FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN

IECON INSTITUTO DE ECONOMÍA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Intergenerational mobility along the income distribution: estimates using administrative data for a developing country

Martin Leites, Xavier Ramos, Cecilia Rodriguez, Joan Vilá

INSTITUTO DE ECONOMÍA
Serie Documentos de Trabajo

Mayo, 2022
DT 5/2022

ISSN: 1510-9305 (en papel)
ISSN: 1688-5090 (en línea)

investigación fue financiada por la CSIC, convocatoria I+D 2018 de la Universidad de la República. Xavier Ramos y Martín Leites agradecen además el apoyo del proyecto PID2019-104619RB-C43 (Ministerio de Ciencia e Innovación de España). Asimismo, los autores reconocen los aportes del equipo de la División Planificación, Estudios y Coordinación de la Dirección General Impositiva y del equipo de Planificación e Información del Banco de Previsión Social, que fue un apoyo fundamental en la construcción de la base de datos. También agradecen al equipo de investigadores del grupo de desigualdad y pobreza del Instituto de Economía de la Universidad de la República, quienes han realizado avances sustantivos en el trabajo con registros administrativos; sin esta acumulación previa hubiera sido muy dificultoso poder realizar esta investigación. Los errores y omisiones son responsabilidad de los autores

Forma de citación sugerida para este documento: Leites, M. Ramos, X., Rodriguez, C. y Vilá, J. (2022) "Intergenerational mobility along the income distribution: estimates using administrative data for a developing country". Serie Documentos de Trabajo, DT 05/2022. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.

La movilidad intergeneracional a lo largo de la distribución: estimaciones en base a registros administrativos para un país en desarrollo

Martin Leites, Xavier Ramos, Cecilia Rodriguez, Joan Vilá

Resumen

Esta investigación aporta nuevas medidas sobre la movilidad intergeneracional de ingresos en Uruguay, siendo hasta el momento la primera que brinda estimaciones para países en desarrollo basadas en registros administrativos. Se estima el grado de asociación intergeneracional entre rankings (IRA por su sigla en inglés) y se explora la presencia de no linealidades en esta relación. Se aplican distintas estrategias para abordar la presencia de un segmento informal en el mercado laboral, desafío adicional para obtener medidas precisas sobre la movilidad intergeneracional en los países en desarrollo. Los resultados principales son tres: primero, el nivel de persistencia intergeneracional es más alto cuando se consideran individuos con un vínculo menos estable con el sector formal. Segundo, encontramos la presencia de no linealidades en el grado de persistencia intergeneracional, siendo ésta sustantivamente más alta en los sectores de altos ingresos. Finalmente, las estimaciones sugieren diferencias significativas por género, siendo la transmisión intergeneracional entre madres e hijas la de mayor persistencia.

Palabras clave: Movilidad intergeneracional de ingresos, desigualdad, altos ingresos, no linealidades, mercado de trabajo formal.

Código JEL: D31 J62 E26

Intergenerational mobility along the income distribution: estimates using administrative data for a developing country*

Martín Leites Xavier Ramos Cecilia Rodríguez Joan Vilá

Abstract

In this paper, we provide novel estimates of intergenerational mobility for Uruguay that for the first time in a developing country, are based on administrative tax-social security records. We estimate the Intergenerational Ranking Association (IRA) and explore non-linearities. We explore alternatives to address the role of informal labour market, which is one of the main challenges to obtain precise measure of intergenerational mobility for a developing country. We have three main results: first, the level of persistence is higher when we consider individuals with less attachment to the formal labour market. Second, we find evidence of non-linearities in the degree of intergenerational persistence, being substantially higher for high-income households. Finally, there is heterogeneity by gender on the degree of intergenerational mobility, with mother-daughter transmission being the most persistent.

Keywords: Intergenerational income mobility, Inequality, Top Incomes, Non-linearities, Formal labour market .

JEL Classification: D31 J62 E26

*This research was funded by CSIC I+D 2018 at Universidad de la República. Xavier Ramos and Martín Leites acknowlge financial support of project PID2019-104619RB-C43 (Spanish Ministerio de Ciencia e Innovación)** Corresponding author: M. Leites. IECON-UDELAR and EQUALITAS. email: martin.leites@fcea.edu.uy . X. Ramos: Universitat Autònoma de Barcelona, IZA, and EQUALITAS. email: xavi.ramos@uab.cat. C. Rodríguez: IECON-UDELAR. email: cecilia.rodriguez@fcea.edu.uy. J. Vilá: IECON-UDELAR.email: jvila@fcea.edu.uy,

1 Introduction

The large inequality and poverty levels and unstable long term economic growth were distinct characteristics of developing world economies (Aghion et al., 1999; Breen and García-Peñalosa, 2005; Bourguignon et al., 2007; Alvaredo and Gasparini, 2015). In this context, assessing the degree of intergenerational income mobility provides new evidence to understand the persistent higher levels of inequality and part of their causes. Despite the relevance of intergenerational income mobility and the relationship between intra and intergenerational income inequality, there is still scarce evidence of intergenerational levels of persistence, particularly for Latin American countries.

The intergenerational mobility literature in most of the developing world is based primarily on cross-section surveys, with the limitations associated with this type of estimation: limited sample size of the surveys (homogeneous or non-representative samples) and the imprecise measure of permanent income. [1] The use of administrative records to derive intergenerational levels of mobility in rich countries, mainly in Canada (Corak and Heisz, 1999), the Nordic countries (Björklund and Waldenstöm, 2012; Björklund et al., 2009; Munk et al., 2016), Australia (Deutscher and Mazumder, 2020), and the US (Chetty et al., 2014b,a; ?; Mazumder, 2005)), show the advantages of this data to mitigate the classical problems of measurement error, attenuation and life-cycle biases (Nybom and Stuhler (2017); ?); Chetty et al. (2014a); Jantti and Jenkins (2015)).

Providing evidence based on administrative records in developing countries also contributes to understand the potential differences in intergenerational transmission mechanisms between developed and developing countries. In this paper, we provide novel intergenerational income mobility estimates for Uruguay using unique matched administrative tax and social security records for 2009-2016. More precisely, we study the intergenerational income ranking association (IRA) of matched parents and their offspring, based on a representative sample of more than 180.000 pairs of parents/offspring aged 20 to 39 years. The large size and the high-quality

[1] For Latin American countries see Dunn (2007) for Brazil, Grawe (2004) for Peru, Núñez and Miranda (2007) for Chile, Jiménez (2011) for Argentina, Jiménez (2017) for Argentina and Chile, Azevedo and Bouillon (2010) for Chile, Brazil and Peru, Araya (2019) for Uruguay and Pastore et al. (2019) for Brazil and Panama. The evidence for Uruguay is also scarce and examines educational (Sanroman, 2010; Gandelman and Robano, 2014; Urraburu, 2019) and earnings mobility (Araya, 2019).

dataset allow us to obtain precise estimates of the intergenerational transmission of incomes.

An important challenge that the study of intergenerational income mobility faces in developing countries is the instability of the formal employment and the presence of larger informal labour markets. In Uruguay the informal labour market represent one third of total employment, which is low for Latin American context.[2] Despite the advantages of using administrative records to obtain estimates on the degree of intergenerational mobility, they by construction exclude informal earnings, which may represent an important issue in our case. There is previous evidence that there is a high degree of intergenerational association between informal jobs between generations, while persistence in the informal sector is high when young people initially enter this segment of the labour market (Leites et al., 2020; Carrasco, 2012). As a result, the participation in the informal market is not random and there may be different degrees of intergenerational persistence in this segment of the labour market.

Our intergenerational income mobility measures incorporate three strategies to mitigate the consequences of the presence of individuals with less attachment to the formal labour market, typical of an economy with informal sector. First we use three alternative samples which gradually incorporate workers with less stable links to the formal labour market. The first sample focuses on parent/offspring with most stable links to the formal labour market and consider only those individual that report positive formal income over five years. Estimates based on this sample provides a suitable approximation to the levels of persistence for families with a persistent and stable participation in the formal sector. They generally have higher average incomes, and whose strong attachment to the formal market gives them access to welfare state programs, like unemployment insurance, health coverage and pensions. The others two samples gradually include families that show less attachment to the formal labour market, including some zero incomes in the years of permanent income estimation. Our tax records database prevents us to observe the actual reasons for the zero formal income: non-participation in the labour market (due to being unemployed or inactive), or active participation in the informal market.

[2] Between 2009 and 2016, the informal sector represented in Uruguay about 30% of the labour market. In 2009 the formal sector represented 67.8% of total workers and 80.6% of salaried workers. These rates rose to 74.7% and 87.9% in 2016. Informality affects mostly younger workers and women (Leites et al., 2018).

Second, the use of alternatives samples of parent-offspring pairs allows us to consider alternative definitions of permanent earnings/income. We define permanent income using two extreme assumptions. A first criteria is based on 5-years average including years with zero income. In the second criteria we use the same 5 years, but we excluded the zeros in the calculation of the average income of each generation. The first criterion establishes a lower bound from permanent income, while the second criteria changes the permanent earning and defines an upper bound of the permanent income.[3]

Third, as an alternative strategy to incorporate the presence of informal market, we follow top incomes literature (Atkinson, 2007) and construct a reference distribution that contains both formal and informal earnings for each generation from a combination of tax-survey microdata (Burdín et al., 2019). By ordering individuals in this representative distribution of total income in the economy (formal and informal), this alternative allows us to estimate the degree of intergenerational persistence from individual rankings that approximates their real position in the society.

Our baseline model estimates the average Intergenerational Ranking Association (IRA). Measures based on rankings are deemed to yield more accurate estimates of mobility than those based on income, and provide a good proxy of the link between long-run economic status of parents and their offspring's (Nybom and Stuhler, 2017; **?**; Jantti and Jenkins, 2015). The average IRA of income in the benchmark sample of persistent participation in the formal sector is 0.23, and increases when we include parent-offspring pairs with lower attachment to the formal sector (to approximate 0.27).

Consistent with previous studies, we find income mobility to be lower than earnings mobility and we also confirm the life-cycle effect by age groups. A novel result is the heterogeneity by gender of average IRA. The income persistence is significantly higher between mothers and their daughters than the rest of the pairs, which may represent a transmission of gender roles by mothers. We also examine non-linearities in a flexible manner, using both parametric and non-parametric strategies.[4] Our results show that persistence is higher both at the bottom and

[3]The first criteria is analogue to assume that when an individual declares zero annual income he/she is unemployed or inactive. The second criteria changes the permanent income definition assuming that in the years without formal income, individuals receive average formal income in the informal market.

[4]Most of the previous papers explore non-linearities in the log-log relationships. Dahl and DeLeire (2008)

particularly at the right tail of the distribution, with a IRA two-and-a-half times larger at the top decile.

Our findings are relevant for four reasons. First, they contribute to a better understanding of the income inequality persistence in the long term for a developing country, which is comparable with the more recent evidence based on administrative records for developed countries. To our knowledge, there is no previous evidence on intergenerational income mobility that uses administrative records neither for any other Latin American country nor for any other developing country. This is useful to advance in the discussion about the 'Great Gatsby curve' and the cross-national relationship between mobility and inequality.

Second, our results confirm the presence of non-linearities at the upper tail of the distribution reported in previous studies, and also provide suggestive evidence of greater persistence at the bottom of the distribution. Although there are certain channels of inequality persistence that support the idea that persistence in economic status is especially large at the top and at bottom tails of the income distribution, previous evidence about a non-linear relationship is ambiguous and focuses on developed countries (**?**Corak and Heisz, 1999; Jantti and Jenkins, 2015). We confirm the large degree of persistence for top income groups on a developing country but also novel evidence of large levels of persistence for lower income households. From a theoretical perspective, the presence of non-linearities has been related with certain mechanisms of inequality persistence such as credit constraints, the inheritance of long-term joblessness, poor human capital investment, the transmission of social capital, the inheritance of firms, capital and employers (Nybom and Stuhler, 2017; Corak and Piraino, 2010; Björklund and Waldenstöm, 2012; **?**; Piketty, 2000).

Third, we implement strategies to approach the effects that the existence of the informal market has on our estimates. We show suggestive evidence that by gradually incorporating individuals with less attachment to the formal labour market, the degree of persistence increases, suggesting that the informal sector could represent a relevant mechanism to explain inequality persistence. This evidence is important for developing countries and should thus be considered when studying intergenerational mobility in these countries, but also in developed countries

for the U.S. and Nybom and Stuhler (2017) for Sweden examine the rank-rank relationship, and they find that parental rank is particularly persistent at the very top of the distribution.

with a large share of shadow economy, such as the Southern European countries.

Finally, these results are relevant for social welfare reasons and their public policy implications. From a normative perspective, a larger intergenerational persistence has been interpreted as an imperfect indicator of lower degree of inequality of opportunity. [5] Our finding about long term inequality persistence are a relevant benchmark if an important goal for public policy is to promote equality of opportunity.

Furthermore, the higher level of persistence in the right tail provides additional arguments for public intervention. A very strong concentration of income and wealth may undermine the quality of democratic institutions and even the political equality of citizens, which provides extra argument to reduce the intergenerational persistence at the top of the income distribution (Robeyns, 2019). On the other hand, higher persistence at the bottom of the distribution and the intergenerational transmission of informal jobs may limit the future possibilities of individuals born in lower-income households. In this case, policies related to labour market institutions to achieve more permanent links with the formal sector may lead to higher levels of mobility in the lower and intermediate strata of the income distribution.

The rest of the paper is structured as follows. The next section presents the data, section 3 describes the empirical strategy and section 4 presents our main results on the intergenerational income associations. Finally, section 5 concludes.

2 Data sources

To measure the intergenerational mobility levels from administrative records in Uruguay, we built a novel database from two main sources of administrative data: (i) a sample of parental linkages (parents/sons) from social security records and (ii) the universe of income records from tax agency (*Dirección General Impositiva*). Both data sources were matched from a unique identifier for this paper [6] We start this section with a brief description of these data sources: the family links database on 2.1 and the income tax record on 2.2. Then we present our samples

[5] For a discussion of the inequality of opportunity literature see Ramos and Van de gaer (2016); Roemer and Trannoy (2016); Jantti and Jenkins (2015).

[6] We access a masked personal identifier from social security records.

and their representativeness (2.3) and finally we describe the definition of our interest variables (2.4).

2.1 Family links database

The first database contains information on the set of family links included in the set of policies administered by the social security administration (*Banco de Previsión Social*, BPS). To be part of this database, a member of the household must generate the right to others to receive a social security benefit. The most widespread case is the provision of health coverage by formal workers to other members of the household, particularly their children. However, the database also includes other programs such as Conditional Cash Transfers (CCT's), other social benefit programs for lower-income households, and transfers and benefits for a significant percentage of households that participate in the formal labour market. The diversity of policies included within the database implies a representatives along the income distribution, analysis that we delve into later.

Tables A.1 and A.2 summarize the number of observations by cohorts for sons and parents generation in the sample of family linkages database. Table A.1 shows the increasing coverage of individuals in the offspring generation by cohorts. While the cohorts before 1980, the individuals included in the sample represent approximately 40% of the total population (using the household survey as a representative sample of the total population of Uruguay), as of 1990 the database approached the universe of each cohort. The increasing coverage of the health system could explain a part of this increasing trend in the representativeness of our sample. The same performance is verified by examining the fraction of individuals with formal incomes in comparison with the universe of tax records by cohort (columns 5 to 7 of table A.1). Further, do no find important differences in the female share in each database, being close to half of the sample in most birth cohorts.

In Table A.2 a similar exercise is performed for the parents generation. Again, the fraction that represents the sample is increasing for the recent cohorts, representing less than 50% of the total individuals in the population for the generations prior to 1950, but approaching 80% of the total from the 1960 cohorts onward. In addition to the reduced number of observations, the 1940s cohorts show an over-representation of females in our sample vis-a-vis both household surveys and tax records. This lower representativeness and the higher proportion of mothers represent a potential challenge for the estimates that include these cohorts. However the presence of assortative mating in terms of earning and educational achievement in Uruguay mitigates the potential concerns to define the permanent income from only one of the parents (see Tables B1 and B2 in the online Appendix which for a representative sample of couples summarizes the association of year of education and earning.). In addition, a set of exercises are carried out to analyze the sensitivity of the results to the greater presence of mothers than fathers in our sample (see section 4.2).

2.2 Income tax records

The information about individuals' incomes is obtained from income tax data for the period 2009-2016. This data set includes income from the main formal sources for the entire population –i.e. earnings, capital income, and pensions–, and also information on socio-demographic characteristics of the individuals and on the firms in which they worked. The tax micro-data has been previously used to estimate top income shares in Uruguay (Burdín et al., 2022). As these papers show, the data set includes around 70% of the adult population aged 20 and over in each year, and allows to estimate formal income accurately, particularly at the right tail of the distribution. In turn, despite the potential evasion and elusion problems present in all administrative records, it captures a larger proportion of total income for top income groups, and in particular, for the different capital income sources.[7]

To create the income of parents and children, we merge the tax records with our sample of family linkages with a unique identifier. As noted above, the sample of parents/sons links arises from a set of social security benefits, involving more than one member of the household. One potential concern is that these policies target certain groups of the population, which could potentially result in a selected sample. To examine whether our sample is selected, we compare it with the universe of tax records as a way of measuring the performance of our sample in

[7]Unlike previous work for the estimation of top income shares, our estimations only incorporate the set of incomes reported by each individual or third-party report, without making imputations of non-nominative sources of income.

terms of formal income.

Tables 1 and 2 shows some descriptive statistics for each generation of our sample, the universe of tax records and the household survey for 2012.[8] In the case of sons/daughters, we split the sample into two age groups: 20-29 and 30-39 years old. Given the larger representativeness of the recent cohorts, in the 20 to 29-year-old group, our sample represents close to 75% of the universe present in the tax records, while this share falls to less than 40% for the 30-39 age group. In terms of income, there are slight differences in the average levels between the sample and the universe of tax records in both age groups, with income being barely lower for individuals included in the parents/offspring's database. There are also no relevant differences in terms of age and gender between the three databases used, although the presence of women is greater in the household survey, probably because of the presence of a large number of individuals without labour income.

Table 2 shows the same exercise for the parents' generation. In this case, our sample includes more than half of the universe of tax records aged between 45 and 65 years. Income is again slightly lower, but the difference is expected given the sample includes only parents and we compare with the universe of tax records. As noted above, the participation of women in the sample is larger than in the tax records (54% vs 49% of total). As we mentioned earlier, this over-representation of mothers is explained almost exclusively by the most distant cohorts (prior to 1950).

[8]The selection of the reference year for comparison does not change the main conclusions of this analysis.

