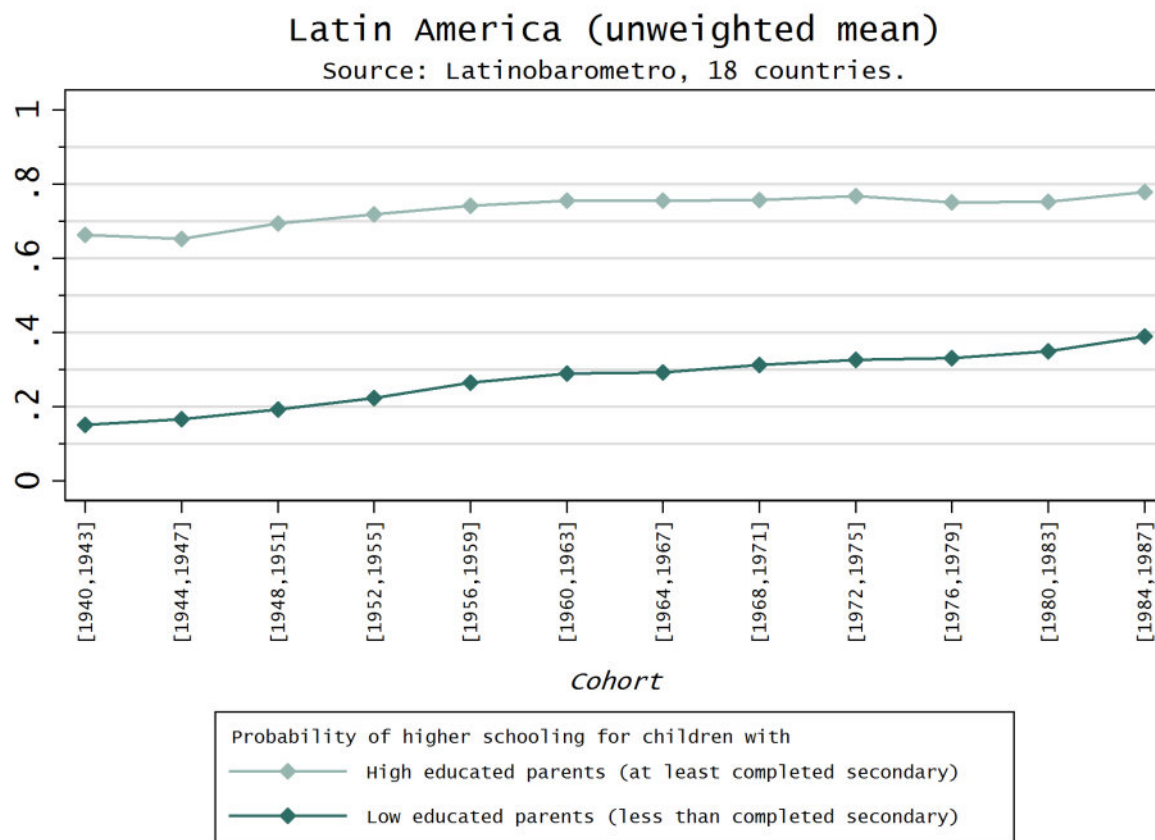
Notes: Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. Bootstrapped confidence interval. *Source:* Latinobarometro 1998-2015, own estimates.

Figure 6: Educational persistence in Latin America: Regression and correlation coefficients.

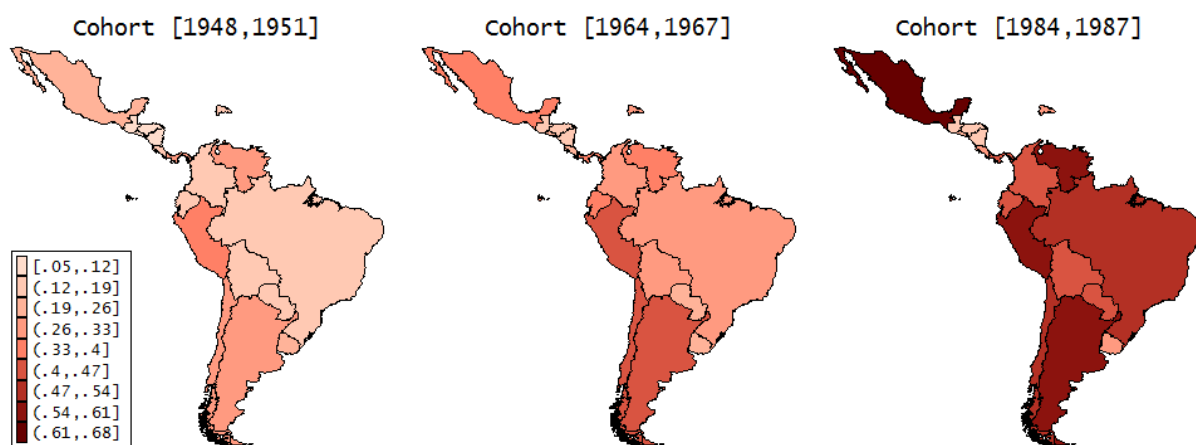


Notes: Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. Bootstrapped confidence interval. *Source:* National Household Surveys 1994-2015, own estimates.

Figure 7: Educational inequality in Latin America: bottom-upward Mobility (*BUM*) and upper class persistence (*UCP*).

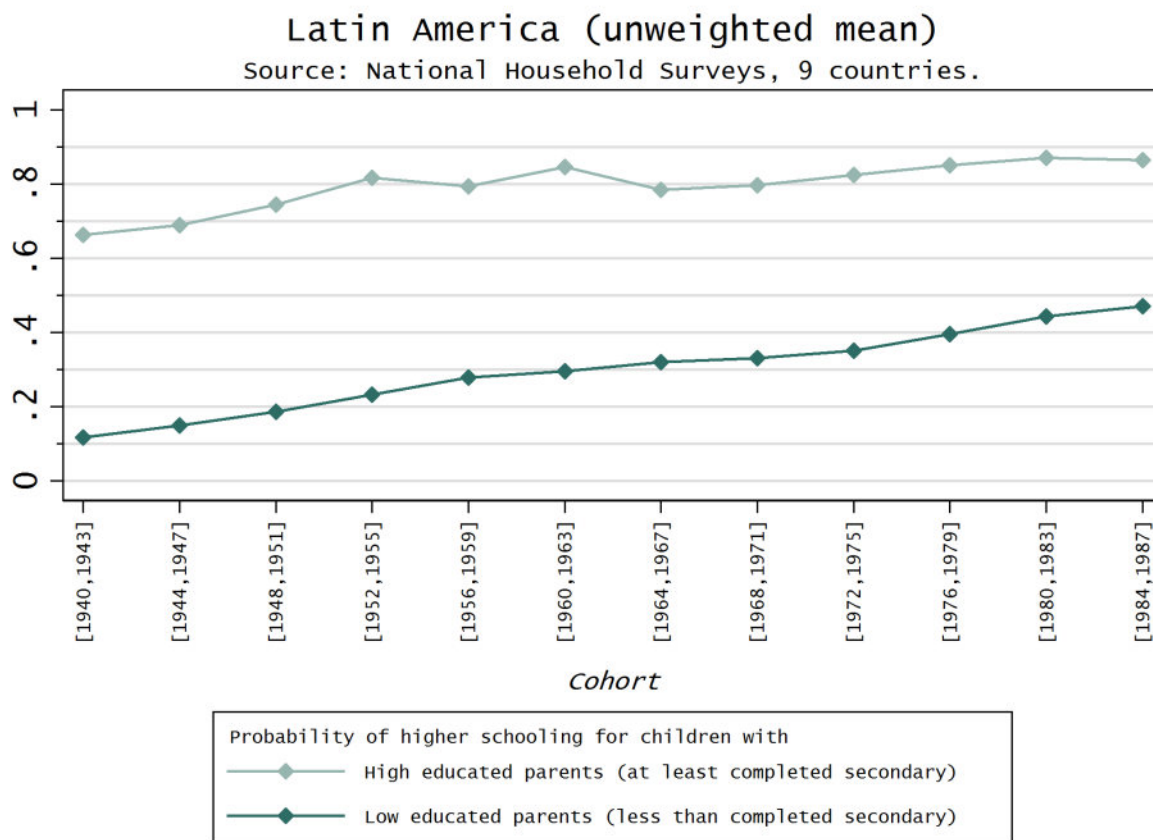


Bottom upward Mobility: Geography and Trends for Latin America



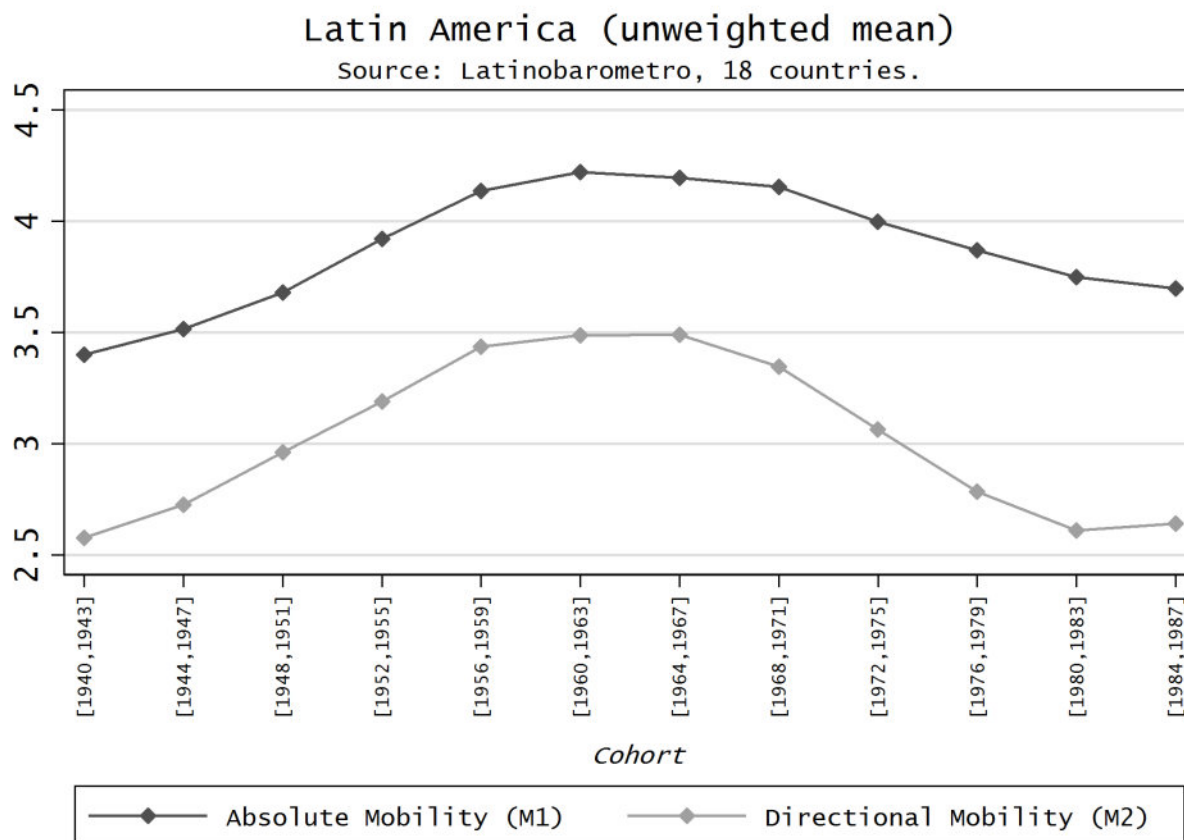
Notes: Estimated probability of higher education (at least completed secondary) of children with different parental educational background. Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. Bootstrapped confidence interval. *Source:* Latinobarometro 1998-2015, own estimates.

Figure 8: Educational inequality in Latin America: bottom-upward Mobility (*BUM*) and upper class persistence (*UCP*).

