3177G

Recent Trends in Intergenerational Educational Mobility in France

Abstract

This paper analyses trends in two-generation educational mobility and patterns of mobility across three generations.

Following the same methodology as Hertz et al. (2007), I find that two-generation persistence has fallen for individuals born in France between 1915 and 1990. The intergenerational regression coefficient showed no trend before 1940 but has thereafter halved, while the intergenerational correlation coefficient fell by 15% over the period.

Three-generation results suggest that exponentiating the two-generation coefficients underestimates long-run persistence by 40-60%, consistent with Braun and Stuhler (2018). The coefficient on grandparents' education is approximately 20% of the size of the coefficient on parents' education.

(7424 words)

1 Introduction

The public education system is designed to promote equality of opportunity and fight social and spacial inequality in academic and educational success, according to the French Education Code (2013). Although the relationship between intergenerational mobility[1] and equality of opportunity is not linear, the former remains a useful indicator of the latter (Jäntti and Jenkins, 2015). Studying intergenerational educational mobility could therefore help assess the effectiveness of the French education system in combatting inequalities based on social background, in a more direct way than studies of earnings mobility. Intergenerational educational mobility is also linked to other desirable outcomes, such as a narrower gender gap in mathematics and science attainment (Breda, Jouini and Napp, 2018). Increasing intergenerational educational mobility could therefore also reduce other inequalities.

This paper analyses trends in two-generation educational mobility and patterns of mobility across three generations. The first part of this paper provides the only analysis of two-generation educational mobility in France over an extended period. Fabre and Moullet (2004) and Ben-Halima, Chusseau and Hellier (2014), have focused on educational mobility at given points in time, whereas Torul and Oztunali (2017) provide estimates for France over a shorter period (for individuals born between 1940 and 1985) as part of their cross-country study.

The historical analysis shows that two-generation persistence was either stable or rising before the Second World War, and has since fallen. The intergenerational regression coefficient showed no trend before 1940, but thereafter has almost halved, whereas the intergenerational correlation coefficient has declined by about 15% throughout the period. Finally, the rank-rank slope coefficient has risen for individuals born before 1950, and fallen for those born after.

Two-generation persistence in France is compared to other European countries. I apply the same methodology as Hertz et al. (2007) to France to provide estimates that

[1] I use the term "persistence" in a descriptive sense to refer to a situation in which the socio-economic status of an individual is similar to that of its ascendants. "Mobility" is defined as the inverse of such persistence, unless otherwise specified.

are directly comparable to their results, as they do not include France in their study of intergenerational educational mobility across 42 nations. The average regression coefficient (between 0.53 and 0.63) and correlation coefficient (between 0.37 and 0.45) place France very close to the regional average for thirteen Western capitalist countries.

The second part of this paper focuses on mobility across three generations. Using a two-sample two-stage least squares (TS2SLS) approach to overcome data limitations, I present the first estimates of educational mobility across three generations for France. In the process, I provide new estimates of the inconsistency of TS2SLS in studies of intergenerational mobility, extending the work by Jerrim, Choi and Simancas (2016). The results suggest that exponentiating the two-generation coefficients overestimates long-run mobility by about 40 to 60%, consistent with estimates for Sweden (Lindahl et al., 2015) and Germany (Braun and Stuhler, 2018). Grandparents' educational outcomes are related to their grandchildren's, even when controlling for parents' education. The coefficient on grandparents' education is approximately 20% of the size of the coefficient on parents' education.

The paper is organised as follows. Section 2 reviews the literature and presents the three main indicators of mobility used in this paper. Section 3 introduces the data sets. Section 4 presents estimates of two-generation educational persistence from a historical and international perspective. Section 5 focuses on three-generation educational mobility. Section 6 concludes.