Table 1: Summary Statistics: Number of observations and participation of females by age group (Sons, 2012)

	Aged 20-29					Aged 30-39				
	Mean	Median	SD	p10	p90	Mean	Median	SD	p10	p90
Panel A: Family linkages sample										
Labour income	151,084	122,587	145,489	9,894	320,891	254,852	187,886	284,235	8,657	542,446
Fraction no income	4.2%	-	-	-	-	7.7%	-	-	-	-
Fraction female	45.6%	0.0%	49.8%	0.0%	100.0%	48.5%	0.0%	50.0%	0.0%	100.0%
Age	24.3	24.0	2.8	20.0	28.0	34.2	34.0	2.9	30.0	38.0
Observations			247866					125770		
Panel B: Tax records										
Labour income	153,927	123,884	152,432	10,083	327,169	253,759	181,994	301,172	9,896	539,131
Fraction no income	4.0%	-	-	-	-	6.9%	-	-	-	-
Fraction female	44.9%	0.0%	49.7%	0.0%	100.0%	46.6%	0.0%	49.9%	0.0%	100.0%
Age	24.6	25.0	2.8	21.0	28.0	34.4	34.0	2.8	30.0	38.0
Observations			335594					333974		
Panel C: Household Survey										
Labour income	124,848	91,487	153,050	-	329,177	205,135	156,674	245,256	-	468,147
Fraction no income	43.4%	-	-	-	-	32.4%	-	-	-	-
Fraction female	50.8%	100.0%	50.0%	0.0%	100.0%	52.1%	100.0%	50.0%	0.0%	100.0%
Age	24.3	24.0	2.9	20.0	28.0	34.5	35.0	2.8	30.0	38.0
Observations			467378					454442		

Notes: The table shows the number of observations according to the birth cohort of the parents generation on the different database: family linkages sample (column 2), household survey (column 3) and tax records (column 6). It also shows the fraction that the observations included in the sample represent with respect to the other databases, and the representativeness by gender. *Source:* Based on social security records (BPS), tax records (DGI), and Household Survey (INE).

Table 2: Summary Statistics: Number of observations and participation of females (Parents aged 45-65, 2012)

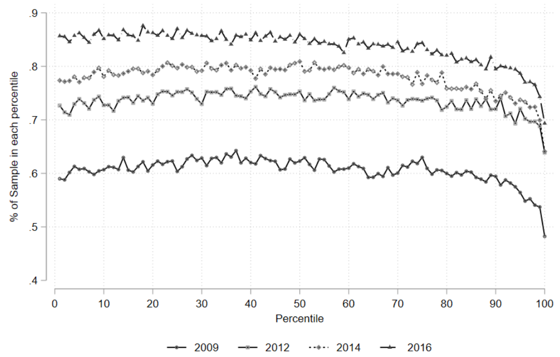
	Mean	Median	SD	p10	p90
Panel A: Family linkages sample					
Labour income	294,001	149,397	522,586	-	716,703
Fraction no income	24.9%	-	-	-	-
Fraction female	54.1%	100.0%	49.8%	0.0%	100.0%
Age	53.5	53.0	6.1	46.0	62.0
Observations			319659		
Panel B: Tax records					
Labour income	323,964	178,508	603,539	24,092	724,160
Fraction no income	4.1%	-	-	-	-
Fraction female	49.4%	0.0%	50.0%	0.0%	100.0%
Age	54.2	54.0	6.3	46.0	63.0
Observations			583061		
Panel C: Household Survey					
Labour income	222,021	139,715	284,950	-	550,699
Fraction no income	28.5%	-	-	-	-
Fraction female	53.5%	100.0%	49.9%	0.0%	100.0%
Age	53.9	54.0	6.2	46.0	63.0
Observations			781577		

Notes: The table shows the number of observations according to the birth cohort of the parents generation on the different database: family linkages sample (column 2), household survey (column 3) and tax records (column 6). It also shows the fraction that the observations included in the sample represent with respect to the other databases, and the representativeness by gender. *Source:* Based on social security records (BPS), tax records (DGI), and Household Survey (INE).

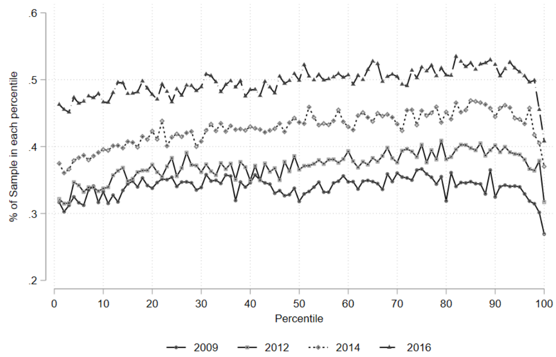
To explore the degree of representativeness of our sample throughout the distribution of formal income, Figure 1 shows the percentage represented by the individuals identified in our sample of family linkages by percentile of formal labour income from tax records. Panels a and b includes the individuals for the generation of children in the same age groups (20-29 and 30-39), and the panel c for the generation of parents.

In all cases, adequate representativeness is achieved throughout the entire income distribution. In the case of the younger generation, there is an under-representation in the upper percentiles, but in any case the percentage located in these strata is similar to the average. In the case of parents, representativeness is close to the average throughout the distribution, with the exception of the last percentile, where the percentage of individuals in our sample increases. In other words, our sample is evenly distributed across the distribution of formal incomes that derive from tax records and hence does not over- or under-represent any income strata of society.

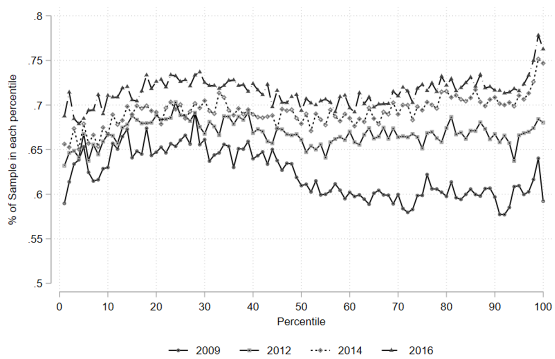
Figure 1: Fraction of individuals from the universe of tax records included in the sample of parents links, by percentile of labor income and year



(a) Sons aged 20-29



(b) Sons aged 30-39



(c) Parents

Notes: The figure shows the percentage of observations included in the sample of parents links database by percentile of labor income in tax records. Panel (a) and (b) shows this fraction for the sons generations (by age group), and panel (c) for parents. *Source:* Based on social security records (BPS) and tax records (DGI).

2.3 Three samples to address degrees of attachment to the formal labour market

To address the presence of individuals with less attachment to the formal labour market, typical of an economy with significant presence of informal sector, we use alternative samples which gradually incorporate workers with less stable links to the formal labour market. Our measures of income mobility are based on three main samples, which differ from each other in the number of formal income records required to be part of it (see Table 3).

The strict sample used in the estimates incorporates the set of individuals which satisfies the following conditions: a parent/offspring link from social security records and five consecutive positive labour income records within the age range we study for each generation (between 20 and 39 years for children and 45 to 65 years for parents). Our choice of 5 years is in line with previous studies, which point out that windows of 5 years provide reasonable estimates of permanent income.

This definition of permanent income represent our baseline sample for estimates, and is comparable with previous international studies. This income measure, however, excludes a significant number of workers who register a less permanent attachment with the formal labour market. The presence of a large informal labour market in the developing world provides an additional alternative to formal employment, causing a group of workers to fluctuate between one labour market or another. To address this issue, we use two alternative samples of paired parents and offspring with more flexible income requirements: a second sample, that we shall call *Extended* sample, includes all individuals with at least two positive labour income records, while a third sample, termed *Universal* sample, includes individuals with at least one formal income. Including these additional individuals in the estimation sample help us a better approximation of the degree of intergenerational persistence at the bottom of the distribution, where individual transitions between the formal and the informal labour markets are more frequent.

Table 3: Criteria for the construction of the three samples

Sample	Income condition	Other requirements	N
Strict	5 positive earnings	Offspring aged 20 - 39; Parents aged 40 - 65	154,030
Extended	2+ positive earnings	Offspring aged 20 - 39; Parents aged 40 - 65	257,790
Universal	1+ positive earning	Offspring aged 20 - 39; Parents aged 40 - 65	300,565

Finally, in order to perform a set of robustness check analyzes we use alternative criteria to the construction of the samples. First, we include the 8 years of tax records for the construction of the samples (2009-2016) instead of the five years windows using on our three main samples. This implies the inclusion of individuals with weaker attachment to the formal sector, particularly in the strict sample. Second, in the case of Universal sample, we sequentially include parents and children that report zero incomes in the period, and in consequence individuals with zero permanent income (see Panel A on Table A.3 for details).

2.4 Definition of income variables and reference distribution

Our main income variable is total income, measured as the sum of three income sources: formal labour income, capital income, and pensions. As it is common in the literature, we will also examine mobility of earnings, which include formal wages and self-employed income. Both concepts are measured before taxes and only incorporate taxable income, which excludes, for example, income from owner-occupied housing and non-contributory public transfers.[9] To avoid temporary income fluctuations we average yearly incomes over 5 consecutive years. This criteria represent our baseline permanent income for the three samples derived in previous section.

In addition to the potential biases generated by measurement error, explained above, one of the main concerns in the mobility literature is the presence of life cycle bias due to the observation of incomes at early ages (Haider and Solon, 2006; Nybom and Stuhler, 2017). Firstly, previous studies show that measures associations tend to stabilize when offspring are about 30 years old, which represents a key aspect for our case (Nybom and Stuhler, 2017; Chetty et al., 2014a,b; Böhlmark and Lindquist, 2006). Due to the reduced period of the tax income database (2009-2016), we average the income at the most advanced age that we observe in the case of children (in most cases for the 2012-2016 period). When computing life-cycle income or earnings we give priority to offspring aged 30 to 39 years and parents aged 45 to 65 years, as we expect life-cycle bias to be minimal at these age windows. In turn, in the average estimates we incorporate a simple weighting by age of the children to correct the possible imbalance at the age level of each group. This weighting adjusts the observations of our sample to replicate the age structure observed in the universe of tax records.

On the other hand, previous work highlights rank-rank estimates tend to be less sensitive to life cycle bias, with less attenuation bias and more stability over age (Nybom and Stuhler, 2017; Chetty et al., 2014a; Mitnik et al., 2015). Based on these recommendations we use percentile

[9] Since we use income ranks in our main analysis, using pre- or post-tax income should yield very similar results as tax rates do not usually exceed 100%, which avoids rerankings from pre- to post-income.

rank as our preferred permanent income measure. The key decision for the elaboration of this indicator is to determine the income distribution to be used for the construction of percentiles. The literature typically uses the sample of matched parents and children as a reference distribution to compute percentiles. Given the presence of the informal sector in Uruguay, we add the individuals with informal incomes to income distribution of tax records, following the spirit of population and income controls from the top incomes literature (Atkinson (2007)).[10] This procedure implies adding about 25% of additional population and income from Household Survey data. Figure A.3 illustrates the way we construct our reference income distribution.

The distribution that result from this combined database (tax and households surveys) is the best approximation to the the overall income distribution in Uruguay for each generation. Then we rank each generation of our samples (strict, extended and universal) on this distribution to derive the rankings for the intergenerational mobility estimations. In this way, the position reached by individuals in the distribution of their generation also depends on the percentage of informal workers and their incomes. We also ensure that the movements observed between percentiles reflect changes in the status of individuals in the entire distribution.

Figure 2 shows the effect of this procedure on the positions reached by the generation of children between 30 and 39 years old in each of the three samples. Figures A.1 and A.2 in the appendix show this same exercise for children between 20 and 29 years old and the parents' generation. In each panel, the three lines show the percentage of individuals from each of the samples, according to the percentiles ordered by the two income concepts. If the individuals of the different samples were uniformly distributed throughout the distribution of total formal income (panel a) or formal and informal income (panel b), they should represent 1% of each percentile.

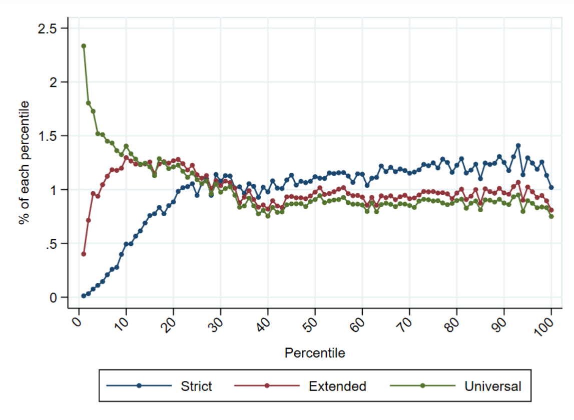
In both panels, we observe a sub-representation of the individuals of our strict sample in the first three deciles of the distribution. This is consistent with the requirements to be part of this group (5 consecutive positive earnings), which implies restricting the sample to workers with a high attachment to the formal labour market. The flexibility of these criteria increases the

[10] An addition concern of using the parent/child sample itself for the analysis could be that our sample of family linkages were not representative of the population in the formal sector. However, as we show in Section 2.3 this is not the case.

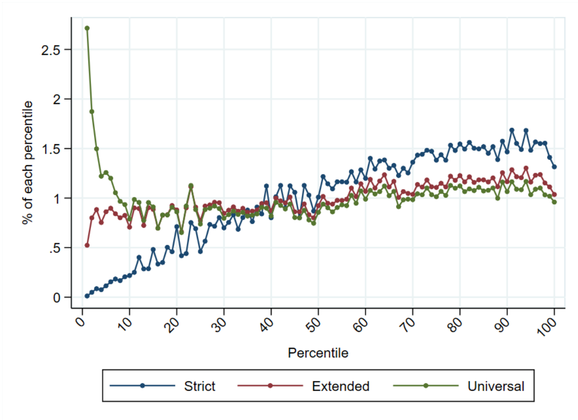
number of observations at the bottom of the distribution in both the extended and universal samples.

Between panels (a) and (b) we can see the difference, for each of the samples, of incorporating informal workers of each generation to build the reference distribution (percentiles where individuals are ordered). The most noticeable difference is the lower number of individuals in the first percentiles in panel (b), since the incorporation of informal income runs to the right of the distribution of the individuals belonging to our samples. The patterns followed by sons/daughters aged 20 to 29 and by the generation of parents are similar both between samples and the effect caused by the incorporation of informal incomes to the reference distribution (Figures A.1 and A.2).

Figure 2: Fraction of individuals in each percentile by sample (Sons aged 30-39)



(a) Formal income distribution



(b) Formal and informal income distribution

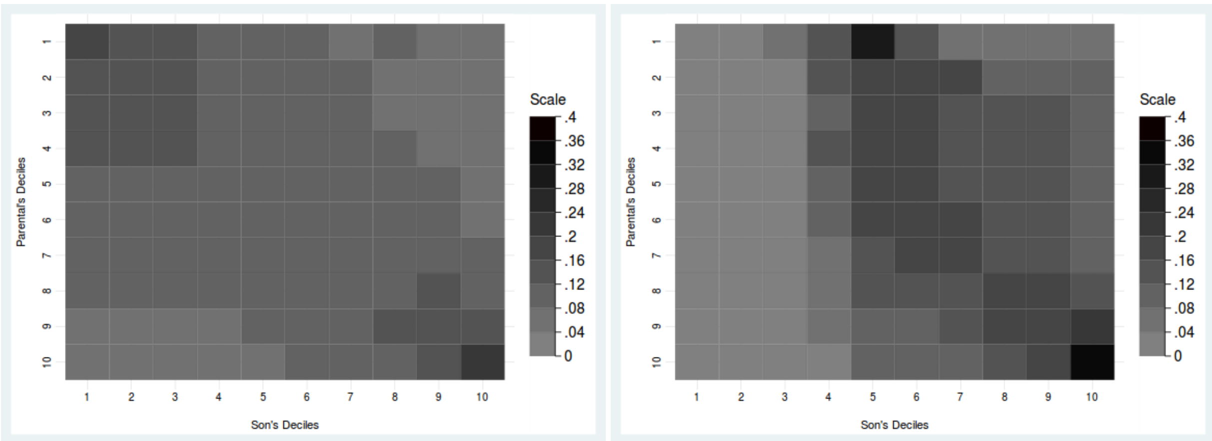
Notes: The figure shows the percentage of observations included in the sample of parents links database by percentile of labor income in tax records. Panel (a) and (b) shows this fraction for the sons generations (by age group), and panel (c) for parents. *Source:* Based on social security records (BPS) and tax records (DGI).

The use of a reference distribution that includes the universe of formal and informal income may also have consequences on the degree of intergenerational persistence. As a first approximation to illustrate this effect, in Figure 3 we show the transition matrices for the generation of children between 30 and 39 years of age for our strict sample using the own sample as reference (panel (a)) and our reference distribution (panel (b)). Appendix Figures A.4 and A.5 show analogous exercises for the extended and universal samples.

In both cases the main diagonals concentrate higher frequencies, which reflects the intergenerational persistence in each decile. However, there are relevant differences in the low tail of the distribution between both panels. The higher degree of persistence observed in the first deciles on panel (a) disappears when we include the set of formal and informal incomes for the construction of the rankings of our samples (panel b). However, this is explained by the scarce number of observations in the first three deciles of both generations when using our reference distribution (as we already commented in Figures 2 and A.2). As expected, in this case we find a higher degree of persistence, which supports that this reference distribution provides more accurate movements in the real positions in the society of parents and children.

Figure 3: Impact on transition matrices from including individuals from the informal market in the income distribution. Strict Sample. Offspring aged 30- 39

(a) Own sample (b) Formal/informal reference distribution



Finally, we consider a set of alternatives definitions of permanent income used as robust checks. First, as the Extended and Universal samples could include yearly observations without formal income, we construct an alternative permanent income for these sample as the average of positive income records. Second, as alternative measure of permanent income we define a new ranking based on real position in the global distribution (for details see Panel B in Table A.3).[11]

[11] The real position orders parents and children in their own generations based on their 5 years average income and it is then normalized by the size of the respective generations.This provides higher variability within each percentile, compared to the used rank-rank estimation.

3 Econometric model

The study of intergenerational income mobility aims at obtaining a characterization of the joint distribution of children's and parents' incomes, $f(y^{parents}, y^{children})$. This distribution can decompose into two components: (i) the joint distribution of parents and children ranks, formally known as the copula of the distribution, and (ii) the marginal distributions of parents and children income (Jantti and Jenkins, 2015). While the marginal distributions ($f(y^{parents})$ and $f(y^{children})$ determine the level of income inequality within each generation, the copula is a key determinant of mobility across generations. Jantti and Jenkins (2015) suggest that global mobility may be interpreted as the transformation linking $f(y^{parents})$ and $f(y^{children})$). To measure intergenerational income mobility prior research has considered both components, the copula and the marginal distributions, but with different emphasis. Our strategy focuses on the copula, but as we explain below it also considers the intragenerational inequality to build the ranking based on the entire income distribution.

Usually $y^{parents}$ and $y^{children}$ represent the vector of income, and ($f(y^{parents})$ and $f(y^{children})$) are the global income distributions of each generation. As we discussed in section 2, we only have information about individuals with formal incomes/earnings, which represent a sub-sample of the entire distribution. We define our vector of formal incomes/earnings as $y^{parents}_{formal}$ and $y^{parents}_{formal}$. A measure of formal income mobility would focus on the transformation linking $f(y^{parents}_{formal}))$ with $f(y^{children}_{formal}))$. However, this strategy does not consider the role of informal sector on intragenerational income inequality. This would not be a problem if $f(y^{parents}_{formal}) = f(y^{parents})$ and $f(y^{children} = f(y^{children}_{formal}))$ but, as we pointed out in section 2, this is not the case. To address this issue, we use information about $f(y^{parents})$ and $f(y^{children})$ when defining the ranking of formal earners in the entire distribution. That is, we define $P_i^{children}$ and $P_i^{parents}$ as the percentile of formal earners in the entire income distribution of each generation. This strategy combines information of the copula of $f(y^{parents}_{formal}, y^{children}_{formal})$ with information of the marginals of the $f(y^{parents}, y^{children})$. This allows us to consider the role of informal incomes/earnings on intragenerational income inequality and allows us to obtain a more complete measure of intergenerational income mobility. Note we use the same reference distribution for the three samples, which allows us to obtain comparable individuals movements when we measure mobility for the alternatives samples. .