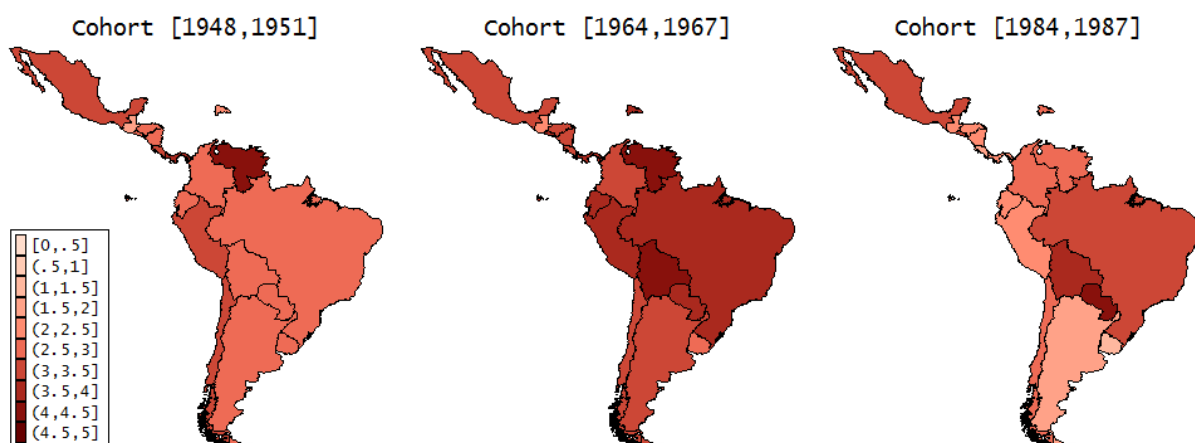


Notes: Estimated probability of higher education (at least completed secondary) of children with different parental educational background. Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. Bootstrapped confidence interval. *Source:* National Household Surveys 1994-2015, own estimates.

Figure 9: Educational mobility in Latin America: absolute ($M1$) and directional ($M2$) mobility in years of education.

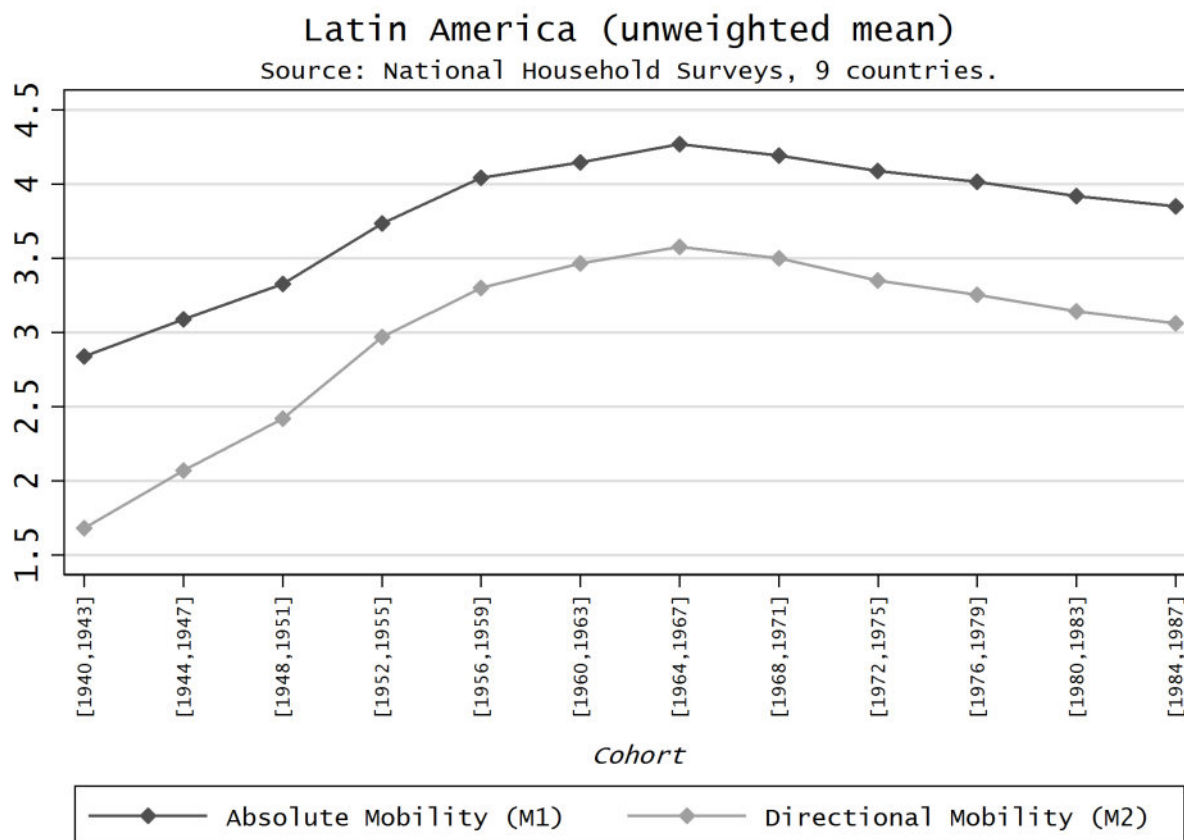


Directional Mobility: Geography and Trends for Latin America



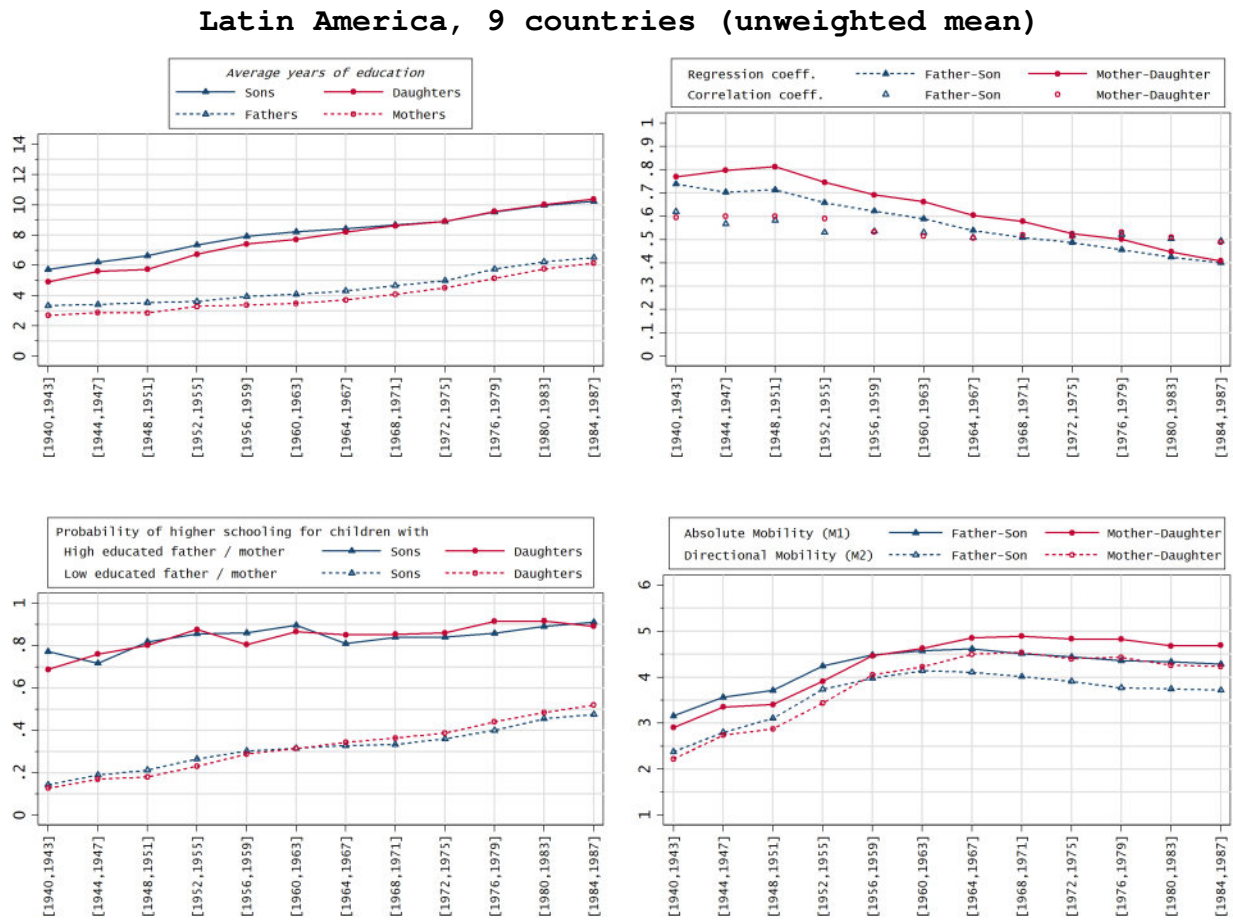
Notes: Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. Source: Latinobarometro 1998-2015, own estimates.

Figure 10: Educational mobility in Latin America: absolute ($M1$) and directional ($M2$) mobility in years of education.



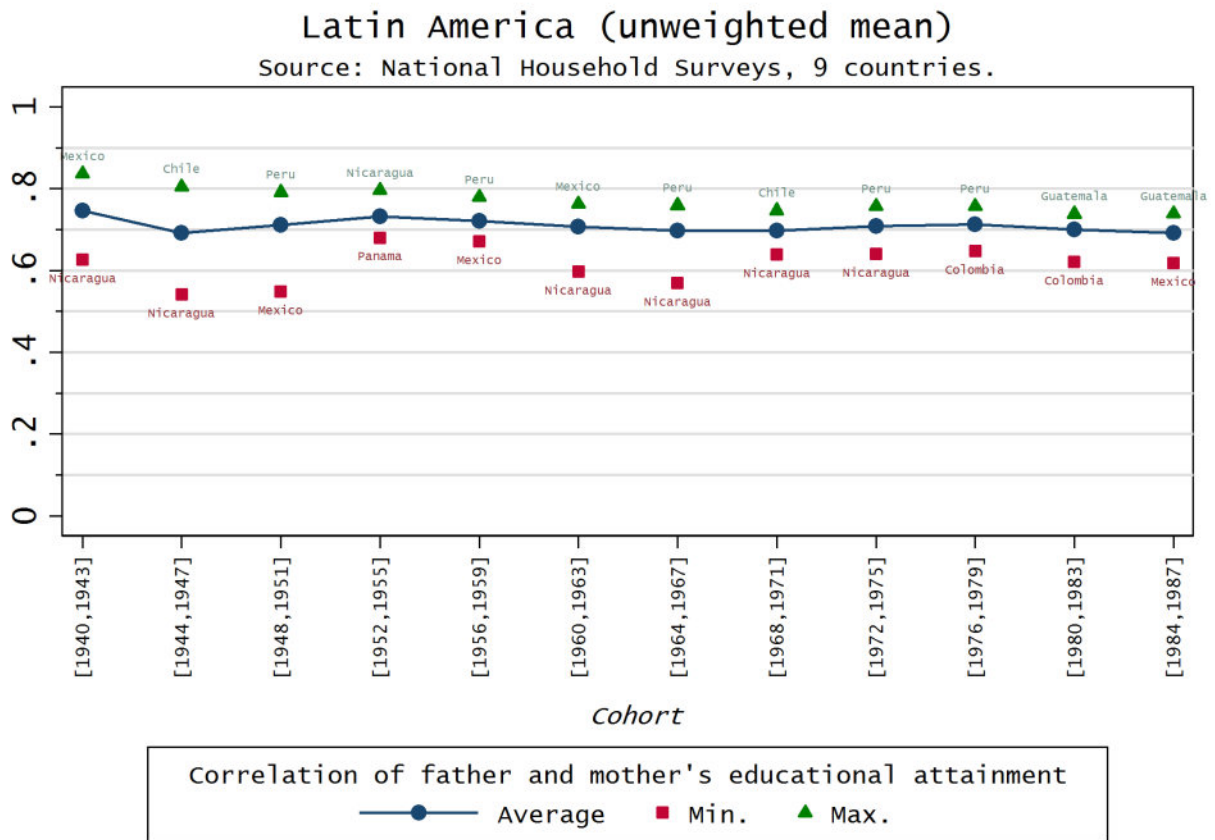
Notes: Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. *Source:* National Household Surveys 1994-2015, own estimates.