2 Measuring intergenerational educational mobility

The analysis of intergenerational educational mobility is at the intersection of two traditions: the "causal" literature and the social mobility literature. This paper places itself in the second tradition, focusing on documenting historical changes in mobility in France for cohorts born between 1915 and 1990. After reviewing the literature, this section presents three empirical measures of persistence: grade, standardised, and rank persistence.

2.1 The causal effect of parents' schooling on children's education

Becker and Tomes (1979, 1986) have developed the theoretical framework for studying the causal impact of parents' schooling on children's schooling. Empirically, Haveman and Wolfe (1995) have highlighted the importance of family background on children's educational attainment. In the case of France, Maurin and McNally (2008) estimate the impact of father's education on the children's probability of grade repetition. Using the university reform of 1968 as an instrument for father's education, they find that the probability of repeating a grade is 33 per cent lower for children whose father is more educated.

A limitation in this literature is that instrumental variables, twins, and adoption studies yield systematically different results, even when using the same data set (Holmlund, Lindahl and Plug, 2011). Moreover, there is a lack of a historical perspective, which is especially true for studies that use policy shocks as instruments. The nature of the identification strategy relies on a one-time exogenous change, making comparability across time difficult.

2.2 Empirical measures of intergenerational mobility

The intergenerational mobility literature restrains from making causal claims, but seeks to document changes in the association between parents and children's socio-economic status, over time or across countries. Different measures of socio-economic status have been used such as income, earnings, wealth, occupation or educational attainment (Black and Devereux, 2011). A practical advantage of focusing on educational mobility is that data on schooling is more readily available and less sensitive to life-cycle bias than earnings, as formal schooling tends to stay fixed once individuals reach adulthood (Blanden, 2013; Hertz et al., 2007).

I now present three measures of intergenerational mobility: the intergenerational regression coefficient, the intergenerational correlation and the intergenerational rank-rank slope.

2.2.1 Intergenerational regression and correlation coefficients

The benchmark regression relates parents' schooling S^i_{t-1} to child's schooling S^i_t, where υ^i_t is a well-behaved error term:

$$S^i_t = \alpha + \beta S^i_{t-1} + \upsilon^i_t \tag{1}$$

t indexes the generation, and i refers to a given family. The intergenerational income mobility literature traditionally uses logarithms, such that β becomes the intergenerational elasticity. In the case educational mobility, Hertz et al. (2007) use levels, and I will follow a similar methodology and terminology. S^i_t represents years of schooling, so β is not an elasticity but a measure of "grade persistence". The intergenerational correlation r (or "standardised persistence") is related to the regression coefficient β such that

$$r = \frac{\sigma_{t-1}}{\sigma_t}\beta \tag{2}$$

where σ is the standard deviation of years of schooling.

2.2.2 Intergenerational rank-rank slope

Rank-based measures come from decomposing the joint distribution of parent and child outcomes into their own marginal distributions and the copula, which is the joint distribution of parent and child ranks (Chetty et al., 2014*b*). The marginal distributions determine the level of inequality within a generation, whereas the copula determines relative intergenerational mobility.

The rank-rank slope ρ (or "rank persistence") is the correlation coefficient on the ranks of the child's and parents' education (Chetty et al., 2014*b*). As a measure solely based on the copula, it can be used on ordinal data and does not impose a linear relationship between variables, which overcomes some of the limitations of the intergenerational elasticity and correlation.

2.2.3 Comparing measures of persistence

The choice of the indicator of persistence depends on the conception of status. If we are interested in the association in absolute years of education between parents and children, β would be the preferred measure. If we believe that interpersonal differences are more important when they represent a larger proportion of the observed difference among people, then this would provide a basis for preferring the standardised measure of persistence, r (Hertz et al., 2007). If we believe that it is the relative position in the distribution that matters, this would justify focusing on the rank-rank slope, ρ.