Our baseline model estimates the average Intergenerational Ranking Association (IRA):

$$P_i^{children} = \beta' P_i^{parents} + f(\gamma,\, age^{children}, age^{parents}) + v_i \tag{1}$$

where $P_i^{children}$ and $P_i^{parents}$ identify the relative position of the family i in the income distribution of the generations of children and parents, respectively. Measures based on ranks within each generation provide a good proxy of the long-run economic status of both parents and their offspring. Note that the magnitude of the IRA is insensitive to changes in the formal income inequality within each generation (if there is no exchange of positions).

We characterize mobility based on the slope (β') of this rank-rank relationship, which provides an average measure of the strength of the association in the copula of the the joint distribution (**?**). The assumption of linearity assures that (β') is both locally and globally informative. According to Chetty et al. (2014c), β' can be interpreted as the average difference in the mean percentile/position rank of children from the richest families vs. children from the poorest families.

From the point of view of the empirical approach, the estimate of the IRA has some advantages in measuring income mobility. First, attenuation bias and life cycle bias are considerably weaker in rank-based measures (compared with the standard intergenerational log-log intergenerational elasticity (IGE) of child income with respect to parent income (Nybom and Stuhler, 2017)). Second, previous papers suggest that estimates of the IRA are comparatively more stable, less sensitive to the samples (and the presence of outliers in the tails), and to the specification choices (e.g. the way in which earnings/income are defined and to the treatment of zero incomes in particular) (Dahl and DeLeire, 2008; Nybom and Stuhler, 2017; Chetty et al., 2014c; Mazumder, 2005). Third, as Nybom and Stuhler (2017) noted, classical measurement error attenuates log-linear measures through its effect on the variance of observed incomes, but the variances of observed and true ranks are equal by definition. However, a drawback is that v_i follows a non-classical error distribution as top (bottom) ranks cannot be overstated (understated). Furthermore, since IRA looks at rank movements across generations, similar

IRA estimates could have different welfare implications as income distances between ranks may differ. This weakness could be problematic when making comparisons across countries with different distributions of income.

Following Björklund et al. (2009), to assess the IRA at different points of the distribution, we extend equation (1) using non-linear regressions by means of a spline function with pre-defined knots, which identify the position in the distribution of parental incomes at which the slope is allowed to change. Following previous work, we employ four (P25, P50, P75 and P90) or five knots (we add P99). [12]

$$P_i^{children} = \beta' P_i^{parents} + \delta_p \sum_{p=25}^{p=99} (P_i^{parents} - P_p) + u_i \qquad (2)$$

When we use the extended and the universal samples we include an extra knot point in equation (2) at the lower tail of the distribution (P10), as sample sizes are larger.

As Hertz (2009) noted, the interpretation of the coefficients δ_p in equation (2) is different than the average IRA coefficient from equation (1). In this case, the comparison of local slopes alone does not provide information about the intergenerational persistence or about the presence of differences in expected incomes of offspring from poor, middle, or rich households. For instance, we can not conclude anything about the differences in intergenerational persistence between both groups when the local slope is steeper for one group than for the other. The coefficients provide information about the local relationship between offspring's and parents' ranks. For instance, it allows us to assess if the transmission of a given increase in the permanent income parents to the expected permanent income of their offspring is equally large for rich and for poor parents. In this sense, the heterogeneity in slopes helps both unpack the average IRA and also provides local marginal effects of parental permanent income for parents at different percentile groups.

The splines model imposes continuity in the relationship between parents' and children's income. To relax the continuity assumption, we use a set of linear specifications that allows

[12] Our baseline specification does not include P99 because we obtain imprecise estimates due to small sample in that fractile.

changing the slope and the intercept for each one of the segments previously considered.

Finally, we also use local polynomial regressions to explore nonlinearities (Cleveland, Devlin and Grosse, 1988). Our model can be written as:

$$E(P_i^{children}|P_i^{parents}) = F(P_i^{parents}) \tag{3}$$

where F is the smoothing function that determines the expected rank of the offspring conditional on parental rank.

4 Main results

This section reports and discusses our main findings. First the average IRA for the alternatives baseline samples is presented in subsection 4.1, while subsection 4.2 explore whether gender is a source of heterogeneity in the degrees of intergenerational persistence. Finally, subsection 4.3 explores the presence of non-linearities in the functional form of intergenerational income persistence. Each main result is followed by the corresponding robustness analysis.

4.1 Average mobility

In this section, we show the average results of the degree of income mobility for our three samples, the age group of sons' generation and incorporating the total income or only labour income. In each regression, we include controls for sex of children and the age and sex of parents and weight the sample to take into account the age composition of our samples.[13]

Table 4 shows average Intergenerational Ranking Association (IRA) estimates (from equation (1)) for children aged 30-39, our preferred age group, and for total income. The average level of persistence is 0.23 in the sample with more stable attachment with the formal market, and increases when gradually include individuals with a less stable participation in the formal

[13] As we noted in Section 2, offspring aged 20 to 29 are over-represented. We weight each individual by the inverse of the number of individuals in the same age range.

market (0.26 in extended and 0.27 in universal sample).[14] This change in the average mobility level is the result of two effects that work in opposite direction: a potential larger persistence between parents/sons pairs with higher instability of their formal job (and probably lower incomes); and, on the other hand, the inclusion of more volatile income/earnings, which could reduce the IRA through the classical attenuating bias. In addition, the average contribution to the IRA also depends on the position in the distribution of income of the new individuals included and their degree of persistence.

The use of the same reference distribution for the three baseline samples allows us to associate differences in mobility levels to the new individuals incorporated into the sample, and not to changes in the income distribution used to build the rankings. In the extended and universal samples we include individuals located in the middle and mainly in the lower tail of the distribution (see Figure 3 in section 2.3), so these results are a first indication of a higher degree of association in the lower end of the distribution. Table A.5 in the Appendix shows some signs of intergenerational transmission of formality status, particularly for children of parents with no contributions in the formal market, that could reflect an additional mechanism of intergenerational transmission for this sub-groups of parents/sons pairs.

Panel B of Table 4 shows the IRA coefficient estimates when we use the own sample as reference distribution. These estimates allow comparison with the previous literature, which usually uses the sample itself for the construction of rankings. Notice that in this case the ranks of the three samples change the reference distribution. The main difference with the previous results is the higher average IRA in the case of the strict sample (0.26 vs. 0.23)[15], which in these estimates approximates the levels of persistence in the two alternative samples (extended and universal). In this case, the average levels of intergenerational mobility found remain stable when we go from the strict to the universal sample.

In Table A.6 in the Appendix we reproduce the previous results but for the youngest cohort of children (aged 20 to 29 years). In this case we found lower levels of persistence, confirming the effect found in most of the previous literature. On the other hand, we find larger level

[14]Difference in estimates are statistically significant, as the p-values in Panel c show.

[15]Difference in estimates are statistically significant, as the p-values in Panel c show.

Table 4: Average IRA for total income by sample. Children aged 30-39

	Samples		
	Strict	Extended	Universal
PANEL A: Global distribution			
IRA	0.2322***	0.2668***	0.2720***
N	30,193	82,519	98,977
PANEL B: Sample distribution			
IRA	0.265***	0.269***	0.268**
N	30,193	82,519	98,977
PANEL C: Mean test difference's			
vs. Strict sample	-	37.16	46.18
Chi (p-value)	-	(0.000)	(0.000)
vs. Extended sample	-	-	6.13
Chi (p-value)	-	-	(0.013)
Global vs sample	119.24	1.70	2.63
Chi (p-value)	(0.008)	(0.193)	(0.105)

Note: The dependent variable is offspring's total income. Coefficients are OLS estimates. Controls: children's age, parental age and sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

of persistence for income than for earnings (see Table A.7 in the Appendix).[16] This result is also consistent with previous findings and reflects the impact of capital income on the intergenerational persistence levels. The novel result is the confirmation of the increasing degree of persistence (IRA) when the samples include individuals with a weak attachment to the formal labour market.

Robustness analysis: Now we present part of the different sensitivity analyzes performed for the average IRA estimates. The tests include (i) the incorporation of a greater number of years of the tax records for the construction of the samples, (ii) the inclusion of additional controls, (iii) the modifications of the definition of permanent income and (iv) the use of alternative measures to estimate the degree of intergenerational mobility.

First we replicate the estimates on average but using the 8 years of tax records for the construction of the samples instead of the five years windows of our benchmark estimations (see Table A.3 in the Appendix for details).[17] This implies the inclusion of individuals with weaker attachment to the formal sector, particularly in the strict sample, and increases the average IRA

[16]The null hypothesis that the income-based and the earnings-based IRA estimates are equal to each other is rejected in all cases (see Table A.8 in the Appendix).

[17]This change does not imply a modification of the permanent income definition (average of 5 years), but includes incomes from years outside the five-year window chosen for each generation.

levels (for details see Table B3 in the online Appendix). Note that estimates based on theses additional samples reinforce the idea that average IRA increase when we include individuals with weaker attachment to the formal sector. Secondly, the inclusion of dummies to control for the number of years with positive earnings for parents generation do not modify the main results presented previously (see B4 in the online Appendix). Thirdly we excluded the zeros in the computation of the permanent income definition (Table B5),[18] and incorporate estimates based on the discrete positions of individuals (Table B6) showing very similar pattern than our preferred specification.

Finally, in Table B7 we estimate the intergenerational income elasticity using log-log specification (IGE). In this case, the levels of persistence are considerably lower than the previous results, potentially due to the sensitivity of this indicator to the definition of permanent income and measurement errors. Chetty et al. (2014a) argue that log-log measure provides more unstable measure of mobility and is more sensitive to the treatment of the zero. The latter is particularly relevant in our case, where by construction we include a significant number of individuals with zero income in our extended and universal sample.

4.2 The gender of parents and children matter

Gender has been found to be a source of heterogeneity in the transmission of economic advantage between parents and children. Previous papers suggest various reasons. In the presence of assortative mating, women from better-off backgrounds tend to marry richer partners and are more likely to work fewer hours or not to work at all. If this is the case, rank-rank slope estimates should be lower for daughters than for sons. In addition, the earnings and income distribution is typically more compressed for women than for men. While the inequality in the marginal income distributions of parents and children do not have any bearing on IRA estimates, this implies lower intergenerational income elasticities for women, ceteris paribus.

Previous studies based on administrative records show that gender differences in intergenerational mobility are sensitive to the way one measures mobility (IGE or IRA) and to the

[18]This criteria changes the permanent income definition assuming that in the years without formal income, individuals receive an equivalent level of income in the informal market. In this sense, it represents an upper bound of the permanent income level.

definition of the outcome (Mitnik et al., 2015; Chetty et al., 2014a; Mazumder, 2005; Dahl and DeLeire, 2008). Using the same methodological choices than ours, Chetty et al. (2014a) finds smaller IRA for daughter than for sons in the US. Given that the disparity in intergenerational mobility by gender found in previous work refers exclusively to rich countries, in this section we examine for the first time whether this is also the case in a poorer country with different institutions and potentially culture differences about the role of gender in society. We analyse and examine whether intergenerational mobility differs by the gender of children, the gender of the parents and both.

The IRA estimates reported in Table 5 indicate that intergenerational persistence is about 28% larger for female than for male children, irrespective of the gender of parents. [19] IRA difference between gender slightly decreases (up to 25%) on the alternative samples. Our estimates for Uruguay, then, are not in line with previous estimates for the US and challenge the standard arguments to understand gender differences in intergenerational mobility. [20]

Table 5: Average IRA for total income by gender of children, and sample (30-39 age group)

	Strict		Extended		Universal	
	Daughter	Son	Daughter	Son	Daughter	Son
PANEL A: Global distribution						
IRA - income	0.251***	0.196***	0.300***	0.239***	0.299***	0.230***
N	15,031	15,162	40,395	42.124	48.649	50.328
PANEL B: Mean test differences						
Diff. between genders	22.10		85.71		108.48	
Chi (p-value)	(0.0000)		(0.0000)		(0.0000)	

Note: The dependent variable is offspring's total income. Coefficients are OLS estimates. Controls: children's age, parent's age and sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table 6 shows the levels of persistence by the gender of the parent with the larger income (fathers in panel A and mothers in panel B), a dimension not explored in previous studies. In both cases, we confirm an increase in the degree of persistence by making the criteria of the samples more flexible (from the strict to the extended and universal sample), although we

[19] The IRA estimates for earning is reported in Table A.9 in the Appendix an confirm this result. Furthermore the alternatives estimates presented in Tables B3, B4, B5, B6 and B7 in the Appendix confirm the mentioned difference by children gender.

[20] Bernuy and Esteve (2019) provide evidence that Uruguay present less educational homogamy in young couples than US. Furthermore, in Uruguay women with tertiary educational levels register a labour rate participation similar to men. Furthermore, for this group of women their labour behavior is similar than men and is not associated with the typical role of "secondary worker" (Espino et al., 2017).

observe a higher average level of persistence in the case of the fathers than in that of the mothers, particularly in the strict sample. On the other hand, only for the sub-sample of mothers (panel B) did we observe IRA levels significantly higher in the case of daughters (between 40% and 50% larger). This evidence suggests a intergenerational transmission of the role model from mothers to daughters which would explain the greater persistence of incomes.

Table 6: Average IRA for total income by gender of parents and children, and sample (30-39 age group)

	Strict			Extended			Universal		
	Average	Daughter	Son	Average	Daughter	Son	Average	Daughter	Son
PANEL A: Fathers (maximum income)									
IRA	0.258***	0.261***	0.255***	0.271***	0.269***	0.272***	0.270***	0.265***	0.275***
	(0.0125)	(0.0189)	(0.0166)	(0.00715)	(0.0104)	(0.00986)	(0.00649)	(0.00930)	(0.00906)
N	8.035	3,795	4,240	22,698	10,765	11,933	27,311	13,148	14,163
F Test Daughters vs Sons		0.05			0.62			0.2849	
(p-value)		(0.8205)			(0.8718)			(0.4318)	
PANEL B: Mothers (maximum income)									
IRA	0.217***	0.250***	0.185***	0.256***	0.298***	0.215***	0.263***	0.309***	0.219***
	(0.00653)	(0.00954)	(0.00889)	(0.00410)	(0.00580)	(0.00578)	(0.00381)	(0.00533)	(0.00543)
N	22,164	11,241	10,923	59,834	29,639	30,195	71,680	35,510	36,170
F Test Daughters vs Sons		25.08			103.19			139.14	
(p-value)		(0.0000)			(0.0000)			(0.0000)	

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Robustness analysis: A possible concern for the results presented so far is the unbalanced gender composition of our parents' sample (mostly made up of women, about 70% of the total). In the first place, to explore the sensitivity of the results to our gender parents' composition we compare the estimations between the sub-sample of households where we have both parents against households in which information of only one parent is available. Table A.10 in the appendix shows the results. Households with both parents show larger persistence levels, particularly in the case of the strict sample. On the other samples the differences are small and the results confirm the commented pattern when we extend the sample to individuals with lower attachment to the formal sector.

On the other hand, the gender imbalance in the composition of our parents could have consequences in the observed patterns of persistence between mothers and daughters. We carry out three additional exercises as robustness [21]: (i) we replicate the results without restricting

[21] As an additional robust check we estimates the specification in Table 6 for the sub-samples of parents from

the sample to the largest income of parents (Table A.11 in the Appendix), (ii) we estimate the IRA coefficients for the sub-sample of households with both parents (Table A.12), and (iii) for the sub-sample of households with at least a son and a daughter (TableA.13). Note that the three strategies addresses the potential compositional gender bias, but the latter case also controls for for unobserved heterogeneity across families. In all cases we confirm that the magnitude of persistence is significantly higher in the daughters/mothers link, while gender differences are not statistically significant in the case of the fathers.

4.3 Mobility differs depending on how rich your parent is

Previous evidence for rich economies shows heterogeneous levels of persistence at different points of the income distribution (Deutscher and Mazumder, 2021). Rich and poor parents do not have the same resources and mechanisms to transmit economic advantage to their children. In the lower tail of the distribution, credit constraints are a plausible mechanisms to explain the concave relationship between child and parent incomes.[22] On the other extreme of distribution, the unequal inheritance of wealth and employers; and the segregation across neighbourhoods can generate larger intergenerational persistence at the top of the income distribution (see for example, Björklund and Waldenstöm (2012); Corak and Piraino (2010); Durlauf and Seshadri (2018); Chetty and Hendren (2018). Our analysis is also motivated by the fact that our alternatives samples include individuals with different attachment to the formal sector, who in turn occupy different positions in the income distribution.

In this section, we use complementary empirical strategies to explore the potential non-linearities in the transmission process. As outlined in Section 3, we first estimate non-linear parametric spline regressions with pre-defined knots (equation (2)). Then, we use a more flexible (but still parametric) strategy and fit a separate standard IRA regression model (equation (1)) for each fractile defined by the knots we use in the spline regressions. [23] Finally, we analyse the

the more recent cohorts. As we mentioned in section 2 the parental gender imbalance disappears for the this group. The results remain unchanged (available upon request of authors).

[22]The early contribution by Becker and Tomes (1986) explore this mechanism. More recent work by Grawe (2004), however, puts into question the relevance of credit constraints to produce a non-linear intergenerational transmission of incomes.

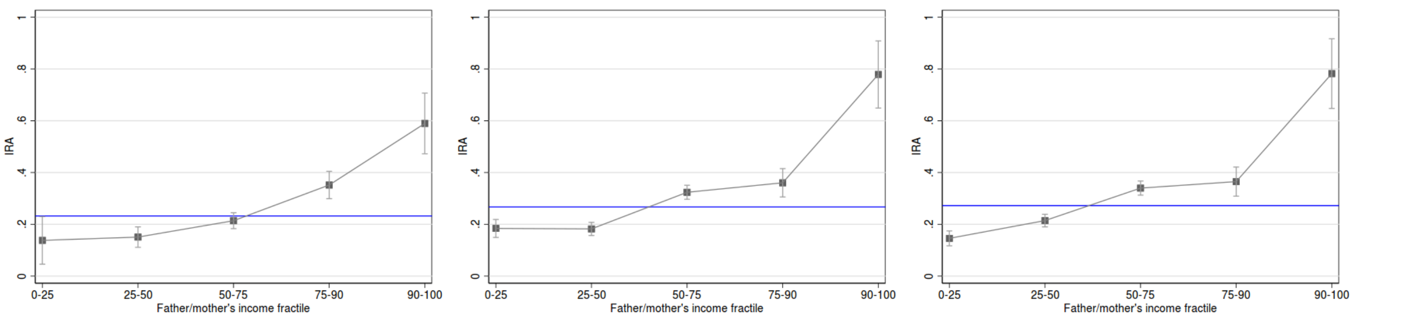
[23]Note that this strategy lifts the assumption of a common intercept for all fractiles.

expected income rank of children conditional on their parents' rank, which results from flexible non-parametric kernel-weighted polynomial regressions (equation (3)).