Figure 11: Average educational attainment by gender and intergenerational mobility for father-son and mother-daughter pairs.



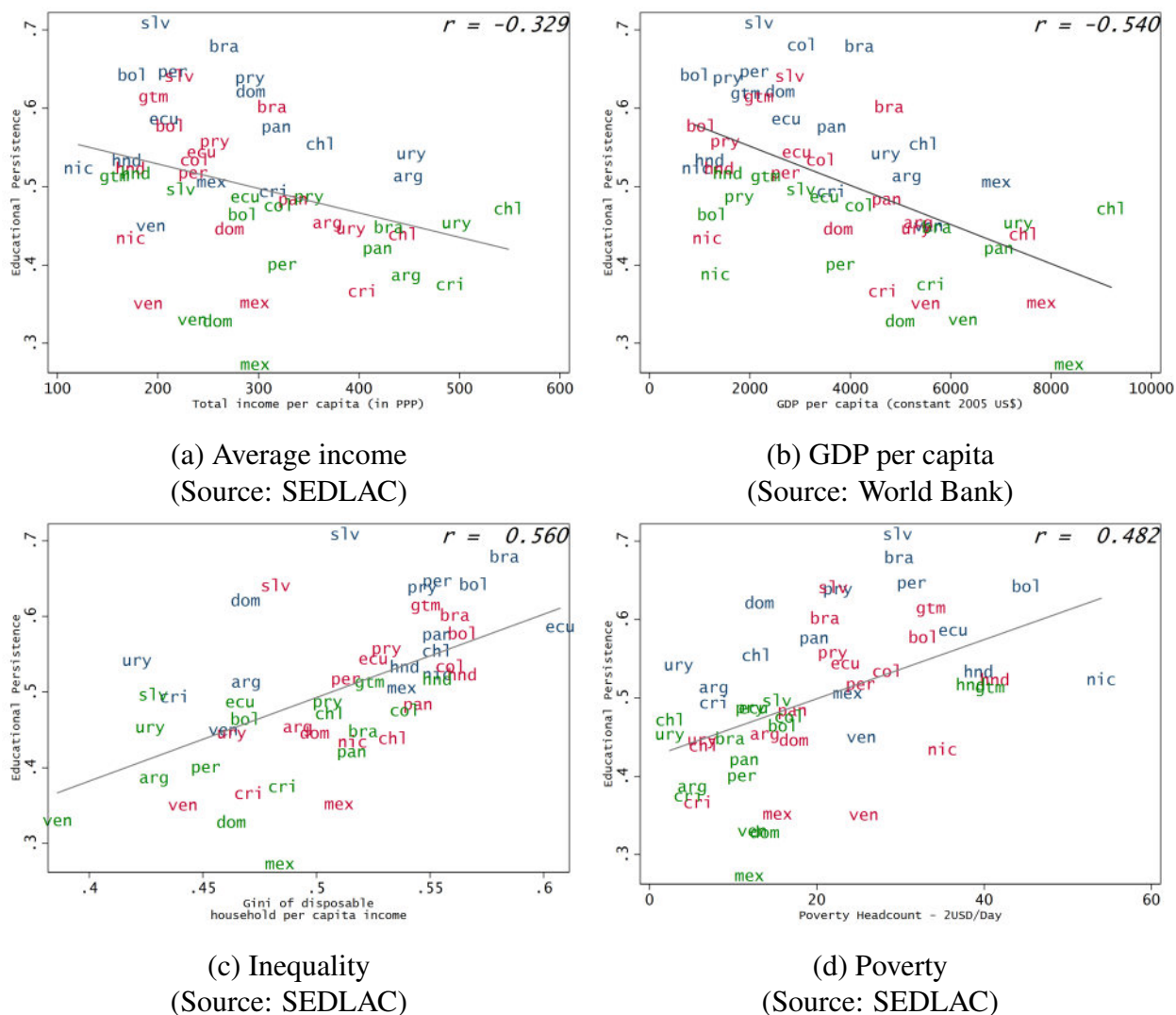
Source: National Household Surveys 1994-2015, own estimates.

Figure 12: Assortative mating – spouse correlation in educational attainments (parental generation).



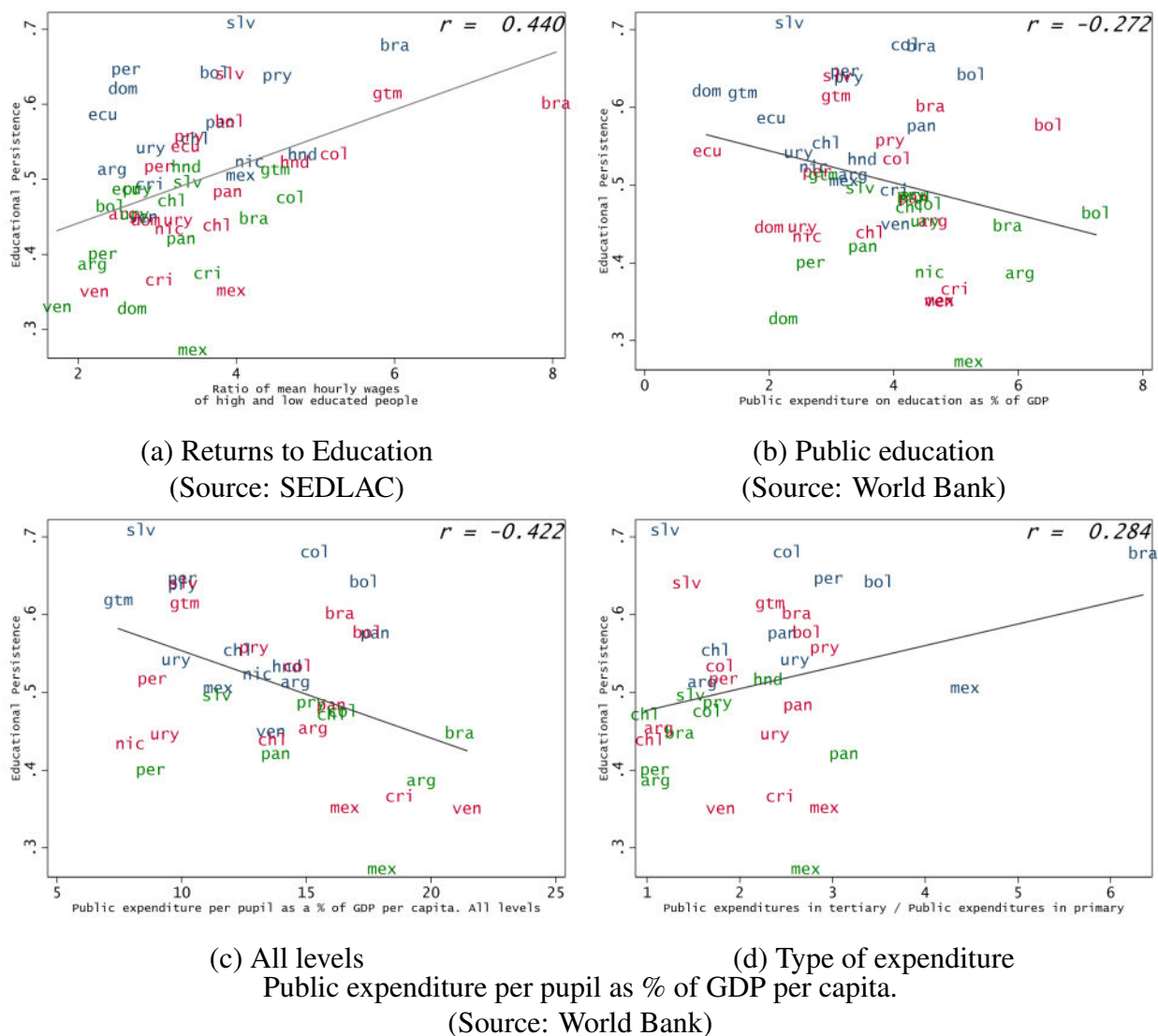
Notes: Points show the unweighted mean over all countries of the estimates for each cohort. Samples for each cohort and country restricted to individuals older than 22. *Source:* National Household Surveys 1994-2015, own estimates.

Figure 13: Educational persistence and economic performance.



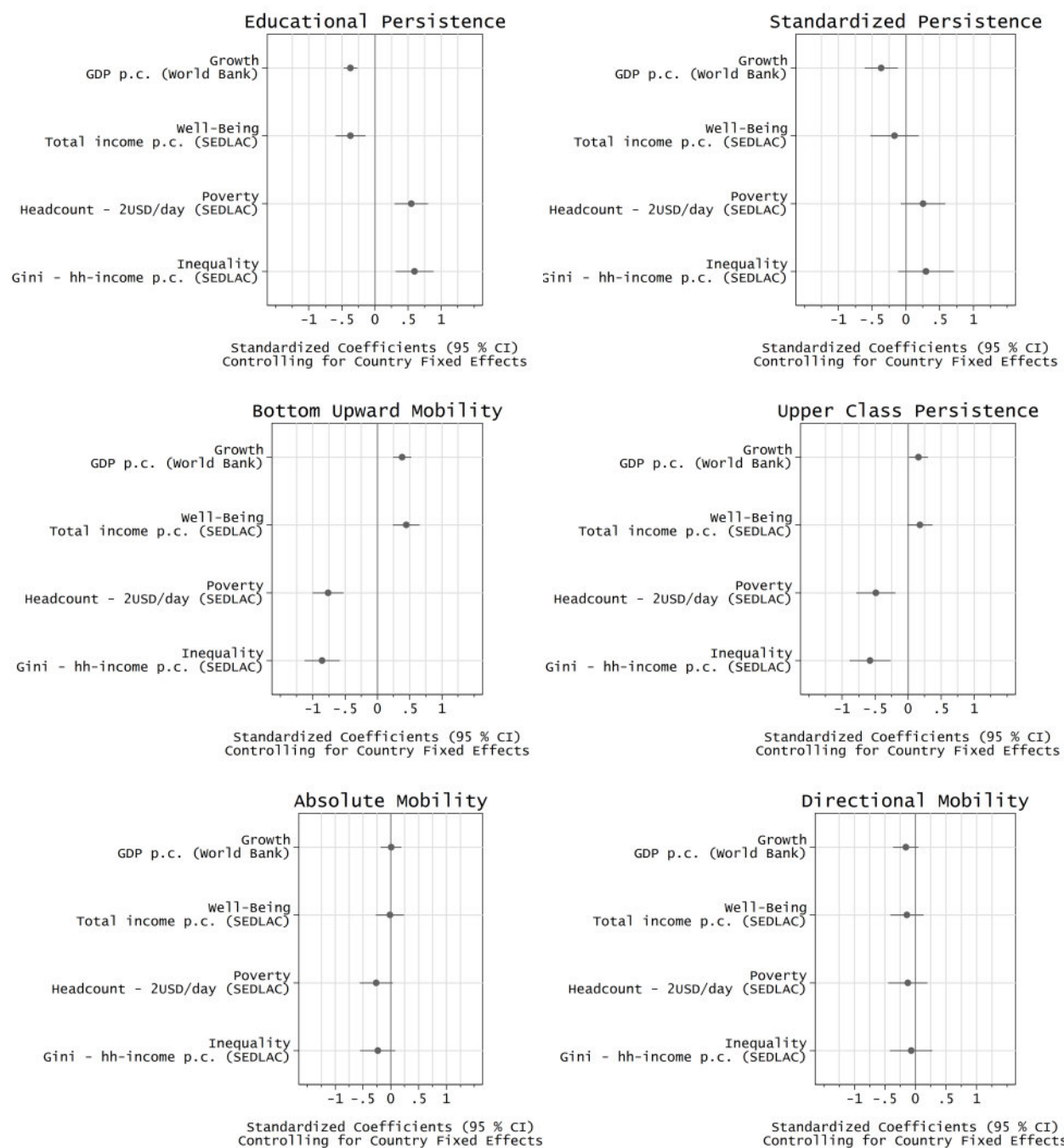
Notes: Intergenerational mobility of the cohorts '40-'54, '55-'69, '70-'84 is associated with the corresponding macroeconomic or institutional characteristic in the years 1990-99, 2000-09, 2010-14. *Sources:* Latinobarometro 1998-2015, own estimates of educational persistence; SEDLAC; World Bank Data.