Practical reasons would suggest preferring the correlation coefficient $\hat{r}$, as it is the most robust measure. An advantage of the regression coefficient $\hat{\beta}$ is that it is not biased by classical measurement error in S_t^i, in contrast to the correlation $\hat{r}$ (Black and Devereux, 2011). However, the analysis confirms the results by Hertz et al. (2007) and Jerrim, Choi and Simancas (2016) that $\hat{r}$ is less sensitive to coding assumptions than $\hat{\beta}$. Contrary to Chetty et al. (2014*a*), I find that $\hat{\rho}$ is not very stable, partly because the data presents large discrete jumps as there are only six categories for education. These practical reasons provide support for choosing the correlation coefficient as the preferred measure of mobility.

3 Data

3.1 Sample based on the *Formation et Qualification Professionnelle* surveys

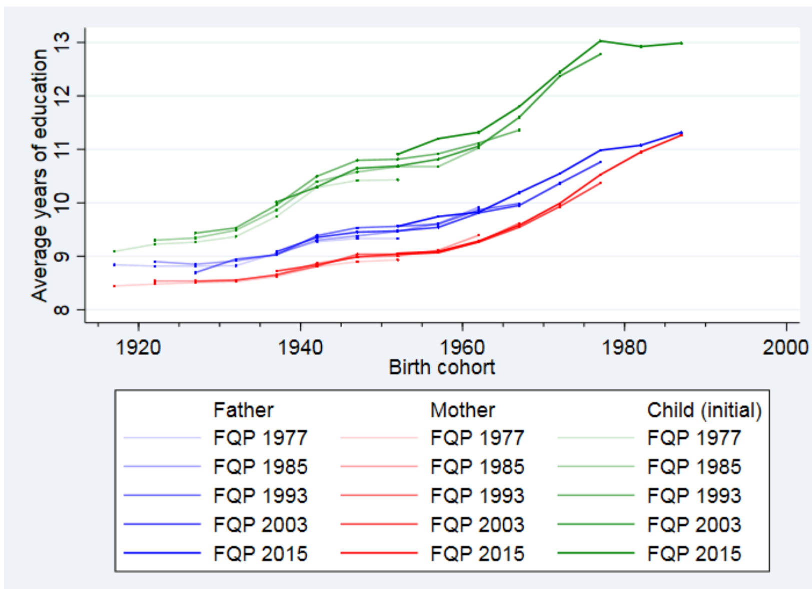
Unless stated otherwise, the analysis is based on the *Formation et Qualification Professionnelle* surveys (*FQP*, which translates to “Education, training and qualification”), organised by the French National Institute of Statistics and Economic Studies (Insee). There have been seven waves of surveys between 1964 and 2014-15. These surveys were designed to enable comparisons over time across the different waves, although there have been changes in the specific questions asked and in the sample design (Goux and Maurin, 1997). The *FQP* surveys contain information about the respondent's family background, such as the education and occupation of the respondent's father, mother, grandparents (occupation only), spouse's parents, and a randomly selected sibling. This makes the *FQP* surveys the main data source for French studies of intergenerational mobility (Goux, 2010; Thévenot and Monso, 2010).

In the two-generation analysis, I pool all waves since 1977 (which is the first wave that includes mother's education), restricting the sample to respondents born in Metropolitan France between 1915 and 1990. I divide the sample into five-year cohorts based on the respondent's year of birth, such that the whole sample contains 15 cohorts. The last cohort has been extended to include individuals born in 1990, covering six years instead of five. I exclude respondents aged less than 25 as a substantial proportion of younger individuals might not have completed their education. I exclude those aged more than 65 to limit measurement error due to inaccurate memories, and to avoid problems that might arise due to the correlation between longevity and education (Hertz et al., 2007). Respondents will be referred to as the children.

Educational attainment is recorded for the respondent and their parents. For both parents, this corresponds to the highest diploma they obtained at the time when the respondent was finishing their studies. For the respondent, this corresponds to the highest diploma completed during their initial training (studying without breaks longer than one year). This variable is preferred to total education as it is less subject to life-cycle bias. It is also preferred to the total number of years spent in school: spending an additional year at school without obtaining a qualification, because of grade repetition or a change in subject, is seen as a "lost year" and a negative signal in France (De Graaf, 2018).