In Figure 4 we examine heterogeneity in slopes along the distribution with non-linear parametric spline regressions with pre-defined knots at percentiles P25, P50, P75 and P90 (estimated coefficients are also shown in Table A.14).[24] In three samples, higher levels of persistence are noticeable in the upper tail of the distribution. For the sample with largest attachment to formal market, the change of expected rank due to a 10 percentile increase in the permanent income of parents goes from 1.4 in the bottom quartile to 5.9 percentile at the top decile. The change in the slope at the upper tail is most evident when we include an extra knot in our model specification at the top 1%.[25]

Figure 4: Intergenerational Income Rank Associations: Non-linear spline regression estimates for income ranks (children aged 30-39). Knots at P25, P50, P75, P90

(a) Strict Sample (b) Extended Sample (c) Universal Sample



Note: The dependent variable is offspring's total income. Coefficients are OLS estimates. Controls: children's age, parental age and sex. 95 Confidence Interval.
Source:

The inclusion of individuals with weaker attachment to the formal sector in the extended and universal samples does not change the main message of non-linear transmission of economic advantage. Note that the extended and universal sample expand the range of variation of both the father's and son's permanent income while the strict sample concentrate patents and children with middle and high income. This changes in the sample allows to better capture slope changes and identify the higher persistence in the high tail in relation to the rest of income

[24] We replicate this estimates with different pre-defined knots: including a p10 knot on the extended and universal sample and also include specifications with a p99 knot for top income groups.

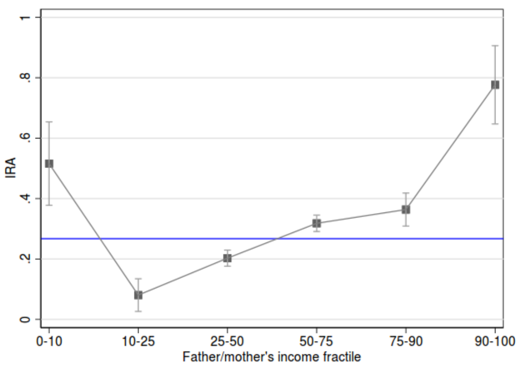
[25] See Figures A.7 and Table A.14). The coefficient at knot P99 is higher than the coefficient at P90, but estimates are imprecise.

distribution.

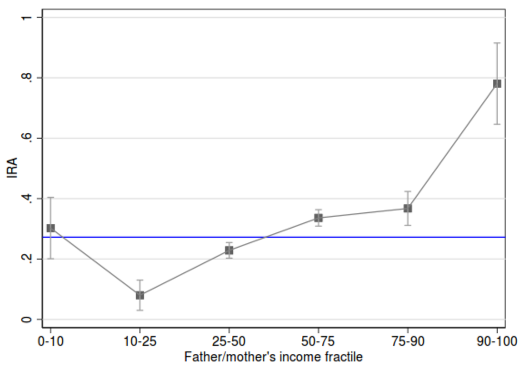
The previous specification includes a unique knot at P25 which could conceal relevant heterogeneity in the intergenerational transmission at the bottom of the distribution. In Figure 5 we introduce an additional knot at P10. Due to the sample size at the bottom of the distribution, we only implemented this specification with the extended and universal samples. This exercises confirm that transmission is heterogeneous at the bottom of the distribution, with a rank-rank slope larger in the first decile. This unveils that rank-rank slopes are not monotonically increasing but are rather J-shaped.

Figure 5: Intergenerational Income Ranking Associations: Non-linear spline regression estimates for income ranks (children aged 30-39). New knot at P10, and old knots at P25, P50, P75, P90

(a) Extended Sample



(b) Universal Sample



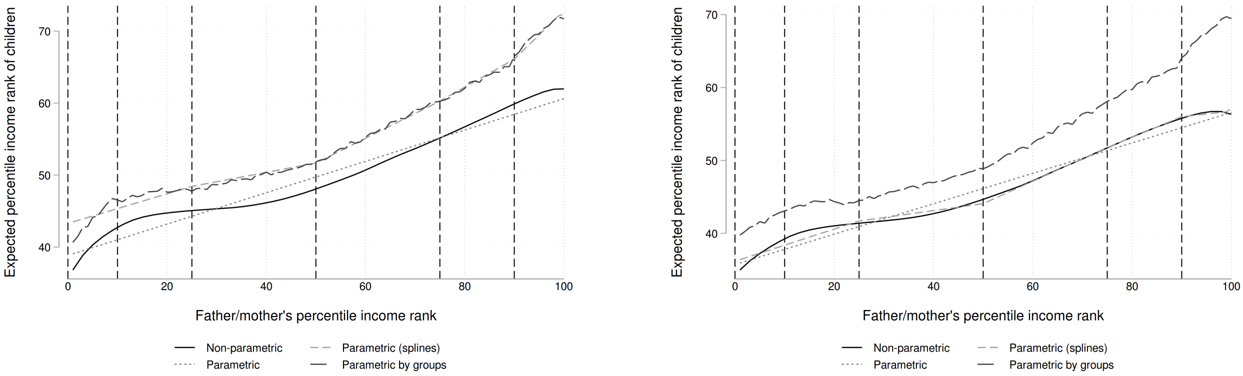
Next we report the estimates of a more flexible parametric strategy, which fits a separate standard IRA regression model to the six groups delimited by the knots of parental income distribution we used in the splines regressions (Figure A.8 in the Appendix shows the estimates of slopes in panels (a) and (c) and intercepts in panels (b) and (d)).[26] This strategy allows different slopes and intercepts for each fractile, and thus relaxes the splines model assumption of continuity in the relationship between parents' and children's income.[27] The IRA point estimates confirm the J-shaped relationship, while estimated intercepts display an inverse pattern, much more pronounced for the extended than for the universal sample. The zero intercept

[26]Table B9 present the coefficients.

[27]We do not present the result for the strict sample, because the sample size at the bottom end of the distribution is not large enough and the estimates we obtain are very imprecise.

Figure 6: Expected income rank of children conditional on parent's income rank

(a) Extended sample (b) Universal sample



of first regression, ant a estimated slope smaller than one for the lowest fractile suggest the existence of an intergenerational poverty trap. No child whose parent belongs to the first decile of the parental income distribution is expected to overcome the first decile.[28]

Figure 6 plots the expected rank of children conditional on parent's rank that result from the linear model and from three flexible models, namely (i) the parametric splines model, (ii) the parametric model with multiple intercepts, (iii) a non-parametric kernel-weighted polynomial regression.[29] The results of the various specifications consistently show a non-linearly relationship between income of both generations, consistent with the findings by Nybom and Stuhler (2017) for Sweden, but is at odds with the lineal relationship for the US Chetty et al. (2014a).[30] The large slop on both extreme of the distribution confirm the higher persistence levels for lower and top income groups.

The joint analysis of these results confirm a differential pattern on intergenerational mobility associated with parents' ranking. We verify a relatively high intergenerational mobility for those children from the bottom of the income distribution. The figures also shows a very steep slope for the higher percentiles, suggesting parental ranking has a much stronger marginal effect on the expected incomes of their kids compared with the rest of the parents. Given the higher

[28]The intercepts associated to the superior fractiles cannot be interpreted alone. The changes on these intercepts are associated with the mechanical relationship with the slopes imposed by the lineal regression model.

[29]The prediction of the splines model is based on the estimated coefficients shown in Table B9, while the prediction of the linear model is based on the estimated coefficients of Table 4.

[30]Figure A.9 and Figure A.10 replicate the estimates for earnings confirming the shape of the relationship.

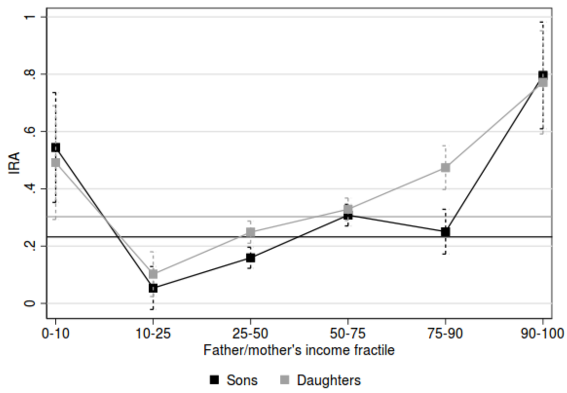
degree of intergenerational transmission in both tails of the distribution, Figure A.6 show the transition probabilities of children with parents at both ends of the income distribution. The transition probabilities for the bottom 10 percent of parental income provide evidence in support of the poverty traps we found in the regression analysis above, especially in sub-figure (f) when we use the universal sample. At the top of the parental distribution we observe the opposite picture (see sub-figures (a) to (c)). Children from high-income parents are far more likely to end up at the top decile and this pattern rises within the top 10 of parents income distribution.

Gender differences Next, in Figures 7 and A.11 we explore the presence of non-linearities by gender of children. The hypothesis of constant IRA were rejected both for sons and daughters. In the case of the women, the magnitude of the coefficients suggests a J-shaped curve, with an increasing slope from the knot P10-P25 to P100. In the case of the sons the coefficients suggests a less pronounced curve. Even, there are significant differences between the coefficients by gender in middle income (knots P25 and P75). We also explore the non-linearities by gender based in the prediction made by the more flexible models (See Figure B1). For both sons and daughters the results confirm the differential pattern on intergenerational mobility associated with parents' position in the income distribution. Given that in the case of the mothers' income we found significant differences in the average IRA between sons and daughters, we explore whether these differences are driven by the non-linearities. In this case, the results reject differences in the shape of the relationship by offspring gender (see Figure B2 in the online appendix).

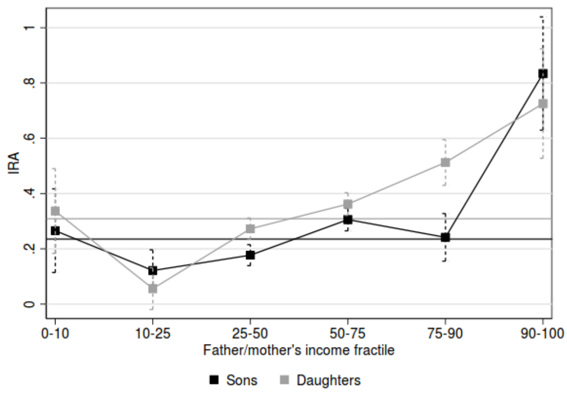
The presence of non-linearities: robustness analyses: We conduct several robustness checks of the previous non-linear results.[31] Firstly, we modify the reference used for the construction of the percentiles (Figure B3 and Table B10 in the online Appendix). Estimates using the sample as reference confirm the larger levels of transmission of income between the P75 and the P100 percentiles and in the bottom of the distribution. Secondly, as was done for the average estimates, we replicate the splines regression estimates using the 8 years of tax records for the construction of the samples (see Table A.3 for details). As the Figures B8 and the Table B11 in the online Appendix show, the estimates confirm previous results.

[31] Because the commented results remain unchanged, we focus on the specification that includes 5 knots: P10, P25, P50, P75 and P90 for extended and universal. All the results for strict sample and for other specifications are available on request.

Figure 7: Intergenerational Ranking Associations by gender: Non linear regression's estimates for income (sons aged 30-39). Knots at P10, P25, P50, P75, P90



(a) Extended



(b) Universal

Thirdly, to test sensitivity of our results to the criteria used to define permanent income, we carried out three strategies: excluded the zeros in the construction of the permanent income concept (Figure B4 in the online Appendix), (ii) include dummies to control for the number of years in which parents declare positive annual earning (see Figure B5 on the online Appendix). and (iii) we construct a new ranking based on real position on the income distribution (see Figure B6 in the online Appendix). Previous results are confirmed for estimates based on the extended sample, although persistence levels tend to be somewhat higher in magnitude. In particular we confirm a significantly and larger coefficient for the higher knots (the coefficients are also available in Table B12). Finally we explore if these results are sensitive to the inclusion of individuals without formal incomes. We sequentially include in the universal sample parents and children with zero permanent income. Te results confirm the convex shape and the larger magnitude of the first slope (from P0 to p10).[32]

5 Final comments

Using administrative data this paper provides precise evidence for Uruguay about the degree of intergenerational income mobility. Our estimates for the average intergenerational mobil-

[32]The difference between the coefficients at P0-P10 and P10-P25 is statistically significant in both cases. See Figure B7 and Table B11 in the online appendix.

ity situates Uruguay as an intermediate case between previous countries' estimates based on administrative records. However, this comparisons should be made with caution, because the differences in the empirical strategies used to measure intergenerational mobility, in the shape of income inequality between countries and the substantial heterogeneity hidden in the average measures of intergenerational mobility.

Most of the previous evidence on intergenerational mobility based on administrative records is for developed countries, where the institutions of the labour market are different and the informal employment has a marginal role. The large presence of informal labour markets in our context, implies a challenge for our estimations. This phenomenon, usually ignored in the developed countries literature, could have consequences on the estimations for the rest of the developing world. The availability of administrative-records data opens new opportunities for the study of intergenerational income mobility in less developed countries, but it also imposes new challenges since the informal sector explains a large part of individuals earning (income). The informal sector trends to increase the effect of non-filers problem and increases the number of individuals with intermittent participation in the formal sector.

We address this important and less explored question and we make a first attempt to advance in the estimation of intergenerational income transmission incorporating part of these particular challenges. We implement a set of strategies to mitigate the consequences of the presence of individuals with less attachment to the formal labour market. Our results suggest that the degree of intergenerational persistence is significantly higher when our measures consider families with less attachment to the formal labour market. Our estimates provide suggestive evidence about the role of the transmission of labour market status suggesting that the informal sector could represent an additional mechanism to explain inequality persistence.

In the paper we also highlight two additional results. First of all, the average mobility hides significant heterogeneity through parental income distribution. On the one hand, our results confirm the existence of important non-linearities in income persistence between generations, being particularly high for top income groups. For this group, the parents ranking improvements are almost completely transmitted to their son. On the other hand, our estimates also confirm a higher persistence in the other extreme of the parents income distribution.

The greater persistence at the bottom of the distribution suggests the existence of an intergenerational low-income trap. These results provides new issues to understand the different the mechanisms of inequality persistence. It is expected that these differences in the tails are explained by different mechanisms and also have different well-fare implications for the public policies. The higher persistence at the top could be related with the transmission of social capital, the inheritance of firms, capital and employers, the human capital investment and the integenerational transmission of preferences and abilities. While the persistence at the bottom could be related with credit constraints, the inheritance of long-term joblessness, poor human capital investment, the transmission of low social capital and problem to access to formal labour market.

Secondly, unlike previous evidence for US, we found a strong persistence between mothers and daughters (much greater than the rest of the father/son pairs) which may be indicating a transmission of gender roles. These results reveal the existence of some inequalities that seem to be transmitted from generation to generation.

Finally, the estimates of this study likely establish a lower bound of intergenerational earnings and income mobility in Uruguay. Our results suggest that the transmission of position is relatively weak in the middle of the distribution, but the persistence is strong at both tails of the distribution. These results reveal the existence of some inequalities that seem to be transmitted from generation to generation. In this sense, our finding about long term inequality persistence are a relevant benchmark if an important goal for public policy is to promote equality of opportunity. In addition, the finding of this paper may support a future avenue of investigation with respect to the potential contribution of informal earning to the intergenerational persistence of income inequality. Also opens new research question regarding the mechanisms that explain the greater intensity of transmission at both extreme of the distribution and the different persistence by gender.

A Appendix

Table A.1: Summary Statistics: Number of observations and participation of females by cohort (Sons, 2012)

Cohort	Sample	HH Survey	% Survey	Sample w/inc.	Tax w/inc.	% Tax	Female share			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	Sample	Sample w/inc.	HH Survey	Tax w/inc.
1970	13638	42039	32.4%	7165	25706	27.9%	51.1%	48.8%	52.8%	47.0%
1971	14798	39596	37.4%	8051	26653	30.2%	50.0%	47.9%	51.7%	46.9%
1972	15614	43335	36.0%	8610	26684	32.3%	50.6%	49.4%	55.1%	47.6%
1973	16499	41161	40.1%	9158	27329	33.5%	50.0%	48.9%	50.7%	47.2%
1974	17981	45859	39.2%	10159	28970	35.1%	49.8%	48.7%	51.7%	46.8%
1975	18969	45503	41.7%	10876	30973	35.1%	49.9%	48.2%	53.6%	46.5%
1976	19745	48286	40.9%	11574	32081	36.1%	49.9%	47.8%	54.1%	46.6%
1977	19886	47981	41.4%	11613	31976	36.3%	50.4%	48.8%	52.8%	46.9%
1978	19296	46016	41.9%	11586	32302	35.9%	50.3%	48.7%	50.1%	46.6%
1979	18982	46392	40.9%	11639	32198	36.1%	50.9%	49.3%	52.4%	46.6%
1980	17671	44599	39.6%	10960	31128	35.2%	50.9%	48.9%	53.3%	46.3%
1981	18580	41281	45.0%	11717	31789	36.9%	50.7%	49.4%	53.4%	46.1%
1982	26545	47364	56.0%	16771	32149	52.2%	50.9%	49.1%	49.3%	46.4%
1983	32596	42300	77.1%	20438	32023	63.8%	49.5%	46.9%	50.5%	45.7%
1984	33985	44844	75.8%	21487	32634	65.8%	49.6%	46.7%	52.2%	45.6%
1985	35721	44102	81.0%	22788	34010	67.0%	49.6%	46.8%	51.9%	45.6%
1986	36826	45269	81.3%	23417	34185	68.5%	49.4%	46.5%	51.4%	45.5%
1987	37310	45146	82.6%	23493	33364	70.4%	49.3%	46.2%	50.9%	45.1%
1988	40181	47727	84.2%	24798	34066	72.8%	49.6%	46.3%	50.5%	45.5%
1989	40842	47396	86.2%	24512	32771	74.8%	50.0%	46.2%	50.1%	45.3%
1990	45142	51396	87.8%	26234	32053	81.8%	49.3%	45.1%	49.8%	44.7%
1991	47252	49767	94.9%	25785	29603	87.1%	49.2%	43.6%	50.7%	43.3%
1992	48297	49431	97.7%	24507	27478	89.2%	49.2%	42.2%	50.6%	41.9%
1993	51687	45902	112.6%	23231	25518	91.0%	49.2%	41.2%	50.0%	40.7%
1994	53638	50707	105.8%	15172	16358	92.7%	49.1%	37.8%	48.6%	37.5%
1995	55212	55054	100.3%	2433	2552	95.3%	49.1%	40.0%	47.3%	39.7%
1996	55961	54490	102.7%	1028	1067	96.3%	49.4%	29.9%	47.3%	30.1%

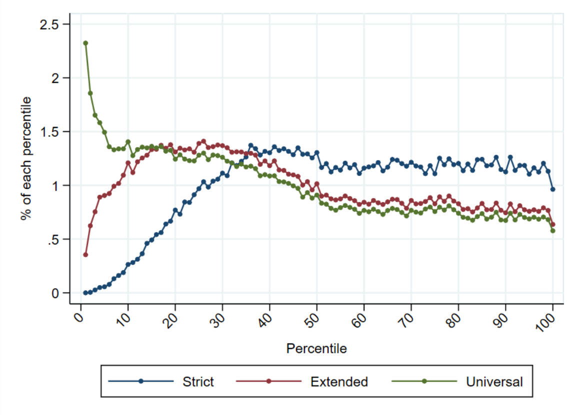
Notes: The table shows the number of observations according to the birth cohort of the sons generation on the different database: family linkages sample (column 2), household survey (column 3) and tax records (column 6). It also shows the fraction that the observations included in the sample represent with respect to the other databases, and the representativeness by gender. *Source:* Based on social security records (BPS), tax records (DGI), and Household Survey (INE).