Figure 14: Educational persistence and institutional characteristics of the education system.



Notes: Intergenerational mobility of the cohorts '40-'54, '55-'69, '70-'84 is associated with the corresponding macroeconomic or institutional characteristic in the years 1990-99, 2000-09, 2010-14. *Sources:* Latinobarometro 1998-2015, own estimates of educational persistence; SEDLAC; World Bank Data.

Figure 15: Intergenerational mobility and economic performance. Regression analysis controlling for country fixed effects.



Notes: Intergenerational mobility of the cohorts '40-'54 , '55-'69, '70-'84 is associated with the corresponding macroeconomic or institutional characteristic in the years 1990-99, 2000-09, 2010-14. Points show the standardized linear regression coefficients and their corresponding confidence interval (95 %) of separate regressions for each of the indicated variables. All regressions control for country fixed effects. In cases where the confidence interval line crosses the zero-line, point estimates are not statistically significant at the 0.05 level. *Sources:* Latinobarometro 1998-2015, own estimates of intergenerational mobility; SEDLAC; World Bank Data.

SUPPLEMENTAL MATERIAL

For Online Publication

APPENDIX A Summary of Data Sources: National Household Surveys	I
APPENDIX B Description of the Database	VII
APPENDIX C Country-Wise Estimates	VIII
APPENDIX D Robustness: Non-linear Correlations	XVIII

A Summary of Data Sources

A.1 Household Surveys

Our main source of information for all 18 Latin American countries in our analysis is the Latino-barometro survey. Using the survey waves 1998 to 2015 our overall sample comprises 211,401 observations. We complement this with National Household Surveys that include information on parental educational achievements collected through retrospective questions. This second data set comprises 1,078,445 observation in total that derive from different data sources.

Data from Brazil comes from the *Pesquisa Nacional por Amostra de Domicílios* (PNAD), which is carried out by the *Instituto Brasileiro de Geografia y Estadísticas* (IBGE) on a yearly basis. This survey included mobility modules in 1982, 1988, 1996 and 2014. Since the coding of the educational variable is not comparable between 2014 and the other three survey waves, we opt to use only the most recent one in our analysis. The survey is nationally and regionally representative, rural and urban, except for the rural areas of the Northern Region, which roughly corresponds to the Amazon rainforest and accounted for 2.3% of Brazil's population in the 2000 Census.

For Chile, we use the *Encuesta de Caracterización Socioeconómica Nacional* (CASEN), which is a nationally and regionally representative household survey carried out by the Ministry of Social Development (in collaboration with the National Institute of Statistics, INE) through the Department of Economics at the *Universidad de Chile*, which is responsible for the data collection, digitalization and consistency checking of the database.¹⁶ The survey has been regularly implemented every two years since 1985 during November and in some cases, up to mid-December. We use surveys for 2006 to 2015, since previous surveys don't provide information about parents.

The same is true for Peru, using the *Encuesta Nacional de Hogares* (ENAHO), which is carried out in four waves since 1997, and continues until today. The fourth wave of the survey is nationally representative, and it is officially used to estimate poverty rates. After year 2000 the survey was enlarged and a new sample frame was used, including questions about parents. We use surveys

¹⁶Before 2011 the survey was carried out by the Ministry of Planning (MIDEPLAN).

for 2001 to 2015. However, from 2002 on the survey asked only the household head about the education of parents. Since most household heads are male the sex composition of our sample is therefore unbalanced.

For the other countries we use different versions of Living Standards Measurement Surveys, originally developed and promoted by the World Bank, which are all nationally representative. Data from Ecuador comes from the *Encuesta de Condiciones de Vida* (ECV) for years 1994, 1995, 1998 and 2006. In the case of Colombia we use the *Encuesta Nacional de Condiciones de Vida* (ECV), which was carried out by the *Departamento Administrativo Nacional de Estadística* (DANE). We use surveys for six years between 2003 and 2013. Although Guatemala is a country with relatively few household surveys, the *Encuesta Nacional sobre Condiciones de Vida* (ENCOVI) have information about individuals' parents (2000, 2006 and 2011). Panama carried out Living Standards Measurement Surveys in 1997, 2003 and 2008, which are called *Encuesta Nacional sobre Condiciones de Vida* (ENV).

The source of information for our estimations of Mexico's statistics is the Mexican Family Life Survey (MxFLS), which is a longitudinal and multi-thematic survey, representative of the Mexican population at the national, urban, rural and regional level. The MxFLS has been developed and managed by researchers from the Iberoamerican University (UIA, per its name in Spanish) and the Center for Economic Research and Teaching (CIDE, per its name in Spanish) in collaboration with researchers from Duke University. Currently, the MxFLS contains information for a 10-year period, collected in three rounds: 2002, 2005-2006 and 2009-2012.

Finally, for Nicaragua the only useful source for our analysis we could find besides Latinobarometro is the 1998 wave of the *Encuesta Nacional de Hogares sobre Medición de Nivel de Vida* (EMNV).

Table A1: *Household surveys used to construct the intergenerational mobility estimates*

Country	Name of survey	Acronym	Coverage	Survey waves
Argentina	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Bolivia	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Brazil	<i>Pesquisa Nacional por Amostra de Domicilios</i>	PNAD	National	2014
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Chile	<i>Encuesta de Caracterización Socioeconómica Nacional</i>	CASEN	National	2006, 2009, 2011, 2013, 2015
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Colombia	<i>Encuesta Nacional de Condiciones de Vida</i>	ECV	National	2003, 2008, 2010, 2011, 2012, 2013
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Costa Rica	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Dominican Rep.	<i>Latinobarometro</i>		National	2004-2011, 2013, 2015

Table A1: *Household surveys used to construct the intergenerational mobility estimates*

Country	Name of survey	Acronym	Coverage	Survey waves
Ecuador	<i>Encuesta de Condiciones de Vida</i>	ECV	National	1994, 1995, 1998, 2006
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
El Salvador	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Guatemala	<i>Encuesta Nacional sobre Condiciones de Vida</i>	ENCOVI	National	2000, 2006, 2011
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Honduras	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Mexico	<i>Encuesta Nacional sobre Niveles de Vida de los Hogares</i>	MXFLS	National	2002, 2005-2006, 2009-2012
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Nicaragua	<i>Encuesta Nacional de Hogares sobre Medición de Nivel de Vida</i>	EMNV	National	1998
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Panama	<i>Encuesta de Niveles de Vida</i>	ENV	National	1997, 2003, 2008
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015

Table A1: *Household surveys used to construct the intergenerational mobility estimates*

Country	Name of survey	Acronym	Coverage	Survey waves
Paraguay	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Peru	<i>Encuesta Nacional de Hogares</i>	ENAHO	National	2001-2015
	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Uruguay	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015
Venezuela	<i>Latinobarometro</i>		National	1998, 2000-2011, 2013, 2015

A.2 Codification of Educational Attainment

0	Illiterate
1	Incomplete primary
2	'
3	'
4	'
5	'
6	Complete primary
7	'
8	Incomplete secondary
9	'
10	'
11	Complete secondary
12	'
13	Incomplete university or technical training
14	Complete technical training
15	Complete university

B Description of the Database

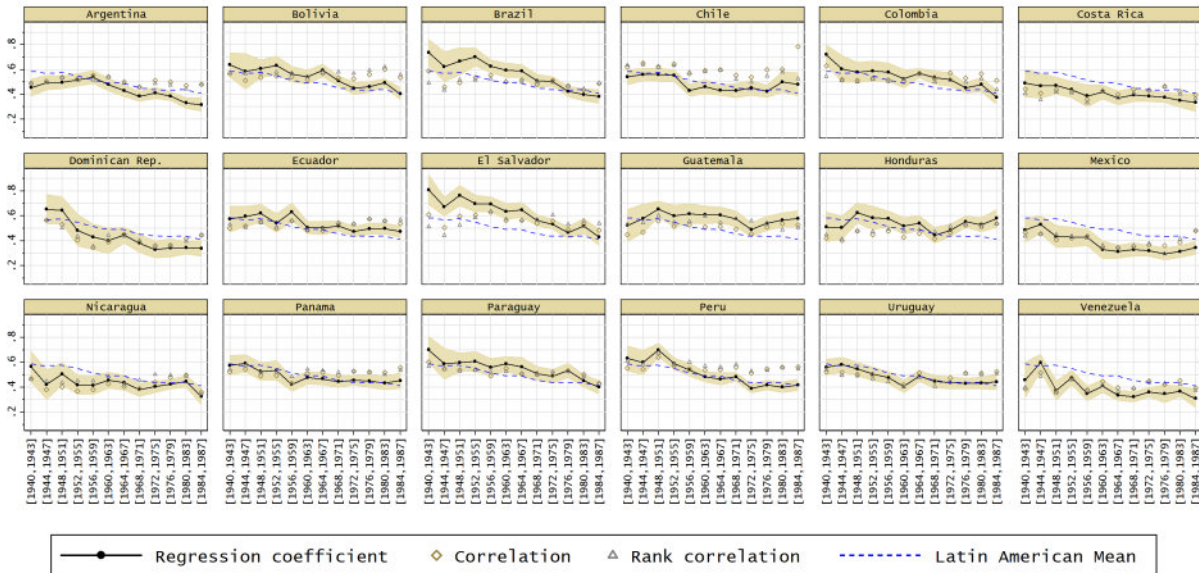
We provide databases containing all mobility indicators described in this project. The variables contained in each database are described in Table B1. The data is divided in four different sets of different periodization of the birth cohorts, separated at intervals of one to four years, respectively. In addition to the main statistics and the identification variables of each country, survey and cohort, we also include complementary variables that may be useful, such as mean and variance of the years of education of individuals and their parents, the average age of individuals and the share of males in the sample for each cohort. Finally, we add a variable that contains the number of observations used for the estimation of mobility statistics to make it possible to evaluate the quality of the estimates.

Table B1: Summary table of the database.