Each survey has a different coding system for recording education. I create a standardised classification based on the minimum number of years of schooling needed to complete the diploma in the current education system, as reported in *The Structure of the European Education Systems* (European Commission / EACEA / Eurydice, 2017), assuming no grade repetition, and starting counting from the first year of compulsory education. I construct six categories: 8 years of schooling (no diploma), 9 years (end of primary school diploma or end of lower secondary school diploma), 11 years (vocational diploma lower than the *baccalauréat*), 12 years (*baccalauréat* and equivalent diplomas, which can be both vocational and academic), 14 years (undergraduate diploma or technician certificates) and 17 years (graduate diploma). 8 years of schooling are assigned to individuals with no diploma, reflecting the median grade attained by those individuals in *FQP* 2003. Further details about the classification are available upon request.

Figure 1: Average years of education, by cohort and survey



Notes: Education is the number of years necessary to obtain the highest diploma, starting from the first year of primary school, assuming no grade repetition. Sample weights have been applied.

Figure 1 shows the average years of education in the sample, by cohort and wave, using sample weights. For each given cohort, the mean years of education when considering different surveys are statistically different. This phenomenon has already been documented by Vallet (2017) for the *FQP* surveys in France. He argues that respondents might find it difficult to admit they have a low level of education in more recent surveys, which have occurred in a "more educated society". Although the difference is statistically significant, it is not sufficiently large to seriously bias the analysis.

3.2 Sample based on the European Social Survey

The second sample is based on the European Social Survey (ESS), which is the source Hertz et al. (2007) use for Estonia, Slovakia and Ukraine. The motivation for using this data set is to enable direct international comparisons with other countries presented in Hertz et al. (2007). The variables used are the highest level of education achieved by respondents and their parents. Answers are recorded into four categories, which are assigned the years of education in parentheses: less than lower education (8 years), lower secondary education completed (9 years), upper secondary education completed (12 years) and tertiary education completed (16 years).

Table 1: Summary statistics

			Sample size		Share enrolled	
Sample	Survey years	Birth years	Total	Min	20-24	25-29
ESS	2004	1935-84	1,095	85	0.33	0.04
FQP (partial)	1977-2015	1935-84	96,021	2,853	N/A	0.03
FQP (full)	1977-2015	1915-90	118,043	1,418	N/A	0.03

Table 1 presents summary statistics of the different samples. The first row presents the ESS sample. 33% of individuals in the last cohort are still enrolled in education, which suggests that results based on this cohort are likely to be biased. I follow Hertz et al. (2007) in including them for the sake of comparability, but do not do so for the broader analysis based on the *FQP* cohorts. The second row presents the restricted *FQP* sample, such that the dates of birth of individuals coincide with the ESS sample. The last row presents the full *FQP* sample.

4 Persistence across two generations

This section presents estimates of mobility across two generations. The first part is an analysis of historical changes in persistence for individuals born between 1915 and 1990. The second part focuses on individuals born between 1935 and 1984 and compares France to countries documented in Hertz et al. (2007).