Table A.2: Summary Statistics: Number of observations and participation of females by cohort (Parents, 2012)

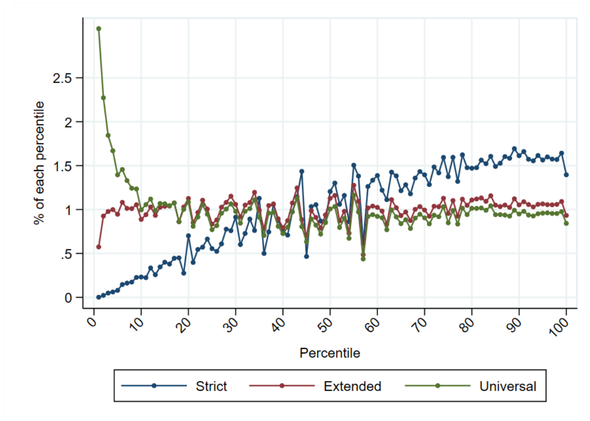
Cohort	Sample	HH Survey	% Survey	Sample w/inc.	Tax w/inc.	% Tax	Female share			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	Sample	Sample w/inc.	HH Surv.	Tax w/inc.
1944	9399	24072	39.0%	1211	3811	31.8%	66.7%	64.8%	56.5%	38.9%
1945	10549	26549	39.7%	1489	4427	33.6%	65.9%	62.2%	57.0%	38.2%
1946	11454	25346	45.2%	1909	5267	36.2%	65.8%	63.1%	55.4%	40.0%
1947	12540	29882	42.0%	2529	6336	39.9%	65.6%	61.6%	55.7%	40.7%
1948	13938	27754	50.2%	3027	7186	42.1%	64.1%	57.4%	54.7%	40.4%
1949	15048	30565	49.2%	3637	8094	44.9%	64.1%	57.7%	52.9%	41.6%
1950	16171	32466	49.8%	4582	9737	47.1%	63.3%	56.5%	54.7%	42.0%
1951	16949	27258	62.2%	5730	11496	49.8%	62.0%	56.0%	51.3%	43.0%
1952	18072	33317	54.2%	7630	14307	53.3%	60.3%	55.7%	55.4%	44.3%
1953	19016	30375	62.6%	8620	15601	55.3%	59.2%	55.0%	55.2%	44.9%
1954	21076	34604	60.9%	10104	17629	57.3%	57.4%	53.4%	52.3%	45.0%
1955	22790	35412	64.4%	11343	19390	58.5%	55.5%	51.0%	53.6%	44.8%
1956	24115	38545	62.6%	11952	20216	59.1%	55.4%	51.7%	54.0%	46.1%
1957	24615	35353	69.6%	12414	20541	60.4%	54.2%	50.8%	53.7%	46.6%
1958	25498	36551	69.8%	12874	21291	60.5%	52.7%	48.2%	54.0%	46.2%
1959	26426	36590	72.2%	13216	21785	60.7%	51.7%	47.7%	52.2%	46.8%
1960	27348	42167	64.9%	13666	22509	60.7%	51.5%	46.2%	51.8%	45.8%
1961	29374	34915	84.1%	14715	23503	62.6%	51.2%	46.5%	54.9%	46.6%
1962	31059	41655	74.6%	15316	24175	63.4%	51.7%	47.3%	55.4%	47.0%
1963	31682	39668	79.9%	15267	23920	63.8%	51.5%	47.5%	50.5%	46.7%
1964	32274	39669	81.4%	15370	24009	64.0%	52.0%	47.8%	51.2%	47.4%
1965	31860	39921	79.8%	14897	23410	63.6%	52.5%	48.7%	54.6%	47.3%
1966	31482	38605	81.5%	14522	22988	63.2%	52.3%	48.3%	52.5%	47.4%
1967	32033	39635	80.8%	14276	23372	61.1%	52.6%	48.0%	53.2%	47.1%
1968	32874	36670	89.6%	14384	23953	60.1%	52.4%	49.0%	53.7%	47.0%
1969	35865	40589	88.4%	15194	25985	58.5%	53.0%	48.5%	52.8%	47.2%

Notes: The table shows the number of observations according to the birth cohort of the parents generation on the different database: family linkages sample (column 2), household survey (column 3) and tax records (column 6). It also shows the fraction that the observations included in the sample represent with respect to the other databases, and the representativeness by gender.
Source: Based on social security records (BPS), tax records (DGI), and Household Survey (INE).

Figure A.1: Fraction of individuals in each percentile by sample (Sons aged 20-29)



(a) Formal income distribution



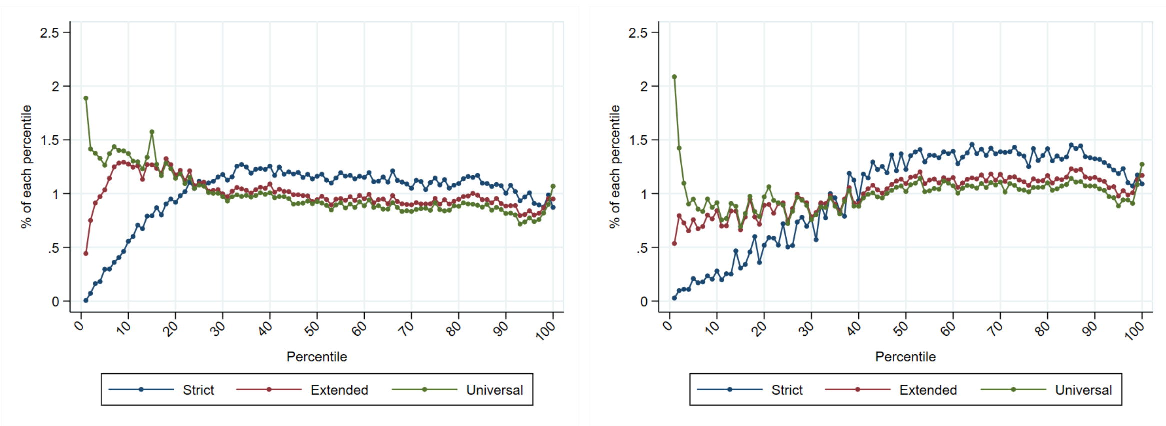
(b) Formal and informal income distribution

Notes: The figure shows the percentage of observations included in the sample of parents links database by percentile of labor income in tax records. Panel (a) and (b) shows this fraction for the sons generations (by age group), and panel (c) for parents. *Source:* Based on social security records (BPS) and tax records (DGI).

Table A.3: Criteria for the construction of the samples and permanent income

Alternatives samples for the robustness analysis				
Sample (labels)	Income condition	Offspring' Ages	Parents' Ages	N
Strict X8	5 positive earnings in 8 years (period 2009-2016)	30 - 39	40 - 65	49,205
Extended X8	2+ positive earnings in 8 years (period 2009-2016)	30 - 39	40 - 65	83,759
Universal X8	1+ positive earning in 8 years (period 2009-2016)	30 - 39	40 - 65	158,424
Universal with zeros (I)	Universal+ parents with zero income	30 - 39	40 - 65	149,943
Universal with zeros (II)	Universal+ parents and children with zeros	30 - 39	40 - 65	191,194
Alternatives definition of permanent income for the robustness analysis				
Sample (labels)	Income condition	Offspring' Ages	Parents' Ages	Definition of permanent income
Extended X5	2+ positive earnings	30 - 39	40 - 65	average only in the years with positive income records
Universal X5	1+ positive earning	30 - 39	40 - 65	average only in the years with positive income records
Extended X5	2+ positive earnings	30 - 39	40 - 65	Real position in the global distribution
Universal X5	1+ positive earning	20 - 39	40 - 65	Real position in the global distribution

Figure A.2: Fraction of individuals in each percentile by sample (Parents aged 45-65)



(a) Formal income distribution (b) Formal and informal income distribution

Notes: The figure shows the percentage of observations included in the sample of parents links database by percentile of labor income in tax records. Panel (a) and (b) shows this fraction for the sons generations (by age group), and panel (c) for parents. *Source:* Based on social security records (BPS) and tax records (DGI).

Figure A.3: Construction of samples and reference distribution

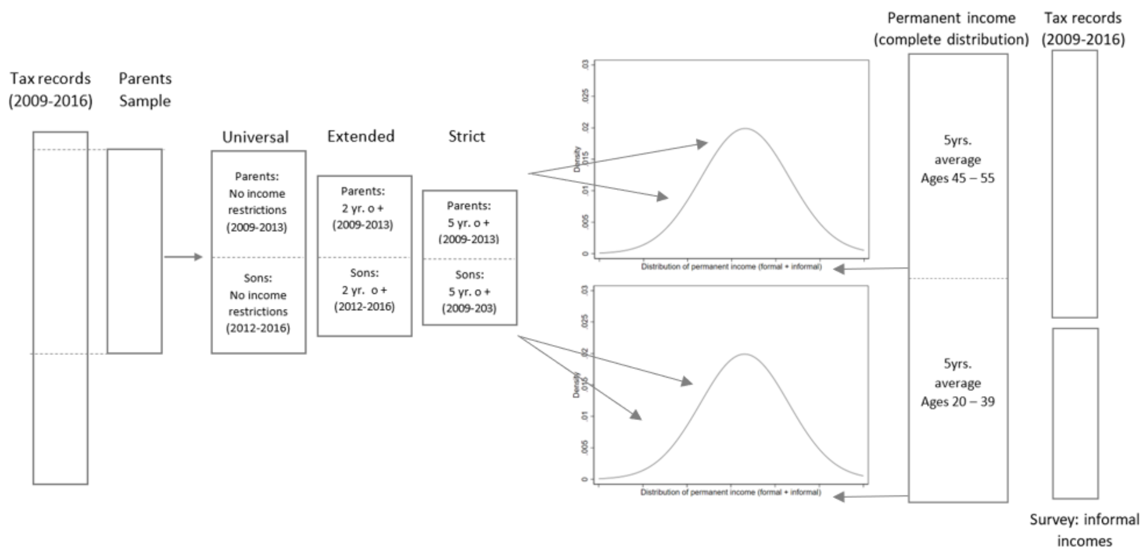
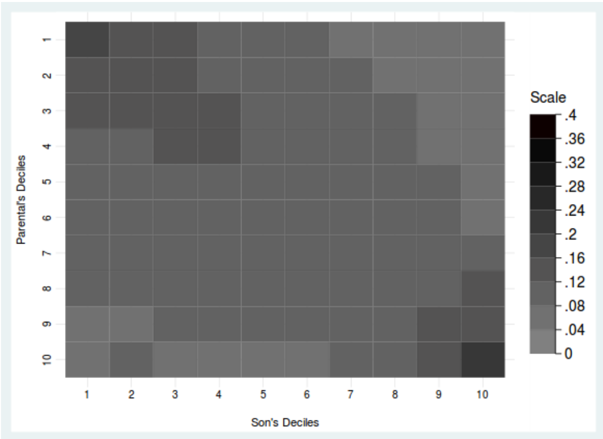


Figure A.4: Transition matrix - Extended Sample - sons aged 30- 39 (Own sample and reference distribution)

(a) Own sample



(b) Reference distribution (formal/informal)

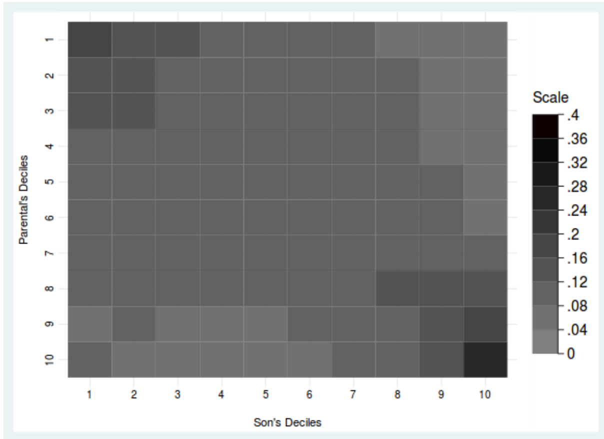
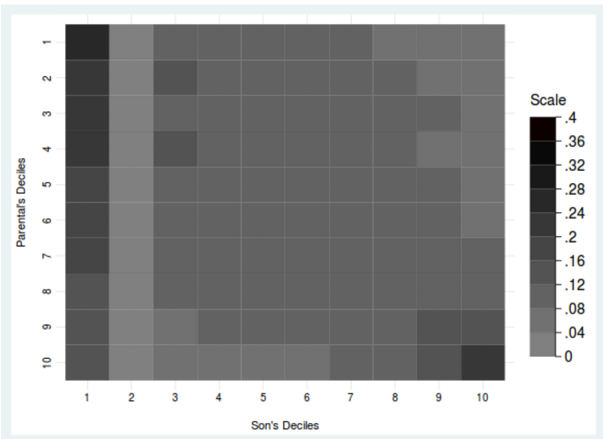


Figure A.5: Transition matrix - Universal Sample - sons aged 30- 39 (Own sample and reference distribution)

(a) Own sample



(b) Reference distribution (formal/informal)

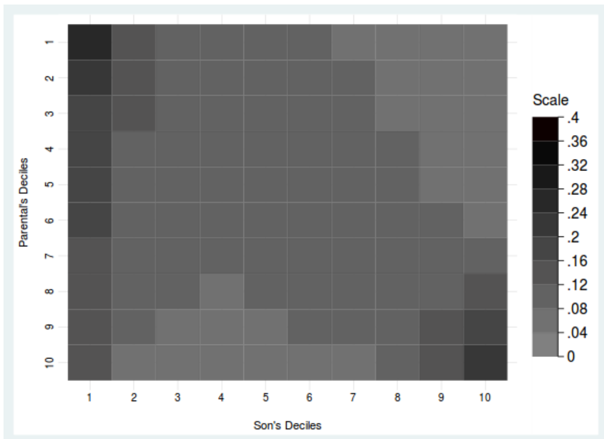


Table A.4: Intergenerational Ranking Association with alternative reference distribution). Non linear regression's estimates for income. Alternative samples

Knots	Stricted sample 8y			Extended Sample 8y			Universal sample 8y		
Average	0.261***			0.291***			0.269***		
	(0.003)			(0.003)			(0.002)		
0-25		0.043	0.043		0.181***	0.181***		0.147***	0.147***
		(0.033)	(0.033)		(0.020)	(0.020)		(0.017)	(0.017)
25-50		0.188***	0.189***		0.156***	0.156***		0.167***	0.167***
		(0.015)	(0.015)		(0.013)	(0.013)		(0.013)	(0.013)
50-75		0.271***	0.270***		0.371***	0.370***		0.393***	0.392***
		(0.014)	(0.014)		(0.013)	(0.013)		(0.013)	(0.013)
75-90		0.375***	0.383***		0.397***	0.402***		0.400***	0.403***
		(0.025)	(0.026)		(0.026)	(0.026)		(0.026)	(0.027)
90-99		0.776***	0.708***		0.824***	0.783***		0.900***	0.872***
		(0.058)	(0.065)		(0.060)	(0.067)		(0.061)	(0.068)
90-99			3.029***			2.224**			1.874*
			(0.979)			(1.062)			(1.091)
R^2	78,597	78,597	78,597	127,786	127,786	127,786	249,185	144,175	144,175
N									

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is son's and daughter's income percentiles. Controls: children's age, parent's age and sex

Table A.5: Number of years with positive formal income (sons and parents)

		N° years - Sons								
		0	1	2	3	4	5	6	7	8
N° years - Parents	0	24.08	6.05	6.53	6.98	7.93	8.03	8.90	10.67	20.83
	1	16.23	6.20	7.09	7.72	8.95	8.92	9.90	11.96	23.04
	2	15.30	6.17	6.94	8.03	9.51	9.31	9.76	12.10	22.88
	3	14.88	6.22	7.11	8.24	9.12	9.64	10.35	11.95	22.49
	4	14.41	6.30	7.19	8.49	10.38	10.01	10.63	12.22	20.38
	5	10.90	4.75	5.66	6.92	8.31	11.45	12.10	14.03	25.88
	6	11.87	5.51	6.63	7.67	9.46	10.51	11.44	13.20	23.71
	7	11.59	5.24	6.42	7.81	9.47	10.86	11.39	13.58	23.65
	8	11.03	5.45	6.79	7.96	9.67	10.93	11.24	13.00	23.92

Table A.6: Average IRA for total income by sample. Children aged 20-29

	Samples		
	Strict	Extended	Universal
PANEL A: Global distribution			
IRA	0.134***	0.152***	0.152***
N	52,621	168,702	202,654
PANEL B: Sample distribution			
IRA	0.157***	0.143***	0.142***
N	52,621	168,702	202,654
PANEL C: Mean test difference's			
vs. Strict sample	-	19.69	18.67
Chi (p-value)	-	(0.0000)	(0.0000)
vs. Strict extend	-	-	0.01
Chi (p-value)	-	-	(0.9155)
Global distribution vs sample	56.40	22.42	22.83
Chi (p-value)	(0.008)	(0.000)	(0.000)

Note: The dependent variable is offspring's total income. Coefficients are OLS estimates. Controls: children's age, parental age and sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table A.7: Average IRA for earnings by sample and age group

	PANEL A: Average IRA by samples					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IRA	0.131***	0.223***	0.145***	0.245***	0.143***	0.245***
N	52,621	30,193	168,702	92,519	202,654	98,977
	PANEL B: Mean test (difference's between reference distribution and samples)					
Diff. vs strict			11.84	14.73	8.17	14.44
F (p-value)			(0.0006)	(0.0001)	(0.0043)	(0.0000)
Diff. vs extended					2.36	0.05
F (p-value)					(0.1243)	(0.8207)
Diff. vs own distribution	56.23	117.66				
F (p-value)	(0.0000)	(0.0134)				

Note: The dependent variable is offspring's earning. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table A.8: Differences between Intergenerational Ranking Association Test: income vs earning

	IRA coefficients Test, income vs earning					
	All	Daugthers	Sons	All	Daugthers	Sons
			Strict			
Diff. income vs earning	9.5215	1.6782	10.0596	30.4218	20.5166	11.1184
F(P-value)	0.0020	0.1952	0.0015	0.0000	0.0000	0.0009
			Extended			
Diff. income vs earning	105.0647	38.7037	72.0934	206.8610	98.6294	106.6500
F(P-value)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
			Universal			
Diff. income vs earning	164.8213	61.4971	109.4385	275.9955	141.8292	133.5508
F(-value)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Based on coefficients estimated from eq. 1 and presented in Tables 4 A.6, 5, A.9 and A.7.