Variable	Label	Definition	Mean	Std. Dev.	Min	Max
country	Country name	Name of country				
idenpa	Country code	World Bank country code				
cohort	Cohort	Cohort indicator				
survey	Survey name	Name of the survey				
N	Number of observations	Number of observations used to estimate indicators	3421.34	7508.60	19	45046
b	Intergenerational persistence parameter	Conditional correlation between years of education of children and parents (beta)	0.49	0.14	0.02	0.91
bstd	Intergenerational correlation (b standardized)	Parameter b weighted by the ratio of standard deviations of years of schooling of children and parents	0.50	0.09	0.06	0.79
corr_spearman	Spearman's correlation	Spearman's rank correlation coefficient (rho)	0.49	0.08	-0.05	0.67
blog	Intergenerational elasticity	Parameter b estimated using the logarithm of the outcome of interest (years of schooling)	0.34	0.12	0.00	0.70
prob_high	Prob(high education) High parental education	Predicted probability of upper class persistence (UCP)	0.75	0.13	0.16	0.97
prob_low	Prob(high education) Low parental education	Predicted probability of bottom upward mobility (BUM)	0.27	0.15	0.03	0.81
M1	Absolute mobility	Absolute mobility (M1)	3.79	0.68	1.60	5.23
M2	Directional mobility	Directional mobility (M2)	2.90	0.83	0.50	4.78
educ	Years of schooling	Average of own years of schooling	8.14	2.19	2.22	14.26
educ_parents	Parental Years of schooling	Average of parents' years of schooling (the highest level of educational attainment among the two)	5.25	2.12	1.39	12.58
var	Variance of years of schooling	Variance of own years of schooling	16.66	5.43	0.84	33.08
var_parents	Variance of parental years of schooling	Variance of parents' years of schooling	17.56	4.31	6.79	32.96
age	Age	Average age of individuals in sample	40.93	13.59	23.00	72.54
male	Share of males	Share of males in sample	0.49	0.06	0.33	0.81

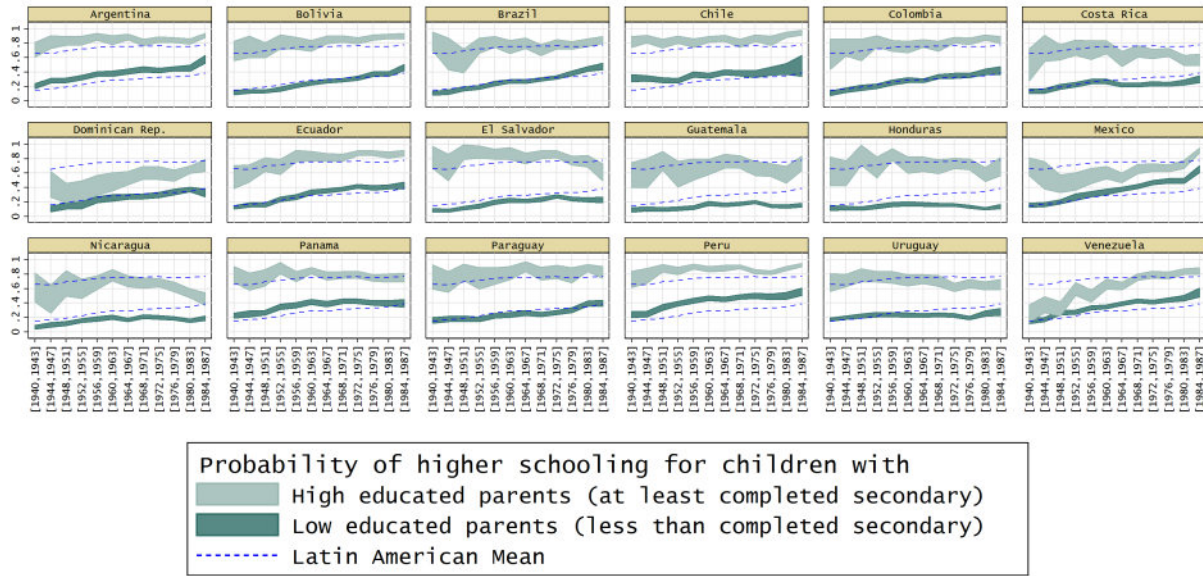
C Country-Wise Estimates

Figure C1: Educational persistence in Latin America: Regression and correlation coefficients by country. *Source:* Latinobarometro 1998-2015, own estimates.



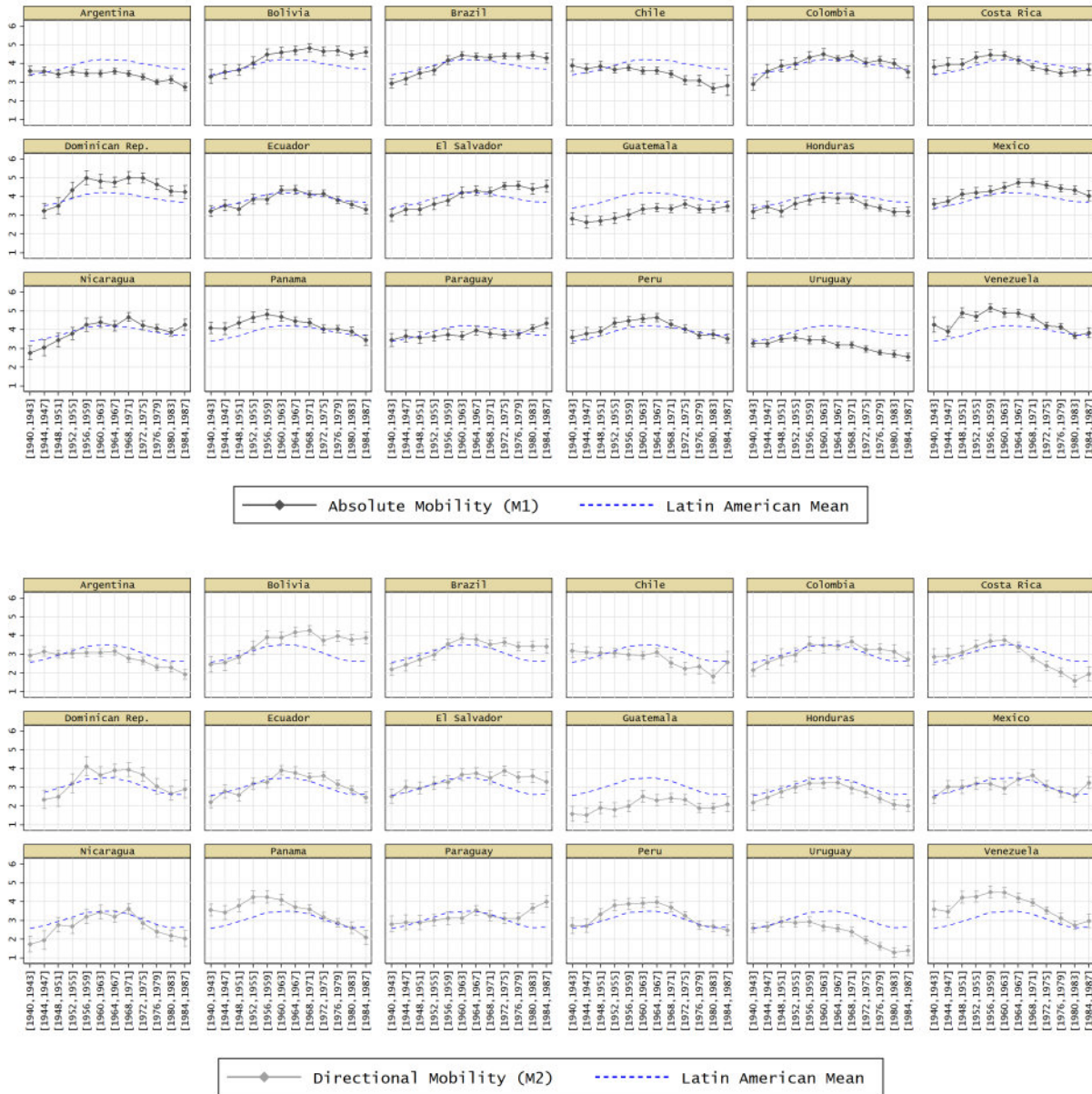
Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C2: Educational inequality in Latin America: bottom-upward Mobility (*BUM*) and upper class persistence (*UCP*). *Source*: Latinobarometro 1998-2015, own estimates.



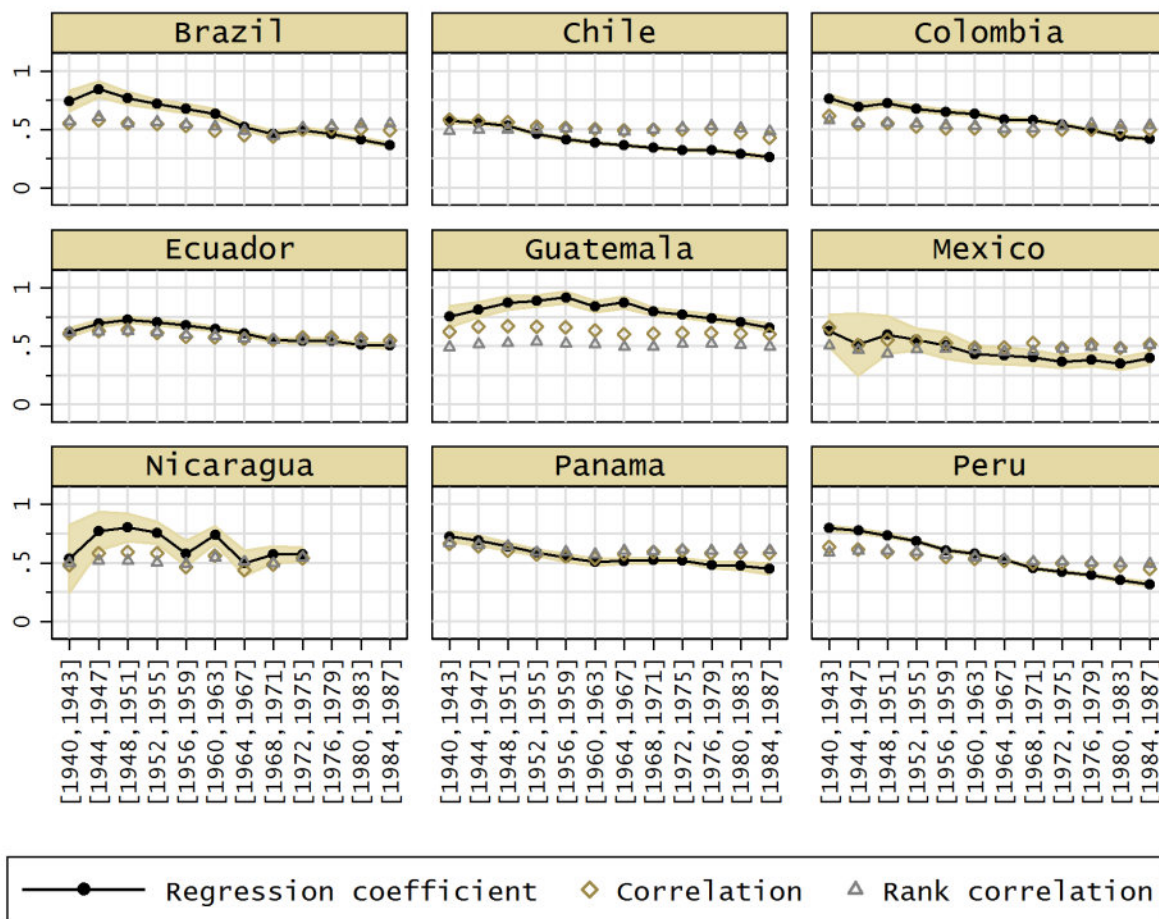
Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C3: Educational mobility in Latin America: absolute ($M1$) and directional ($M2$) mobility in years of education. *Source:* Latinobarometro 1998-2015, own estimates.



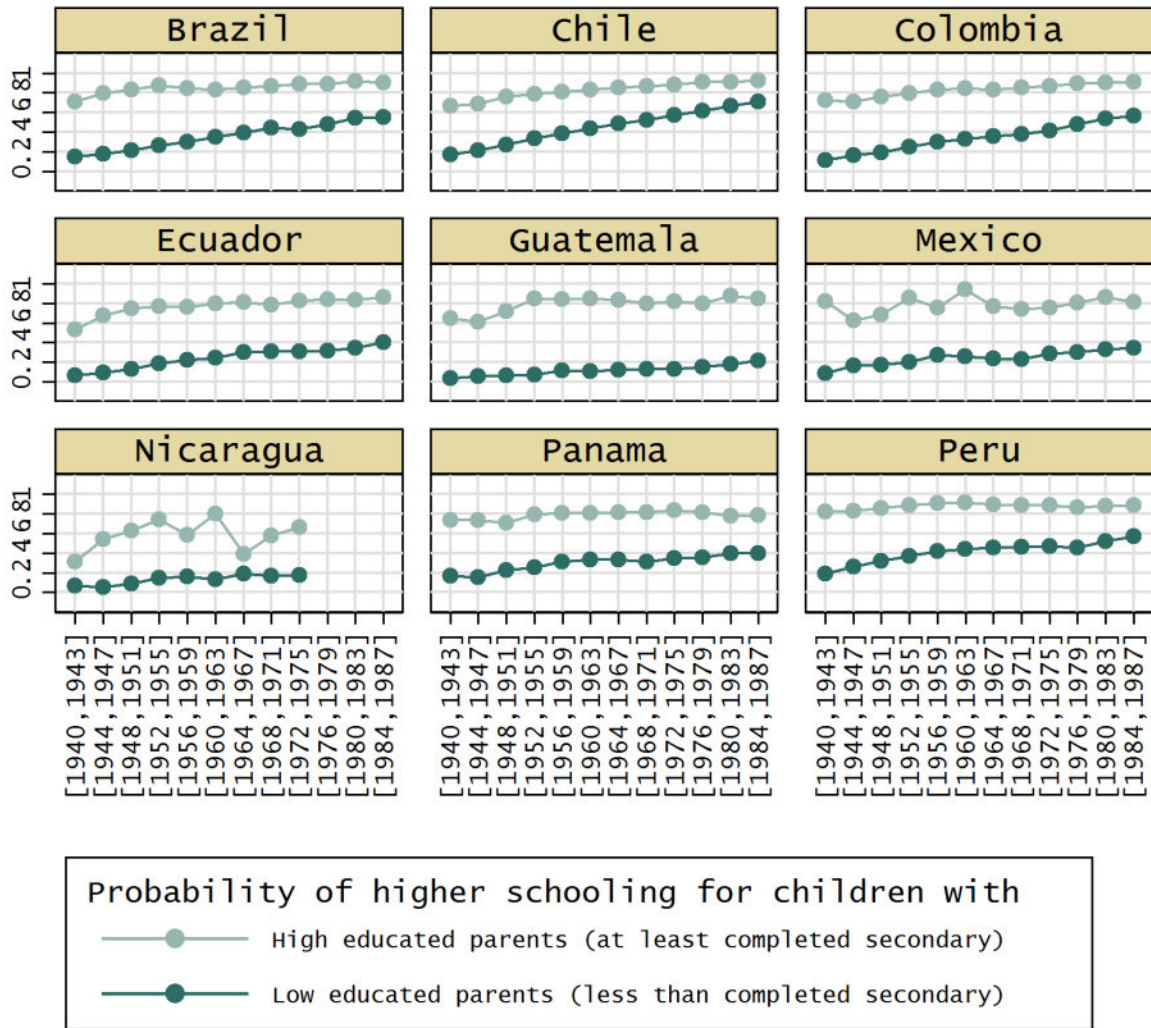
Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C4: Educational persistence in Latin America: Regression and correlation coefficients by country. *Source:* National Household Surveys 1994-2015, own estimates.



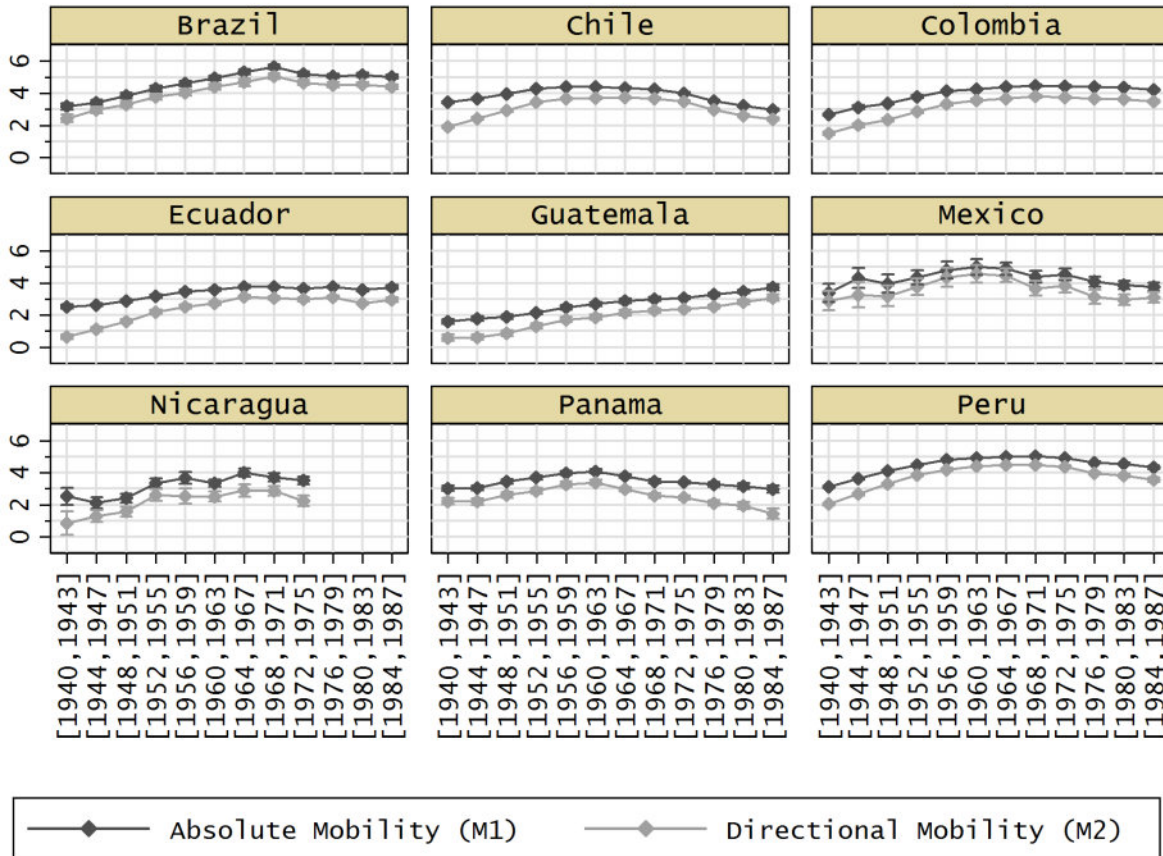
Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C5: Educational inequality in Latin America: bottom-upward Mobility (*BUM*) and upper class persistence (*UCP*). *Source*: National Household Surveys 1994-2015, own estimates.



Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C6: Educational mobility in Latin America: absolute ($M1$) and directional ($M2$) mobility in years of education. *Source:* National Household Surveys 1994-2015, own estimates.



Notes: Samples for each cohort and country restricted to individuals older than 22. Only point estimates displayed relying on at least 200 observations. Bootstrapped confidence interval.

Figure C7: Educational persistence in Latin America for father-son and mother-daughter pairs.
 Source: National Household Surveys 1994-2015, own estimates.

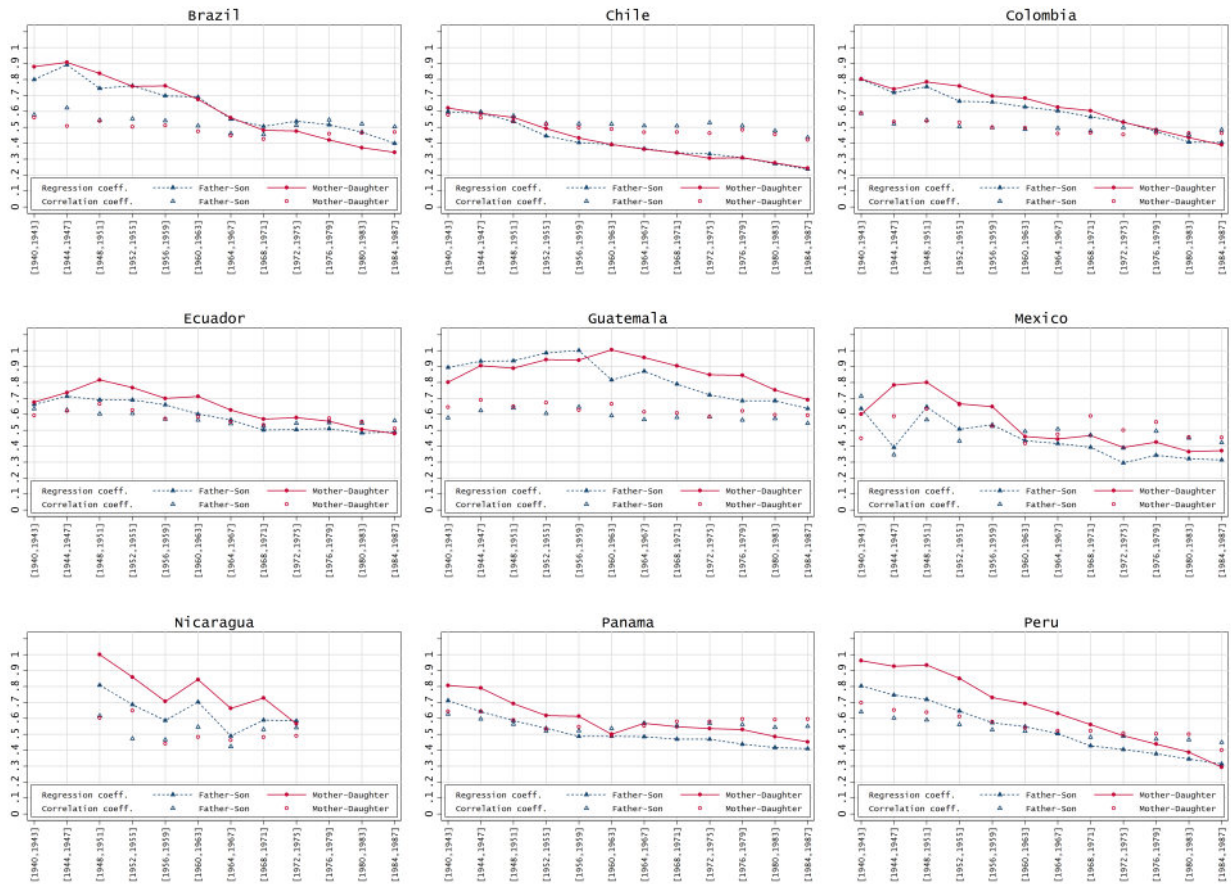


Figure C8: Average educational attainment, intergenerational mobility for father-son and mother-daughter pairs, and assortative mating. *Source:* National Household Surveys 1994-2015, own estimates.