Table A.9: Average IRA for earnings by gender and sample (30-39 age group)

	Samples					
	Strict		Extended		Universal	
	Daughters	Sons	Daughters	Sons	Daughters	Sons
IRA - income	0.136***	0.186***	0.155***	0.208***	0.156***	0.209***
	(23.17)	(23.54)	(48.38)	(42.11)	(54.36)	(45.38)
N	23667	15162	78913	42124	96433	50328
Diff. between genders	23.71		86.64		113.68	
F (p-value)	(0.0000)		(0.0000)		(0.0000)	

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table A.10: Average IRA of total income by sub-sample of parents and gender (without distinguishing maximum income within the household)

	PANEL A: Fathers								
	Strict			Extended			Universal		
	Average	Both	Only father	Average	Both	Only father	Average	Both	Only father
IRA - fathers' income	0.255***	0.310***	0.230***	0.268***	0.306***	0.236***	0.265***	0.269***	0.246***
	(0.012)	(0.026)	(0.014)	(0.007)	(0.012)	(0.008)	(0.006)	(0.007)	(0.011)
N	8,683	2,180	6,503	25,626	9,845	15,781	36,564	27,476	9,088
	PANEL B: Mothers								
	Strict			Extended			Universal		
	Average	Both	Only father	Average	Both	Only father	Average	Both	Only father
IRA - mothers' income	0.217***	0.242***	0.216***	0.259***	0.310***	0.255***	0.267***	0.292***	0.262***
	(0.006)	(0.022)	(0.007)	(0.004)	(0.011)	(0.004)	(0.004)	(0.008)	(0.004)
N	23,690	2,180	21,510	66,738	9,845	56,893	89,889	27,476	62,413
F Test Father vs mother	8.037	0.117	14.440	1.451	8.848	22.468	0.073	26.867	22.612
(p-value)	0.005	0.733	0.000	0.228	0.003	0.000	0.787	0.000	0.000

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table A.11: Average IRA of total income by parents' gender (without distinguishing maximum income within the household)

	PANEL A: Fathers								
	Strict			Extended			Universal		
	Average	Daughter	Son	Average	Daughter	Son	Average	Daughter	Son
IRA - income	0.215***	0.208***	0.221***	0.195***	0.190***	0.199***	0.176***	0.168***	0.182***
	(0.011)	(0.017)	(0.016)	(0.005)	(0.008)	(0.007)	(0.005)	(0.007)	(0.006)
N	10,152	4,607	5,545	38,223	17,248	20,975	48,130	22,251	25,879
Diff. between genders		0.008			0.0679			0.0002	
Chi (p-value)		(0.927)			(0.79)			(0.987)	
	PANEL B: Mothers								
	Strict			Extended			Universal		
	Average	Daughters	Sons	Average	Daughters	Sons	Average	Daughters	Sons
IRA - income	0.218***	0.255***	0.182***	0.248***	0.294***	0.204***	0.246***	0.295***	0.198***
	(0.006)	(0.009)	(0.009)	(0.004)	(0.005)	(0.005)	(0.003)	(0.005)	(0.005)
N	24,534	12,323	12,211	75,144	36,604	38,540	93,267	45,508	47,759
Diff. between genders		26.508			110.504			154.902	
Chi (p-value)		(0.000)			(0.000)			(0.000)	

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table A.12: Gender IRA of total income by sub-sample of parents (sample of households with two parents)

	PANEL A: Fathers (subsample with both parents)								
	Strict			Extended			Universal		
	Average	Daughter	Son	Average	Daughter	Son	Average	Daughter	Son
IRA - income	0.310***	0.288***	0.354***	0.306***	0.305***	0.327***	0.266***	0.265***	0.283***
	(0.026)	(0.038)	(0.036)	(0.012)	(0.017)	(0.018)	(0.007)	(0.011)	(0.010)
N	2,180	1,071	1,109	9,845	5,027	4,818	33,516	16,747	16,769
Diff. between genders		(1.6156)			(0.7960)			(1.5594)	
F (p-value)		0.2038			0.3723			0.2117	
	PANEL B: Mothers (subsample with both parents)								
	Strict			Extended			Universal		
	Average	Daughter	Son	Average	Daughter	Son	Average	Daughter	Son
IRA - income	0.242***	0.288***	0.222***	0.310***	0.279***	0.347***	0.297***	0.244***	0.341***
	(0.022)	(0.032)	(0.032)	(0.011)	(0.016)	(0.016)	(0.008)	(0.011)	(0.011)
N	2,180	1,071	1,109	9,845	5,027	4,818	33,516	16,747	16,769
Diff. between genders		0.1433			2.1431			8.6249	
F (p-value)		(0.1433)			(0.003)			(0.000)	

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

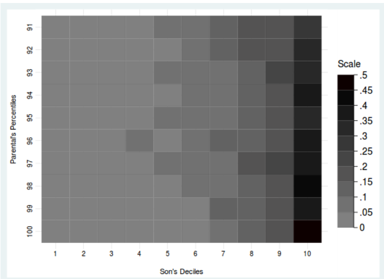
Table A.13: Average IRA for total income by gender of parents and children - households with sons and daughters (30-39 age group)

	PANEL B: Fathers								
		Strict			Extended			Universal	
	Average	Daugthers	Sons	Average	Daugthers	Sons	Average	Daugthers	Sons
IRA - income	0.261***	0.289***	0.229***	0.293***	0.290***	0.296***	0.287***	0.285***	0.289***
	(0.023)	(0.033)	(0.031)	(0.010)	(0.015)	(0.015)	(0.009)	(0.012)	(0.012)
Observations	2,477	1,294	1,183	11,502	5,834	5,668	18,062	9,101	8,961
F Test Diff. between genders		1.77			0.10			0.58	
(p-value)		(0.1835)			(0.7558)			(0.4478)	
	PANEL B: Mothers								
		Strict			Extended			Universal	
	Average	Daugthers	Sons	Average	Daugthers	Sons	Average	Daugthers	Sons
IRA - income	0.259***	0.316***	0.203***	0.297***	0.337***	0.257***	0.302***	0.344***	0.258***
	(0.012)	(0.017)	(0.016)	(0.006)	(0.009)	(0.009)	(0.005)	(0.007)	(0.008)
Observations	7,033	3,553	3,480	28,958	14,573	14,385	41,838	21,021	20,817
F Test Diff. between genders		22.54			43.11			63.39	
(p-value)		(0.0000)			(0.0000)			(0.0000)	

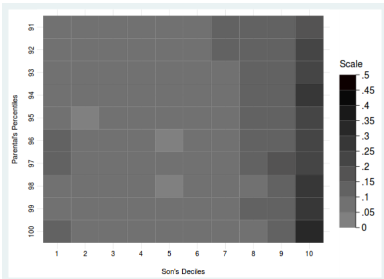
Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Figure A.6: Transition matrices. Deciles of children with parents at top and bottom 10 percentiles (age group 30-39)

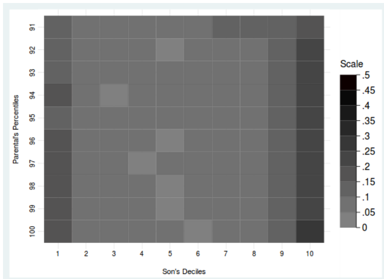
Top 10%



(a) Strict

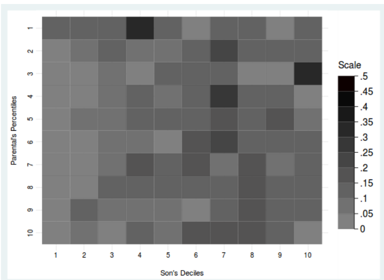


(b) Extended

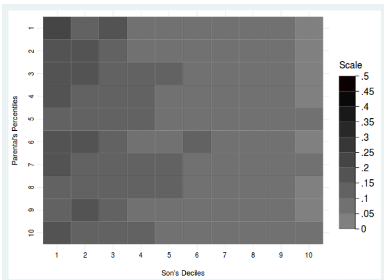


(c) Universal

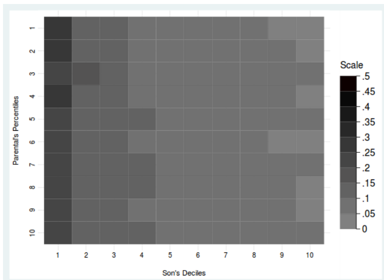
Bottom 10%



(d) Strict



(e) Extended



(f) Universal

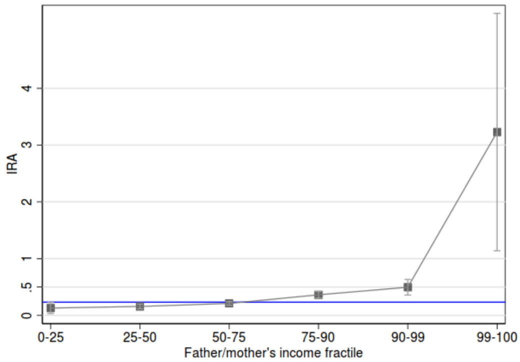
Table A.14: Intergenerational Ranking Association. Non linear regression's estimates for income. Strict, extended and universal samples

Knots	Strict		Extended		Universal	
0-25	0.134**	0.133**	0.182***	0.182***	0.147***	0.147***
	(0.066)	(0.066)	(0.023)	(0.023)	(0.019)	(0.019)
25-50	0.154***	0.155***	0.182***	0.182***	0.215***	0.215***
	(0.028)	(0.028)	(0.017)	(0.017)	(0.016)	(0.016)
50-75	0.211***	0.210***	0.325***	0.324***	0.338***	0.338***
	(0.022)	(0.022)	(0.018)	(0.018)	(0.018)	(0.018)
75-90	0.357***	0.368***	0.358***	0.360***	0.366***	0.368***
	(0.037)	(0.038)	(0.036)	(0.036)	(0.037)	(0.037)
90-99	0.584***	0.498***	0.785***	0.765***	0.786***	0.769***
	(0.080)	(0.091)	(0.083)	(0.093)	(0.085)	(0.096)
90-99		3.458***		1.394		1.329
		(1.313)		(1.413)		(1.446)
N	30,193	30,193	82,519	82,519	98,977	98,977

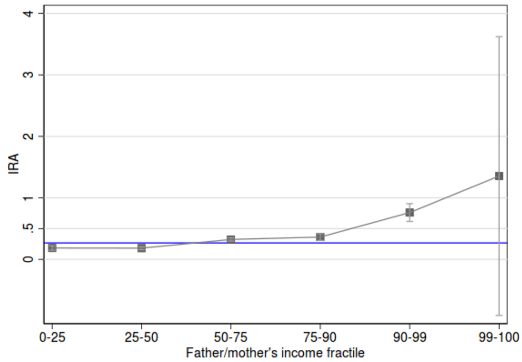
Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort combining the universe of tax records and the Household Survey. Child and parents income is defined as the average of 5 yearly incomes but exclude zeros from the averages. Coefficients are OLS estimates. Controls: children's age, parent's age and sex

Figure A.7: Intergenerational Income Ranking Associations: Non linear regression's estimates for income. Knots at P25, P50, P75, P90 and p99 (sons aged 30-39)

(a) Strict Sample



(b) Extended Sample



(c) Universal Sample

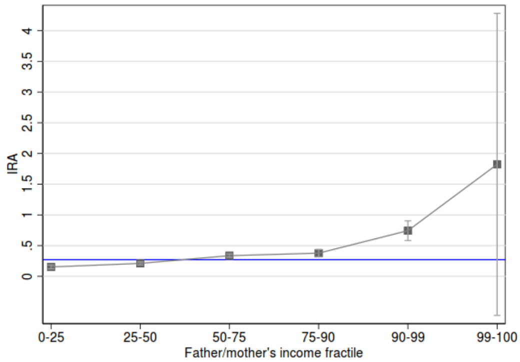
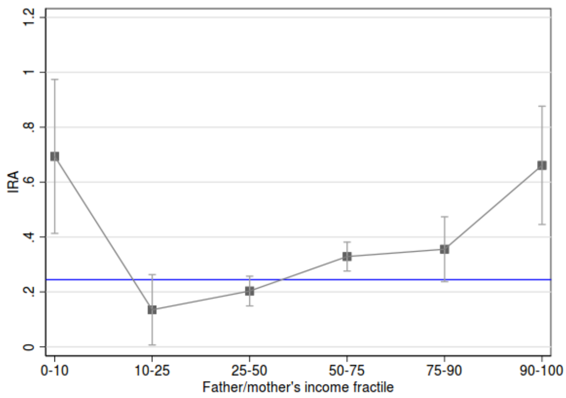
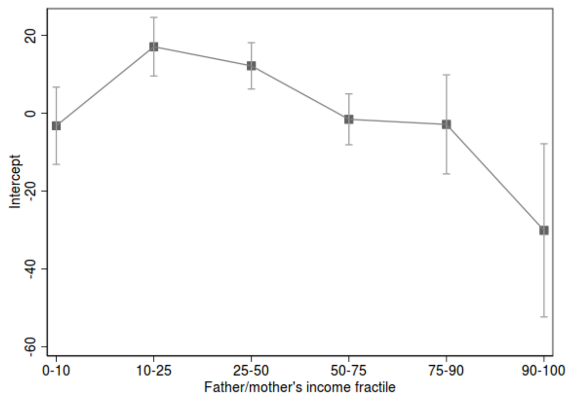


Figure A.8: IRA: Non linear regression's for income groups (P0-P10, P10-P25, P25-P50, P50-P75, P75-P90 and P90-100, sons aged 30-39)

PANEL A: Extended sample

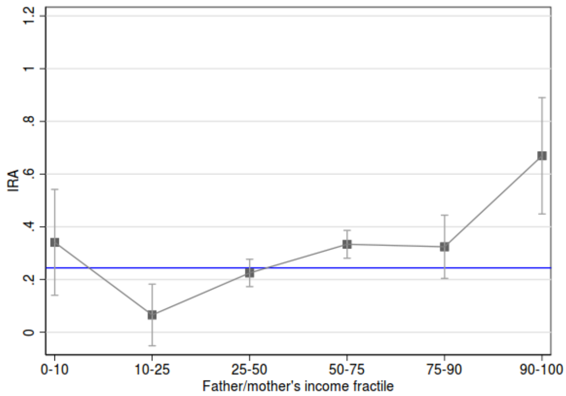


(a) IRA coefficients

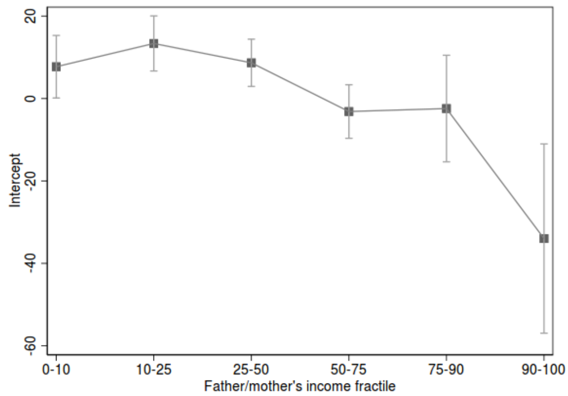


(b) Intercepts

PANEL B: Universal sample



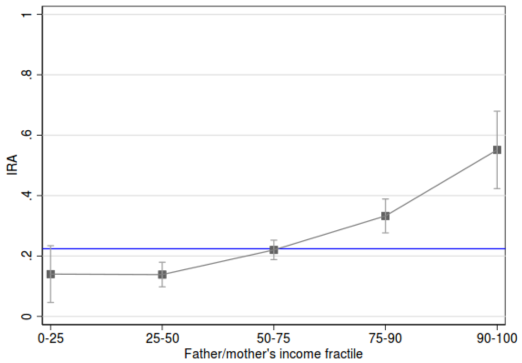
(c) IRA coefficients



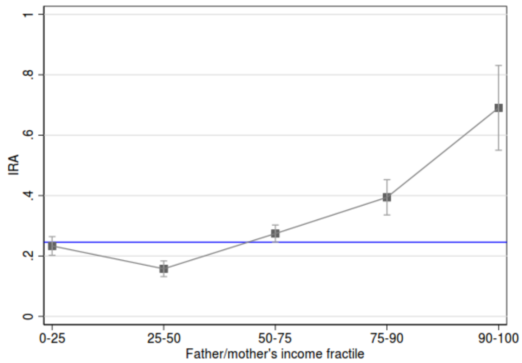
(d) Intercepts

Figure A.9: Intergenerational Income Ranking Associations: Non linear regression's estimates for earning. Knots at P25, P50, P75, P90 and p99

(a) Main Sample (sons aged 30-39)



(b) Extended Sample (sons aged 30-39)



(c) Universal Sample (sons aged 30-39)

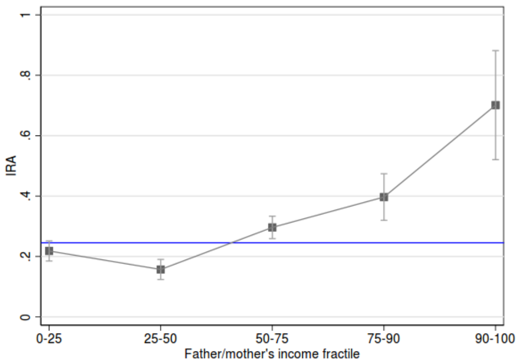
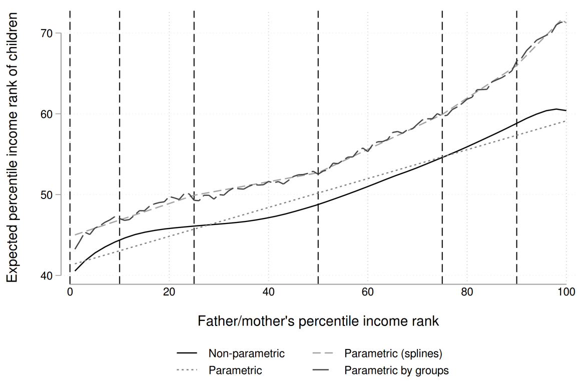


Figure A.10: Expected earning rank of children conditional on parent's earning rank

(a) Extended sample



(b) Extended sample

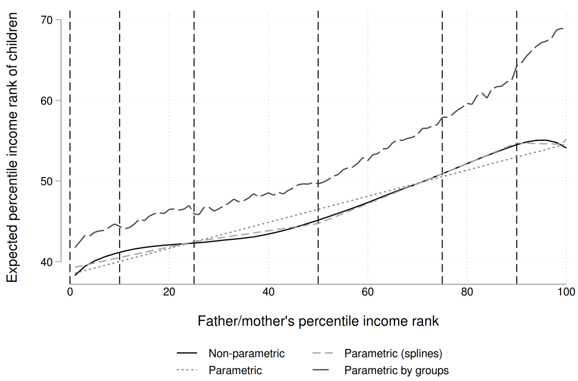
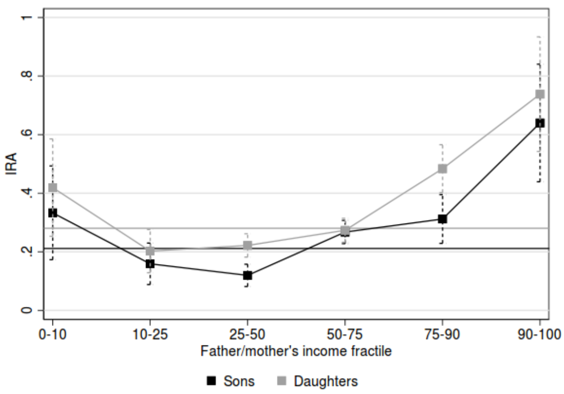
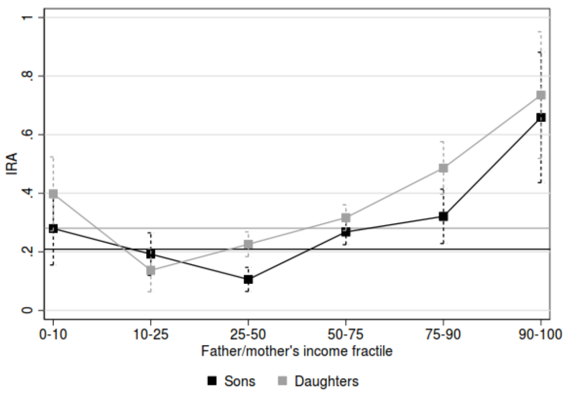


Figure A.11: Intergenerational Ranking Associations by gender: Non linear regression's estimates for earning (sons aged 30-39). Knots at P10, P25, P50, P75, P90



(a) Extended



(b) Universal

References

Aghion, P., Banerjee, A., and Piketty, T. (1999). Dualism and macroeconomic volatility. *The Quarterly Journal of Economics*, 114(4):1359–1397.