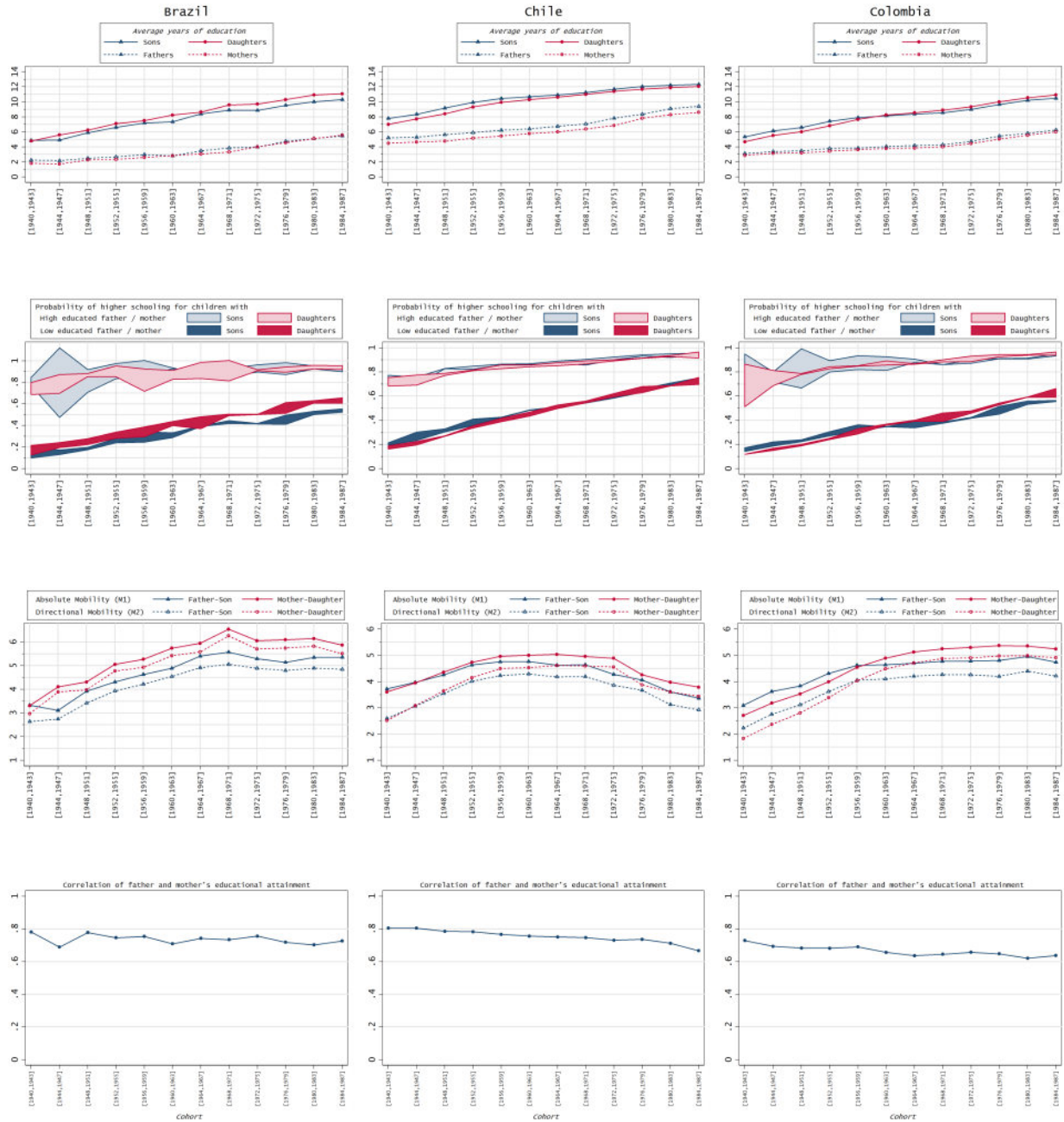


Figure C9: Average educational attainment, intergenerational mobility for father-son and mother-daughter pairs, and assortative mating. *Source:* National Household Surveys 1994-2015, own estimates.

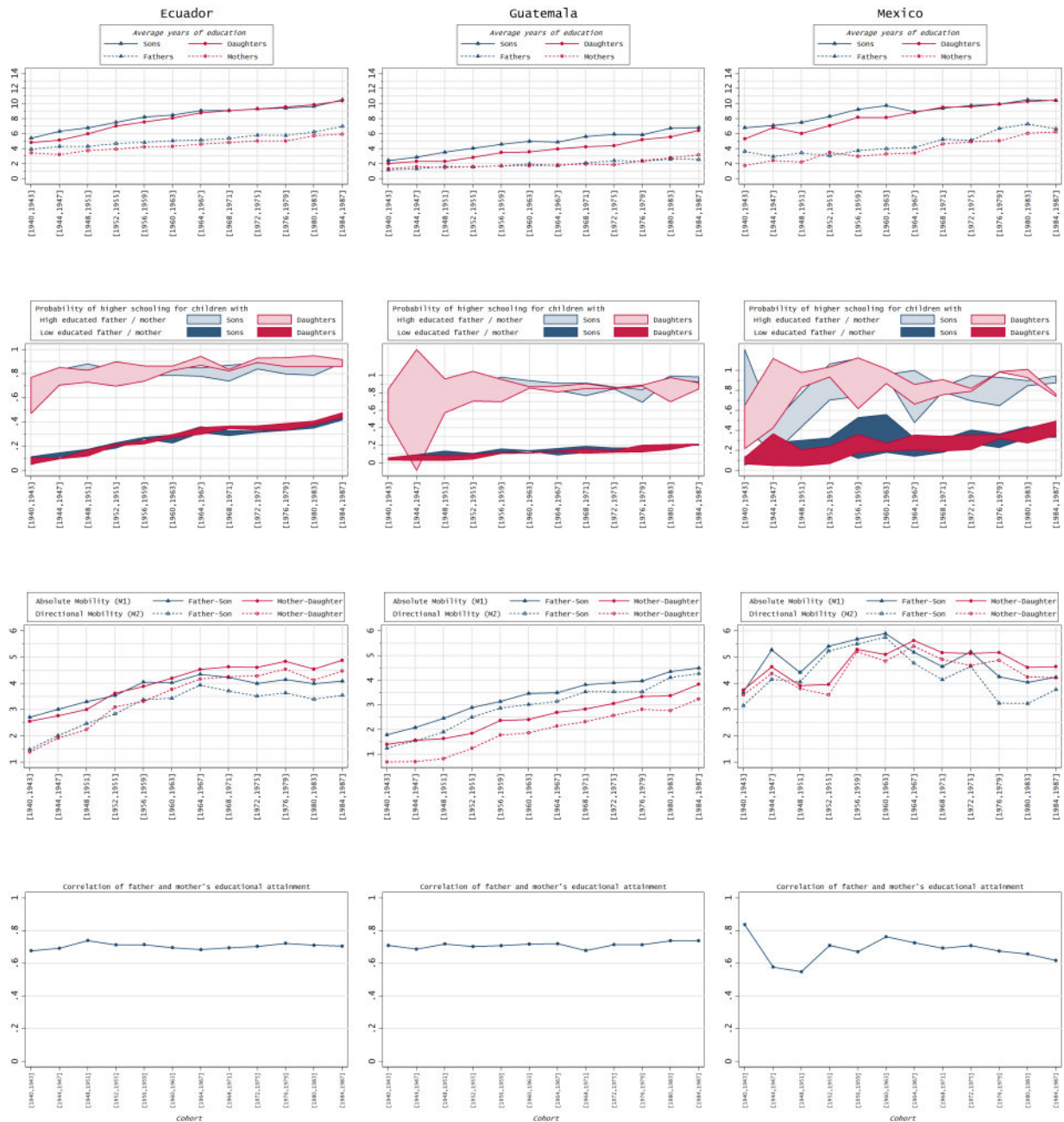
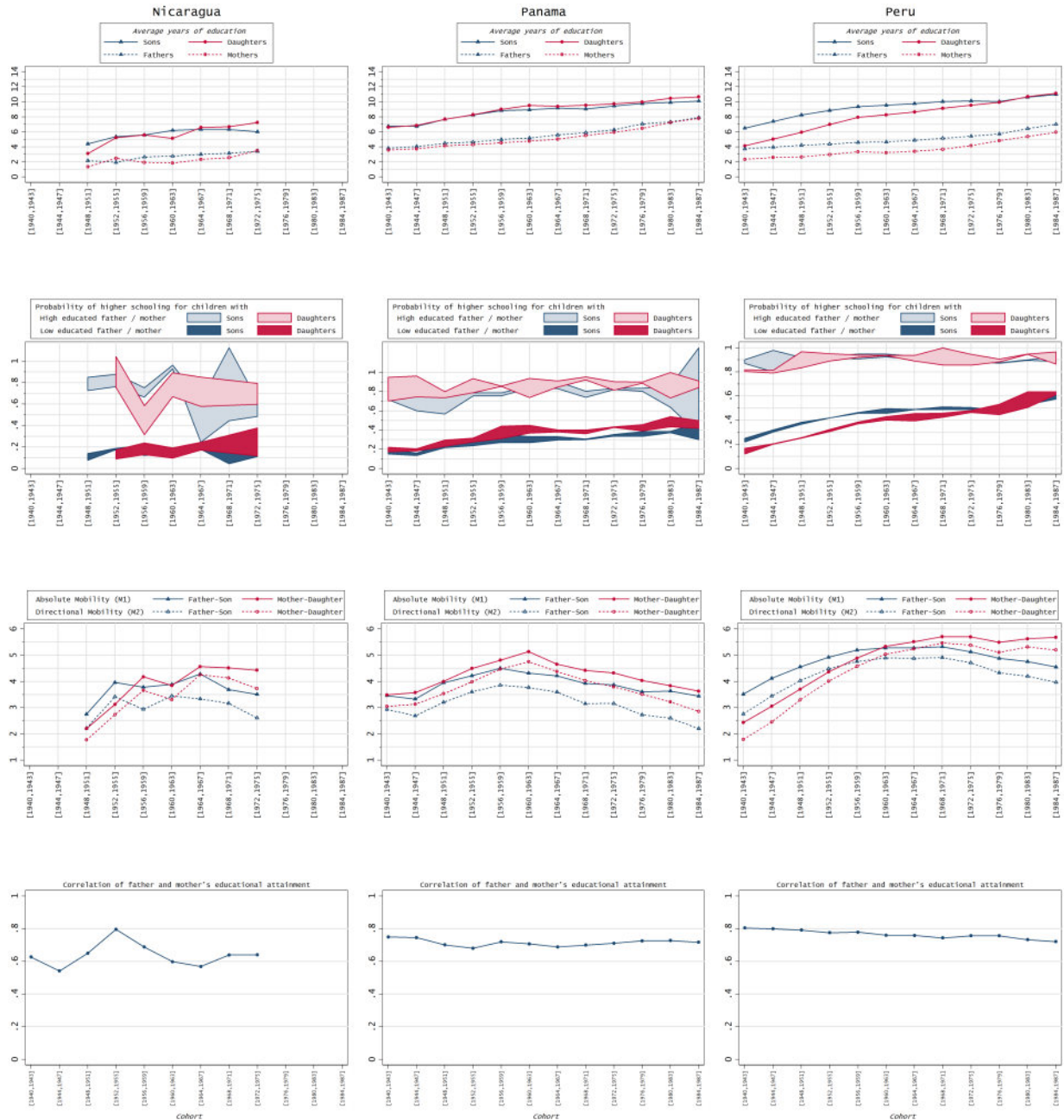


Figure C10: Average educational attainment, intergenerational mobility for father-son and mother-daughter pairs, and assortative mating. *Source:* National Household Surveys 1994-2015, own estimates.



D Non-linear correlation of educational levels

Some of the measures that are usually applied to study intergenerational mobility assume that the relationship between the outcomes of parents and children is linear. However, this assumption has been questioned by recent analyses showing that the slope coefficients might vary for families in different parts of the distribution Bratberg et al. (2017). Especially measuring educational attainment, the assumption of years of education as a cardinal measure and of an underlying monotonic and linear relationship between parents' and children's schooling has been questioned. However, cross country studies show a high correlation between linear and non-linear measures of relative intergenerational mobility (see Blanden, 2013). Figure D1 shows an evaluation of non-linear patterns in the correlation of parents' and children's years of education. Generally, the issue certainly requires particular attention that would go beyond the scope of this work. For the sake of completeness, we here show the robustness of our cross-country estimates applying a measure that takes into account that the correlation of educational levels might be of non-linear nature.

The applied measure is the correlation of error terms in a bivariate ordered probit model. The model estimates the joint probability distribution of two ordered categorical variables, in our case parents' and children's education in levels. This method has been used e.g. by Magee et al. (2000) to estimate assortative mating patterns in educational levels.¹⁷ The outcome variables in our application both have six categories: illiterate, incomplete primary, complete primary, incomplete secondary, complete secondary, incomplete higher education, complete higher education.

Assume that the two latent variables defining the educational level y of parents (p) and children (c) are determined by:

$$y_{pi}^* = X_{pi}'\delta_p + \varepsilon_{pi} \quad (8)$$

$$y_{ci}^* = X_{ci}'\delta_c + \varepsilon_{ci} \quad (9)$$

¹⁷For further examples, see Sajaia (2008).

where i denotes the family. δ_p and δ_c are vectors of parameters for X_p and X_c that include age and sex and satisfy the exogeneity conditions $E[X_{pi}\varepsilon_{pi}] = E[X_{ci}\varepsilon_{ci}] = 0$. ε_p and ε_c are the error terms, distributed as a bivariate standard normal. Denote the cutoffs of the observed categorical variables indicating parents' educational level $j \in (1, 2, 3, 4, 5, 6, 7)$ as c_{pj} , where $c_{pj-1} < c_{pj}$, and let $c_{pj} = -\infty$ for $j = 0$ and $c_{pj} = \infty$ for $j = 7$. The indicator for the child is determined in the same way. Then the probability that the parent and the child have the same educational level m is

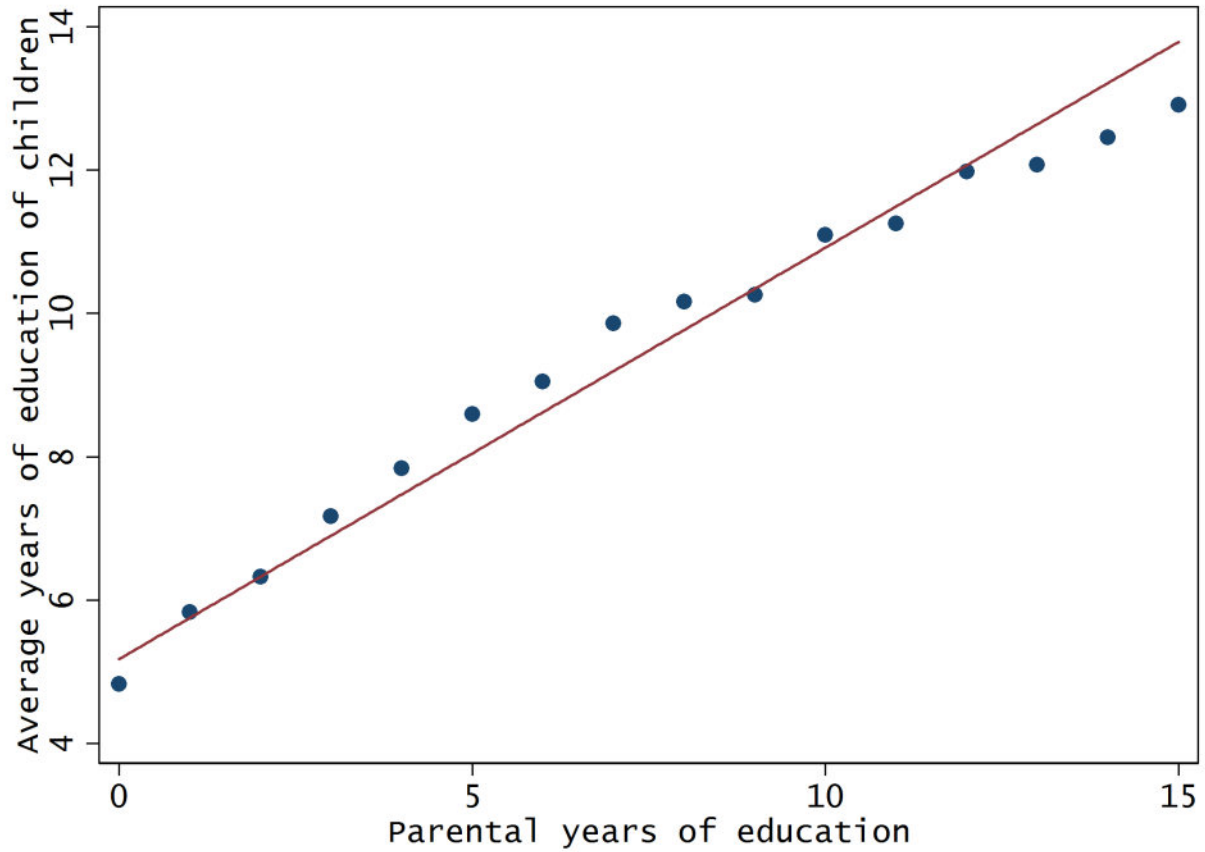
$$Pr(y_{pi} = m, y_{ci} = m) = Pr(c_{pm-1} < y_{pi}^* \leq c_{pm}, c_{cm-1} < y_{ci}^* \leq c_{cm}).$$

The parameter of interest here is the association measure ρ^ε that is the correlation between the two error terms ε_p and ε_c . Figures **D2** and **D3** show ρ^ε estimated separately for each cohort and compare it with the Pearson correlation coefficient ρ measured on the same ordered variables. As is evident, ρ^ε is always higher than ρ in all countries and surveys, but the trends are almost constantly parallel.

References

- Blanden, J. (2013), Cross-country Rankings In Intergenerational Mobility: A Comparison Of Approaches From Economics And Sociology. *Journal of Economic Surveys*, 27: 38–73.
- Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D. D. and Vaage, K. (2017), A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US. *Scandinavian Journal of Economics*, 119: 72–101.
- Magee, L., Burbidge, J., Robb, L. (2000), The Correlation Between Husband's and Wife's Education: Canada, 1971-1996. SEDAP Research Paper No. 24
- Sajaia, Z. (2008), Maximum likelihood estimation of a bivariate ordered probit model: implementation and Monte Carlo simulations. *The Stata Journal*, 4(2): 1-8.

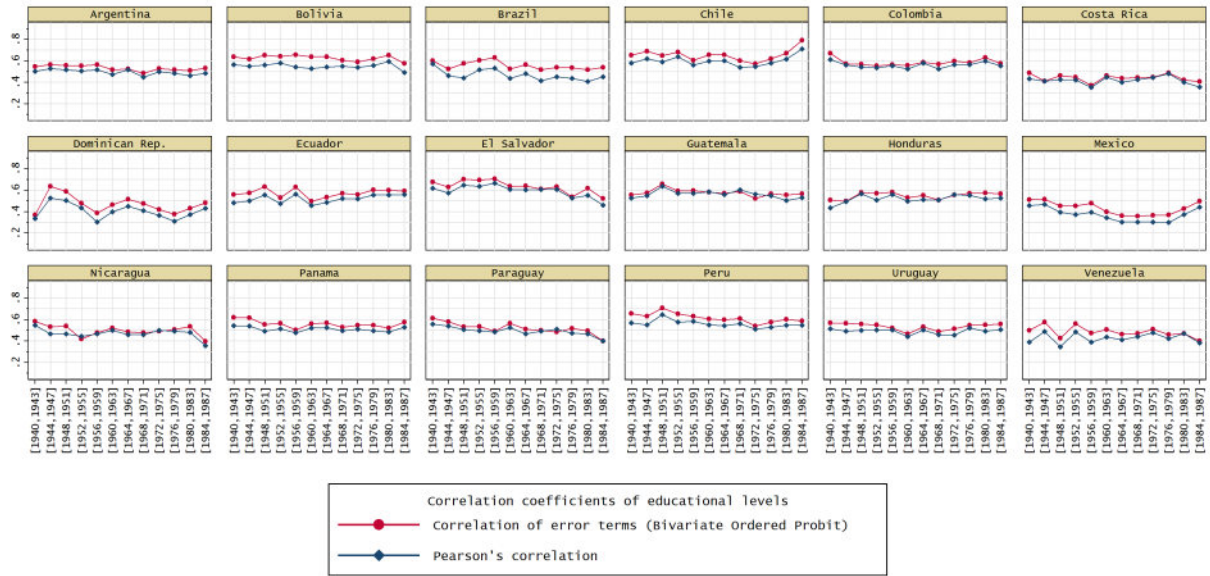
Figure D1: Children's average years of education for each level of parental education.



Notes: Samples for each cohort and country restricted to individuals older than 22. *Source:* Lati-nobarometro 1998-2015, own estimates.

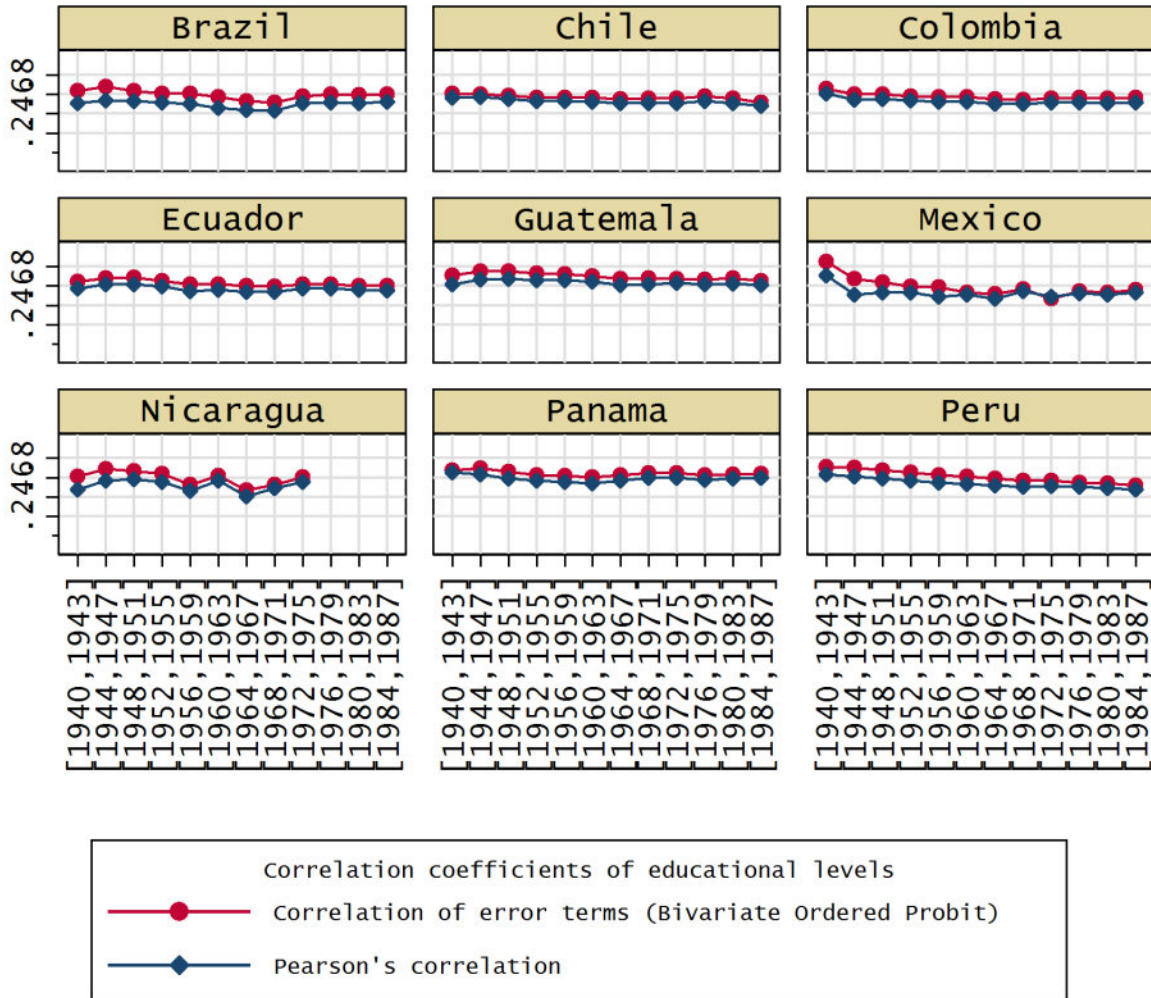
D. NON-LINEAR CORRELATION OF EDUCATIONAL LEVELS

Figure D2: Educational persistence in Latin America: Correlation coefficients by country. Latino-barometro.



Notes: Samples for each cohort and country restricted to individuals older than 22. *Source:* Latino-barometro 1998-2015, own estimates.

Figure D3: Educational persistence in Latin America: Correlation coefficients by country. National Household Surveys.



Notes: Samples for each cohort and country restricted to individuals older than 22. Source: National Household Surveys 1994-2015, own estimates.