Alvaredo, F. and Gasparini, L. (2015). Recent trends in inequality and poverty in developing countries. *Handbook of income distribution*, 2:697–805.

Araya, F. (2019). Evidencia sobre la movilidad intergeneracional de ingresos laborales para un país en desarrollo: el caso de uruguay. *El Trimestre Económico*, 86:265.

Atkinson, A. B. (2007). Measuring top incomes: methodological issues. *Top incomes over the twentieth century: A contrast between continental European and English-speaking countries*, 1:18–42.

Azevedo, V. M. R. and Bouillon, C. P. (2010). Intergenerational Social Mobility in Latin America: A review of existing evidence. *Revista de Analisis Economico – Economic Analysis Review*, 25(2):7–42.

Becker, G. S. and Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3, Part. 2):S1–S39.

Bernuy, V. S. J. and Esteve, A. (2019). Amores imposibles: la brecha entre universitarios y el resto de grupos educativos en los mercados matrimoniales de américa latina, 1970-2010. *Notas de Población*, 46(108):11–36.

Björklund, A., Jäntti, M., and Lindquist, M. J. (2009). Family background and income during the rise of the welfare state: Brother correlations in income for swedish men born 1932-1968. *Journal of Public Economics*, 93(5-6):671–680.

Björklund, A., R. J. and Waldenstöm, D. (2012). Intergenerational top income mobility in sweden: Capitalist dynasties in the land of equal opportunity?. 96:474–484.

Böhlmark, A. and Lindquist, M. J. (2006). Life-cycle variations in the association between current and lifetime income: Replication and extension for sweden. *Journal of Labor Economics*, 24(4):879–896.

Bourguignon, F., Ferreira, F. H., and Walton, M. (2007). Equity, efficiency and inequality traps: A research agenda. *Journal of Economic Inequality*.

Breen, R. and García-Peñalosa, C. (2005). Income inequality and macroeconomic volatility: an empirical investigation. *Review of Development Economics*, 9(3):380–398.

Burdín, G., De Rosa, M., Vigorito, A., and Vila, J. (2019). Is the recent inequality fall in latin america a datadriven illusion? top income shares and mobility patterns in uruguay 2009-2016. *Serie Documentos de Trabajo; 30/19.*

Burdín, G., De Rosa, M., Vigorito, A., and Vilá, J. (2022). Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data. *World Development*, 152:105783.

Carrasco, P. (2012). El efecto de las condiciones de ingreso al mercado de trabajo en los jóvenes uruguayos: Un análisis basado en la protección de la seguridad social. *Serie Documentos de Trabajo; 13/12.*

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014a). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014b). Is the united states still a land of opportunity? recent trends in intergenerational mobility. *American Economic Review*, 104(5):141–47.

Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014c). Is the us still a land of opportunity? recent trends in intergenerational mobility. *The American Economic Review*, 104.

Corak, M. and Heisz, A. (1999). The intergenerational earnings and income mobility of canadian men: Evidence from longitudinal income tax data. pages 504–533.

Corak, M. and Piraino, P. (2010). Intergenerational Earnings Mobility and the Inheritance of Employers. IZA Discussion Papers 4876, Institute of Labor Economics (IZA).

Dahl, M. and DeLeire, T. (2008). The association between children's earnings and fathers' lifetime earnings: Estimates using administrative data. *Discussion Paper No. 1342-08.*

Deutscher, N. and Mazumder, B. (2020). Intergenerational mobility across australia and the stability of regional estimates. *Labour Economics*, 66:101861.

Deutscher, N. and Mazumder, B. (2021). Measuring intergenerational income mobility: A synthesis of approaches. *Journal of Economic Literature.*

Dunn, C. E. (2007). The Intergenerational Transmission of Lifetime Earnings: Evidence from Brazil. *The B.E. Journal of Economic Analysis & Policy*, 7(2):1–42.

Durlauf, S. N. and Seshadri, A. (2018). Understanding the great gatsby curve. *NBER Macroeconomics Annual*, 32(1):333–393.

Espino, A., Isabella, F., Leites, M., and Machado, A. (2017). Do women have different labor supply behaviors? evidence based on educational groups in uruguay. *Feminist Economics*, 23(4):143–169.

Gandelman, N. and Robano, V. (2014). Intergenerational Mobility, Middle Sectors and Entrepreneurship in Uruguay. *Latin american journal of economics*, 51:195 – 226.

Grawe, N. (2004). *Intergenerational mobility for whom? The experience of high- and low-earning sons in international perspective.* Cambridge: Cambridge University Press.

Haider, S. and Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, 96(4):1308–1320.

Hertz, T. (2009). *Chapter Five. Rags, Riches, and Race*. Princeton University Press.

Jantti, M. and Jenkins, S. (2015). Income mobility. In *Handbook of Income Distribution. Vol. 2*.

Jiménez, M. (2011). Un análisis empírico de las no linealidades en la movilidad intergeneracional del ingreso. el caso de Argentina. *Documento de trabajo del CEDLAS No. 114*.

Jiménez, M. (2017). La movilidad intergeneracional del ingreso y sus métodos de estimación. un análisis comparativo para argentina y chile. *Cuadernos de Economía*, 41.

Leites, M., Salas, G., Rivero, L., Suarez, L., and Vigorito, A. (2018). *Trayectorias educativas y laborales de los jóvenes en Uruguay*. BID-Espacio Público-IDRC.

Leites, M., Sena, E., and Vila, J. (2020). *Movilidad intergeneracional de ingresos en Uruguay. Una mirada en base a registros administrativos*. Cuaderno sobre Desarrollo Humano N°12 del PNUD, Serie El Futuro en Foco.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *The Review of Economics and Statistics*, 87:235–255.

Mitnik, P., Bryant, V., Weber, M., and Grusky, D. B. (2015). New estimates of intergenerational mobility using administrative data. *Statistics of Income Division working paper, Internal Revenue Service*.

Munk, M., Bonke, J., and Hussain, M. (2016). Intergenerational top income persistence: Denmark half the size of sweden. *Economics Letters*, 140.

Nybom, M. and Stuhler, J. (2017). Biases in Standard Measures of Intergenerational Income Dependence. *Journal of Human Resources*, 52(3):800–825.

Núñez, E. and Miranda, L. (2007). Recent findings on intergenerational income and educational mobility in chile. *University of Chile, Department of Economics, Working Papers*.

Pastore, F., Doruk, , and Yavuz, H. (2019). Intergenerational mobility: An assessment for latin american countries. *IZA Discussion Papers 12312,*.

Piketty, T. (2000). Chapter 8 Theories of persistent inequality and intergenerational mobility.

Ramos, X. and Van de gaer, D. (2016). Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, 30(5):855–883.

Robeyns, I. (2019). What, if anything, is wrong with extreme wealth? *Journal of Human Development and Capabilities*, 20(3):251–266.

Roemer, J. E. and Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature*, 54(4):1288–1332.

Sanroman, G. (2010). Intergenerational educational mobility: evidence from three approaches for brazil, chile, uruguay and the usa (1995-2006). *Documento de trabajo N.º 01/10*.

Urraburu, J. (2019). Movilidad educativa y ocupacional intergeneracional en uruguay. *Tesis de maestría. Facultad de Ciencias Sociales, Universidad de la República, Montevideo.*

B Online Appendix: Additional Tables and Figures

Table B1: Educational assortative mating by year of formal education. Couples aged 25 and +

		Women					
	Yrs. education	0-6	6-9	9-12	12-16	> 16	Total
Men	0-6	**20%**	9%	4%	1%	1%	35%
	6-9	8%	**13%**	8%	2%	3%	33%
	9-12	2%	5%	**6%**	2%	3%	18%
	12-16	0%	1%	1%	**2%**	2%	6%
	> 16	0%	1%	1%	1%	**5%**	9%
	Total	30%	29%	19%	8%	13%	100%

Note: Frequencies based on National Household Survey.

Table B2: Earning assortative mating by Decile. Couples aged 25 and +

		Women									
	Decile	1	2	3	4	5	6	7	8	9	10
Men	1	**2,1%**	1,5%	1,2%	1,1%	0,8%	0,7%	0,7%	0,5%	0,4%	0,4%
	2	1,6%	**1,4%**	1,4%	1,0%	0,9%	0,6%	0,4%	0,3%	0,3%	0,2%
	3	1,4%	1,0%	**1,3%**	1,2%	1,5%	0,8%	0,7%	0,6%	0,5%	0,3%
	4	1,3%	0,8%	1,2%	**1,4%**	1,1%	1,5%	0,8%	0,7%	0,4%	0,4%
	5	1,2%	1,0%	1,1%	1,2%	**1,1%**	1,2%	1,2%	1,0%	0,7%	0,4%
	6	0,9%	1,1%	1,1%	1,1%	1,2%	**1,3%**	1,3%	1,1%	0,8%	0,4%
	7	1,2%	1,0%	1,1%	1,0%	1,0%	1,2%	**1,2%**	1,2%	1,1%	0,6%
	8	1,0%	0,6%	0,8%	0,9%	1,0%	1,3%	1,1%	**1,2%**	1,7%	1,0%
	9	0,9%	0,6%	0,7%	0,7%	0,7%	0,9%	1,3%	1,7%	**1,8%**	1,8%
	10	0,6%	0,4%	0,5%	0,4%	0,5%	0,7%	1,0%	1,3%	1,9%	**3,9%**

Note: Frequencies based on National Household Survey. Sample includes couples of women and men aged 25 years old an more.

Table B3: Average IRA for income by sex and age group (samples constructed with 8 years of tax records)

	By age group					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IRA	0.179***	0.260***	0.189***	0.288***	0.189***	0.289***
	(0.003)	(0.005)	(0.002)	(0.004)	(0.002)	(0.003)
	103,009	49,205	199,428	86,018	270,337	143,901
	By sex (30-39 age group)					
	Strict		Extended		Universal	
	Daughter	Son	Daughter	Son	Daughter	Son
IRA - income	0.293***	0.230***	0.332***	0.245***	0.340***	0.237***
	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
N	23,796	25,409	42,312	43,706	72,220	71,681

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B4: Average IRA for income by sex and age group controlling by participation in the labor market

	By age group					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IRA	0.139***	0.229***	0.142***	0.281***	0.141***	0.291***
	(0.004)	(0.006)	(0.003)	(0.004)	(0.003)	(0.004)
N	52,621	30,192	168,702	82,515	202,654	98,977
	By sex (30-39 age group)					
	Strict		Extended		Universal	
	Daughters	Son	Daughters	Son	Daughters	Son
IRA - income	0.253***	0.205***	0.318***	0.244***	0.332***	0.250***
	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)
N	15,031	15,161	40,395	42,120	48,649	50,328

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: parent's labor market participation children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B5: Average IRA for income by sex and age group. Alternative permanent income (based only on positive income)

	By age group			
	Extended		Universal	
	20-29	30-39	20-29	30-39
IRA	0.166***	0.259***	0.164***	0.264***
	(0.002)	(0.004)	(0.002)	(0.003)
N	168,915	83,014	202,981	99,991
	By sex (30-39 age group)			
	Extended		Universal	
	Daughters	Sons	Daughters	Sons
IRA	0.301***	0.218***	0.308***	0.221***
	(0.005)	(0.005)	(0.005)	(0.005)
N	40,641	42,373	49,164	50,827

Note: The dependent variable is an upper bound of the offspring's permanent income. The permanent income is based on the average of use the same 5 annual income, but it excludes the zeros in the calculation. We proceed in analogous manners in the case of the permanent income of the parents. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B6: Average Intergenerational Position Association (IPA) for income by sex and age group

	By age group					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IPA	0.088***	0.209***	0.108***	0.233***	0.110***	0.238***
	(0.004)	(0.006)	(0.002)	(0.004)	(0.002)	(0.004)
N	52,621	30,193	158,342	66,813	187,114	75,377
	By sex (30-39 age group)					
	Strict		Extended		Universal	
	Daughters	Sons	Daughters	Sons	Daughters	Sons
IPA	0.231***	0.187***	0.269***	0.199***	0.278***	0.200***
	(0.008)	(0.008)	(0.006)	(0.005)	(0.005)	(0.005)
N	15,031	15,162	32,607	34,206	36,953	38,424

Note: The dependent variable is offspring's real position. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B7: Average IGE for income by sex and age group

	By age group					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IGE	0.110***	0.206***	0.125***	0.223***	0.123***	0.208***
	(0.003)	(0.005)	(0.002)	(0.003)	(0.002)	(0.004)
	52,621	30,193	168,702	82,519	202,654	98,977
	By sex (30-39 age group)					
	Strict		Extended		Universal	
	Daughters	Sons	Daughters	Sons	Daughters	Sons
IGE	0.221***	0.192***	0.257***	0.190***	0.247***	0.171***
	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
N	15,031	15,162	40,395	42,124	48,649	50,328

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B8: Average IRA for income by sex and age group for alternative reference distribution (own sample)

	By age group					
	Strict		Extended		Universal	
	20-29	30-39	20-29	30-39	20-29	30-39
IRA	0.1576***	0.2649***	0.1434***	0.2692***	0.1421***	0.2680***
	(0.0041)	(0.0061)	(0.0023)	(0.0036)	(0.0021)	(0.0033)
	52,621	30,193	168,702	82,519	202,654	98,977
	By sex (30-39 age group)					
	Strict		Extended		Universal	
	Daughters	Sons	Daughters	Sons	Daughters	Sons
IRA	0.2784***	0.2516***	0.2998***	0.2391***	0.3016***	0.2350***
	(0.0087)	(0.0087)	(0.0051)	(0.0051)	(0.0046)	(0.0047)
N	15,031	15,162	40,395	42,124	48,649	50,328

Note: The dependent variable is offspring's total incomes. Coefficients are OLS estimates. Controls: children's age, father's age, and parent's sex. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% level, respectively.

Table B9: IRA: Non linear regression's of income groups: P0-P10, P10-P25, P25-P50, P50-P75, P75-P90 and P90-P100 (30-39 age group)

	Extended sample					
	<10	10- 25	<25-50	50-75	75-90	90-100
IRA coefficients	0.658***	0.140**	0.199***	0.318***	0.375***	0.618***
	(0.134)	(0.062)	(0.026)	(0.025)	(0.057)	(0.102)
Intercepts	-3.590	15.047***	12.831***	-1.799	-6.829	-26.339**
	(4.849)	(3.688)	(2.912)	(3.135)	(6.079)	(10.829)
N	6,429	12,362	20,675	22,490	12,588	7,109
Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Universal sample					
IRA coefficients	0.339***	0.076	0.224***	0.329***	0.349***	0.654***
	(0.098)	(0.057)	(0.025)	(0.026)	(0.058)	(0.106)
Intercepts	7.780**	12.229***	9.398***	-2.837	-6.062	-32.837***
	(3.781)	(3.316)	(2.837)	(3.169)	(6.257)	(11.277)
N	11,451	16,553	24,054	24,702	13,481	7,639
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort combining the universe of tax records and the Household Survey. Child and parents income is defined as the average of 5 yearly incomes but exclude zeros from the averages. Coefficients are OLS estimates. Controls: children's age, parent's age and sex

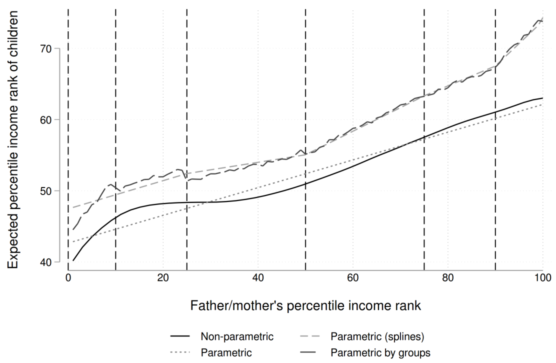
Table B10: Intergenerational Ranking Association with alternative reference distribution: own sample & without informal income. Non linear regression's estimates for income. Strict, extended and universal samples (30-39 age group)

Knots	Strict			Extended			Universal		
Average	0.265***			0.269***			0.268***		
	(0.006)			(0.004)			(0.002)		
0-25		0.266***	0.265***		0.196***	0.196***		0.151***	0.151***
		(0.037)	(0.037)		(0.021)	(0.021)		(0.019)	(0.019)
25-50		0.186***	0.186***		0.222***	0.222***		0.217***	0.217***
		(0.031)	(0.031)		(0.017)	(0.017)		(0.015)	(0.015)
50-75		0.259***	0.256***		0.298***	0.297***		0.316***	0.315***
		(0.032)	(0.032)		(0.019)	(0.019)		(0.017)	(0.017)
75-90		0.432***	0.455***		0.351***	0.355***		0.347***	0.352***
		(0.063)	(0.064)		(0.040)	(0.040)		(0.036)	(0.036)
90-99		0.646***	0.463***		0.806***	0.777***		0.825***	0.787***
		(0.135)	(0.153)		(0.092)	(0.102)		(0.084)	(0.093)
90-99			6.353***			1.783			2.170
N	30,193	30,193	30,193	82,519	82,519	82,519	98,977	98,977	98,977

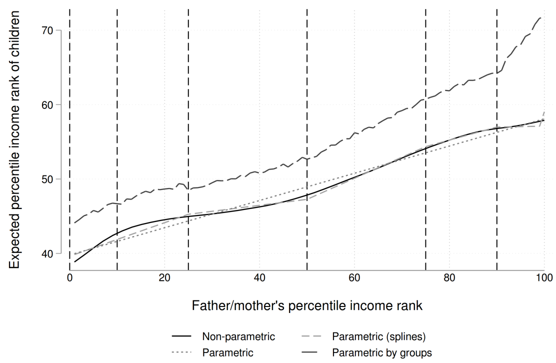
Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort of the sample of tax records. Coefficients are OLS estimates. : Knots at P25, P50, P75, P90, P99. Controls: children's age, parent's age and sex

Figure B1: Intergenerational Ranking Associations by gender: Non linear regression's estimates for income (30-39 age group). Knots at P10, P25, P50, P75, P90

Son

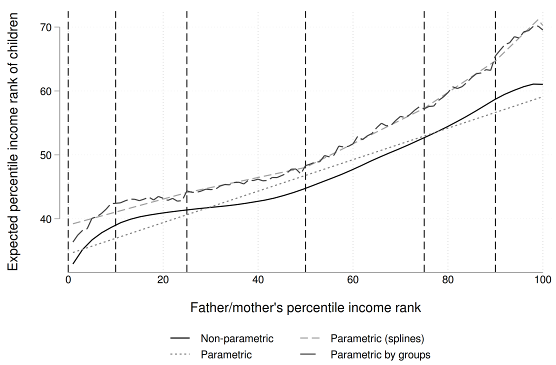


(a) Extended

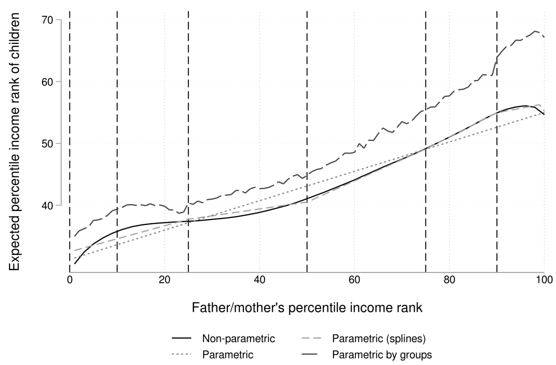


(b) Universal

Daughters



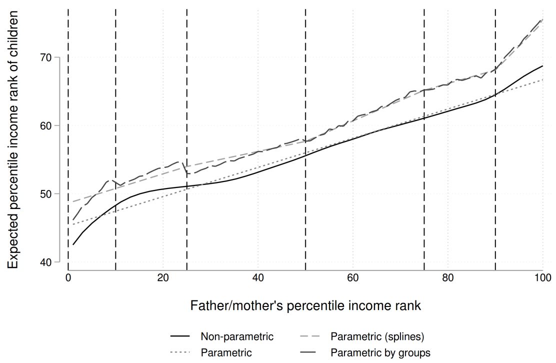
(c) Extended



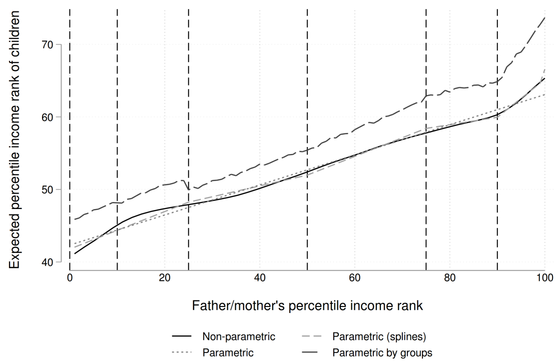
(d) Universal

Figure B2: Intergenerational Ranking Associations of mother's income by gender: Non linear regression's estimates for income (30-39 age group). Knots at P10, P25, P50, P75, P90

Son

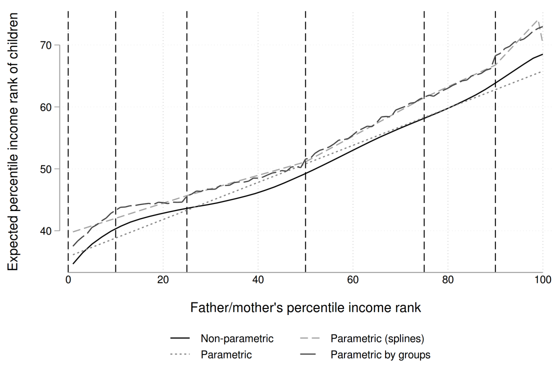


(a) Extended

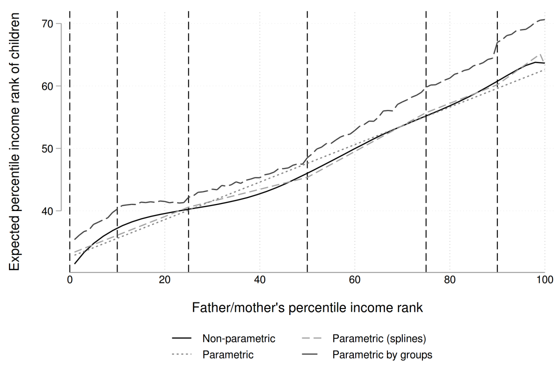


(b) Universal

Daughters



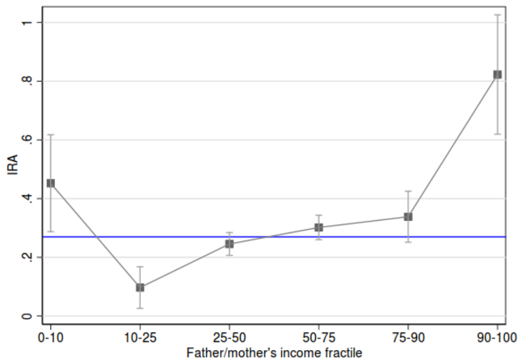
(c) Extended



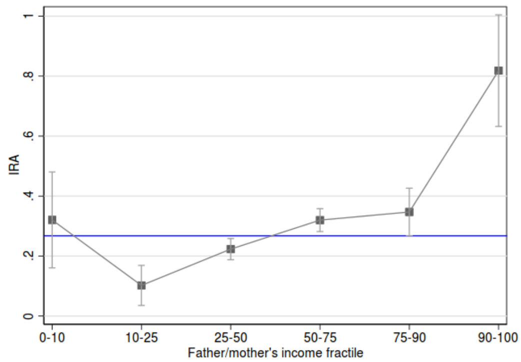
(d) Universal

Figure B3: Intergenerational Ranking Association with alternative reference distribution: own sample & without informal income. Non linear regression's estimates for income: knots at P10 P25, P50, P75, P90. Extended and universal samples (30-39 age group)

(a) Extended Sample



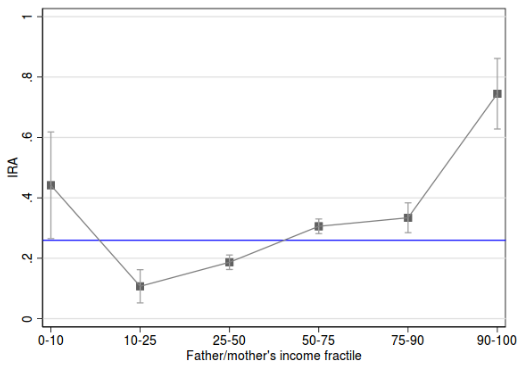
(b) Universal Sample



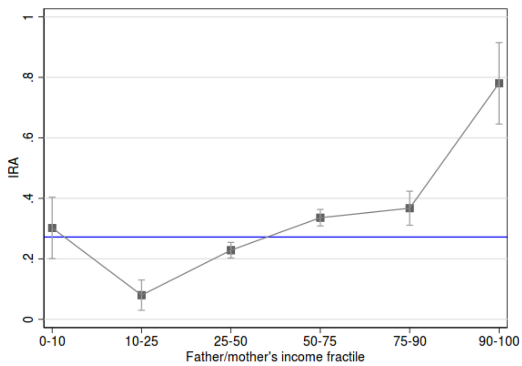
Coefficients are OLS estimates: Knots at P10, P25, P50, P75, P90. Blue line represents the average. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort of the universe of tax records. Controls: children's age, parent's age and sex

Figure B4: Intergenerational Ranking Association with alternative permanent income (average based on only positive income). Non linear regression's estimates for income (30-39 age group).

(a) Extended Sample



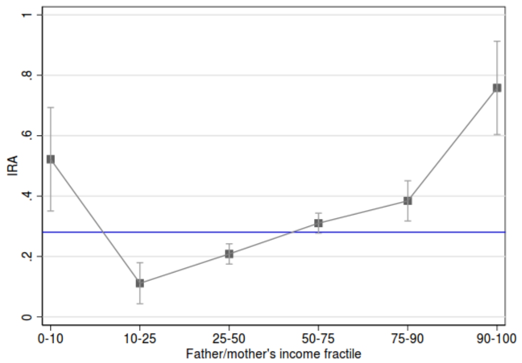
(b) Universal Sample



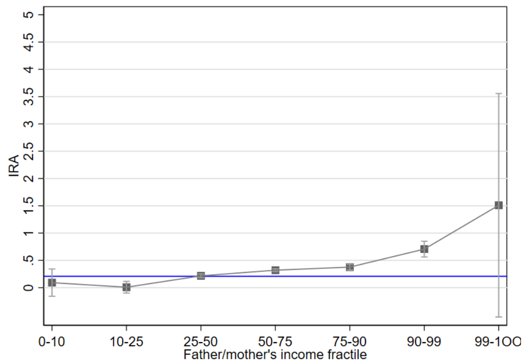
Coefficients are OLS estimates: Knots at P10, P25, P50, P75, P90. Blue line represents the average. The dependent variable is son's and daughter's income percentiles.Ranking are based on the individuals in the same cohort combining the universe of tax records and the Household Survey. Child and parents income is defined as the average of 5 yearly incomes but exclude zeros from the averages. Controls: children's age, parent's age and sex

Figure B5: Intergenerational Ranking Association: controlling by participation in the labor market. Non linear regression's estimates for income: knots at P10 P25, P50, P75, P90. Extended and universal samples (30-39 age group)

(a) Extended Sample



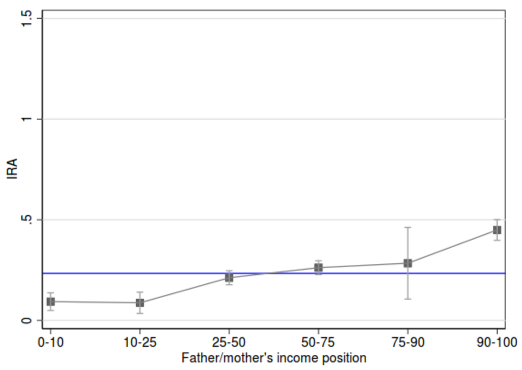
(b) Universal Sample



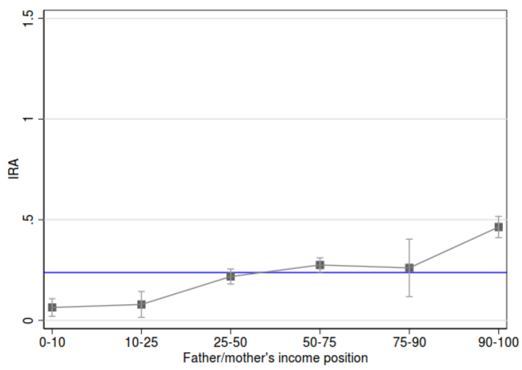
Coefficients are OLS estimates: Knots at P10, P25, P50, P75, P90. Blue line represents the average. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort of the universe of tax records. Controls: Parent's labor market participation, children's age, parent's age and sex

Figure B6: Intergenerational Ranking Association based on real position. Non linear regression's estimates for income (30-39 age group)

(a) Extended Sample



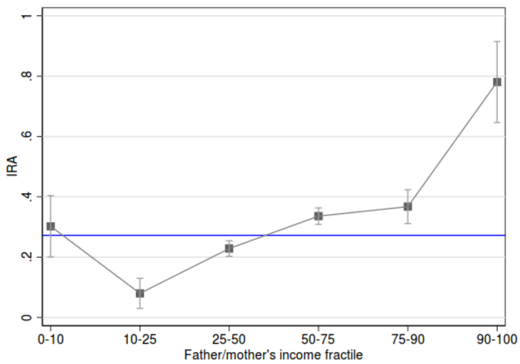
(b) Universal Sample



Coefficients are OLS estimates: Knots at P10, P25, P50, P75, P90. Blue line represents the average. The dependent variable is son's and daughter's income position. In both cases we averaged 5 yearly incomes. Position are based on the individuals in the same cohort of the complete income distribution. Controls: children's age, parent's age and sex

Figure B7: Intergenerational Income Ranking Associations: Non linear regression's estimates for income. Universal Sample X5 years. Knots at P10 P25, P50, P75, P90 (30-39 age group)

(a) Only parents with positive income



(b) Includes zeros (Parents and children)

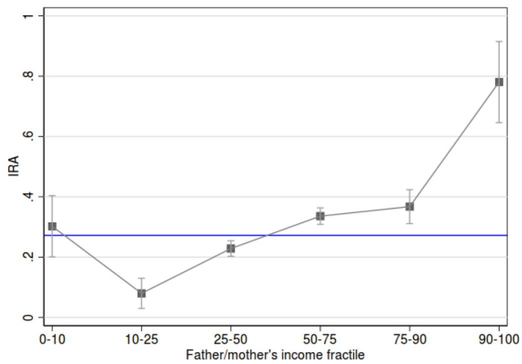
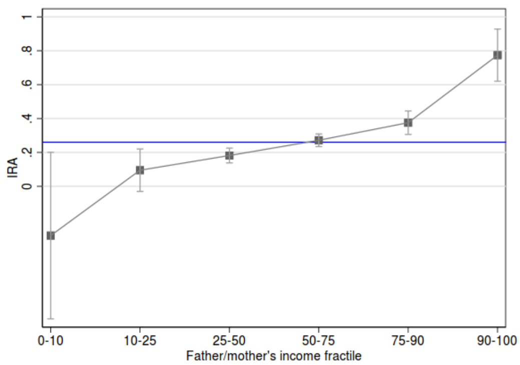
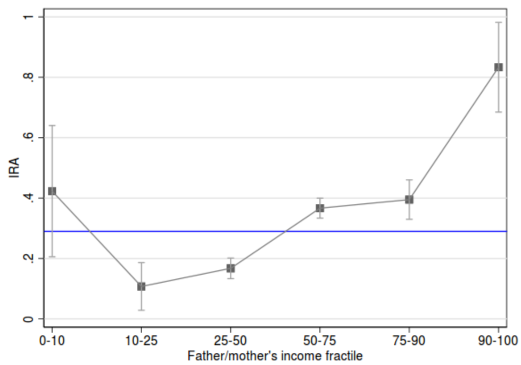


Figure B8: Intergenerational Ranking Association with alternative reference distribution: own sample & without informal income. Non linear regression's estimates for income: knots at P10 P25, P50, P75, P90, Extended X8 and universal X8 samples (30-39 age group)

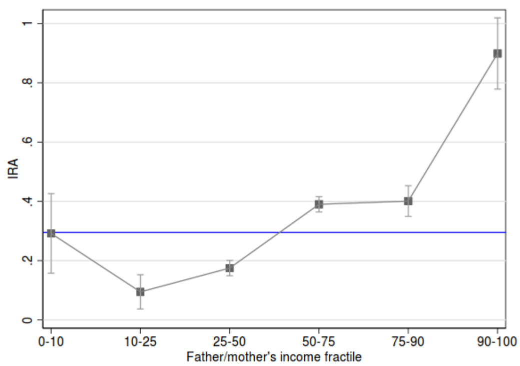
(a) Strict X8 sample



(b) Extended sample X8



(c) Universal X8 sample



Coefficients are OLS estimates: Knots at P10, P25, P50, P75, P90 in the cases of Extended X8 and Universal X8 samples. knot at P10 is excluded in the case of the Strict X8 sample. Blue line represents the average. The dependent variable is son's and daughter's income percentiles. Ranking are based on the individuals in the same cohort of the universe of tax records. Controls: children's age, parent's age and sex.

Table B11: Intergenerational Ranking Association with alternative reference distribution. Non linear regression's estimates for income. Alternative sample

	Extended (5y)	Universal 5y			Extended (8y)	Universal 8y		
Knots		Only positive	Parents w/zero	Both w/zero		Only positive	Parents w/zero	Both w/zero
0-10	0.492***	0.329***	0.297***	0.111***	0.417***	0.289***	0.292***	0.163***
	(0.101)	(0.051)	(0.055)	(0.037)	(0.086)	(0.065)	(0.069)	(0.036)
10-25	0.102**	0.121***	0.092***	0.053*	0.115***	0.135***	0.095***	0.036*
	(0.040)	(0.026)	(0.027)	(0.027)	(0.031)	(0.029)	(0.029)	(0.022)
25-50	0.249***	0.201***	0.223***	0.329***	0.166***	0.170***	0.175***	0.210***
	(0.020)	(0.014)	(0.014)	(0.017)	(0.013)	(0.013)	(0.013)	(0.011)
50-75	0.329***	0.337***	0.333***	0.386***	0.368***	0.385***	0.390***	0.401***
	(0.020)	(0.015)	(0.015)	(0.020)	(0.013)	(0.014)	(0.013)	(0.013)
75-90	0.474***	0.340***	0.376***	0.493***	0.399***	0.409***	0.401***	0.402***
	(0.039)	(0.033)	(0.031)	(0.045)	(0.026)	(0.029)	(0.026)	(0.028)
90-100	0.772***	0.765***	0.776***	0.726***	0.822***	0.917***	0.899***	0.900***
	(0.092)	(0.082)	(0.073)	(0.111)	(0.060)	(0.070)	(0.061)	(0.068)
R^2	61,509	165,376	134,145	133,811	127,786	169,052	144,175	249,185
N								

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is son's and daughter's income percentiles. Controls: children's age, parent's age and sex

Table B12: Intergenerational Ranking Association alternative estimates. Non linear regression's estimates for income. Alternative sample

Knots	(I)Own sample		(II)Control by Part.		(III)Positive		(IV)Real order	
Average	0.453***	0.319***	0.517***	0.371***	0.432***	0.273***	0.087***	0.062**
	(0.076)	(0.073)	(0.091)	(0.070)	(0.117)	(0.084)	(0.028)	(0.026)
0-10	0.107***	0.096***	0.101***	0.100***	0.105***	0.125***	0.091***	0.082**
	(0.033)	(0.030)	(0.036)	(0.034)	(0.037)	(0.035)	(0.034)	(0.038)
10-25	0.242***	0.229***	0.212***	0.239***	0.188***	0.135***	0.211***	0.220***
	(0.018)	(0.016)	(0.018)	(0.017)	(0.016)	(0.016)	(0.023)	(0.023)
25-50	0.292***	0.312***	0.320***	0.337***	0.305***	0.306***	0.265***	0.276***
	(0.019)	(0.017)	(0.018)	(0.018)	(0.016)	(0.016)	(0.022)	(0.021)
50-75	0.355***	0.349***	.371***	0.374***	0.335***	0.401***	0.279**	0.287***
	(0.040)	(0.036)	(0.036)	(0.037)	(0.032)	(0.031)	(0.113)	(0.088)
75-90	0.804***	0.823***	0.777***	0.789***	0.748***	0.648***	0.452***	0.449***
	(0.092)	(0.084)	(0.082)	(0.085)	(0.074)	(0.068)	(0.034)	(0.032)
90-100	82,519	98,977	82,519	98,977	83,014	97,252	66,731	75,224

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. In (I) the dependent variable is son's and daughter's income percentiles using own sample as reference. In (II) the dependent is son's and daughter's income percentiles. In (III) the dependent variable is son's and daughter's income percentiles using the average of positive income. In (IV) the dependent variable is son's and daughter's real position in the whole income distribution . Controls: children's age, parent's age and sex. Specifications (II) includes as control parent's labor market participation