4.1 Historical analysis

4.1.1 Grade and standardised persistence

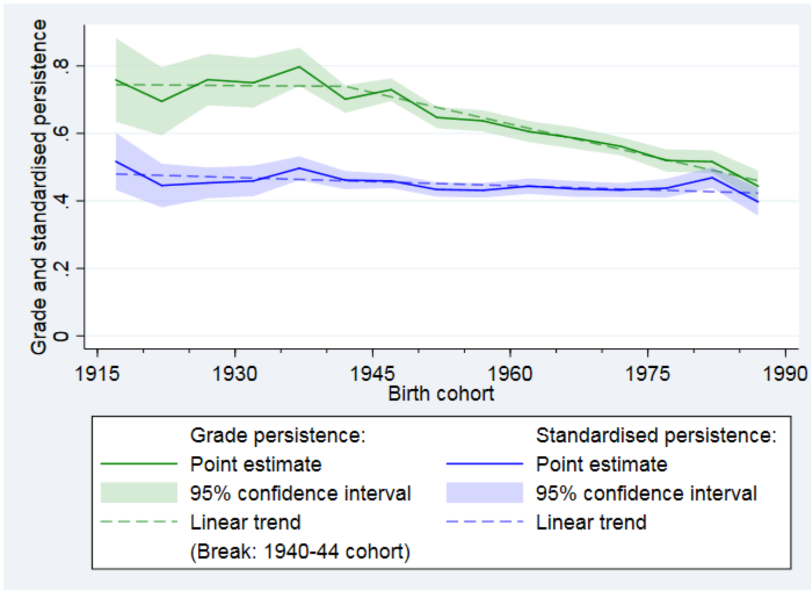
I calculate grade persistence following Hertz et al. (2007), regressing the child's years of education on the average years of education of the parents (or a single parent's if only one value is reported). The same method is used with standardised data to compute standardised persistence.

Coefficients are estimated separately for each five-year cohort using ordinary least squares (OLS) and applying sample weights. Figure 2 shows the estimated coefficients and 95% confidence intervals.

There is a substantial decline in grade persistence over the period, from 0.76 to 0.44. I allow for different linear trends before and after the sixth cohort (which is the date of the structural break estimated by supremum Wald and likelihood-ratio tests), restricting the coefficients so that there are no jumps in the levels. Results are shown in the top panel of table 2. While there was no significant trend before 1940, the trend in grade persistence becomes statistically significant after 1940. It implies that after 50 years, grade persistence would decrease by 0.311, such that grade persistence has almost halved in the second part of the twentieth century.

I conduct the same tests for standardised persistence, but the null of no structural break is not rejected, even at the 10% level. The slope of the coefficient is negative and statistically significant at the 5% level. However, the economic significance of the coefficient is small: it implies that intergenerational correlation would decrease by 0.041 after 50 years. At this rate, almost 300 years would be required for persistence to halve.

Figure 2: Grade and standardised persistence, 1915-90



Notes: Grade persistence corresponds to the coefficient in a regression of child's years of education on parents' years of education. Standardised persistence corresponds to the correlation coefficient.

4.1.2 Rank persistence

I calculate rank persistence following Chetty et al. (2014*b*). First, I rank children relative to others in their birth cohort based on their highest diploma. Similarly, I rank each parent relative to all parents of children in the same birth cohort, based on their highest diploma. These percentile ranks are calculated using weights to take into account the stratified nature of the sample. I regress the child's percentile rank on the the average of parents' percentiles (or a single parent's value if information is missing), and report the results in figure 3.

The null of no structural break is rejected at the 1% level. The supremum Wald test estimates the break date at the eighth cohort (1950-54), whereas the supremum likelihood-ratio test estimates it at the fifth cohort (1935-39). The bottom panel of table 2 shows the results from using the two different break dates while estimating

Table 2: Estimates of linear trends in grade, standardised and rank persistence

	Grade persistence		**Standardised persistence**
	Before break	After break	
Linear trend (per 5-year)	-0.0010 (0.0056)	-0.0311*** (0.0029)	-0.0041** (0.0018)
Constant (or change in constant)	0.746*** (0.025)	0.180*** (0.046)	0.484*** (0.016)
Observations	15		15
Pseudo R-squared	0.936		0.401

	Rank persistence			
	Break in 1935-39		Break in 1950-54	
	Before break	After break	Before break	After break
Linear trend (per 5-year)	0.0142*** (0.0035)	-0.0060*** (0.0013)	0.0051* (0.0025)	-0.0084*** (0.0025)
Constant (or change in constant)	0.416*** (0.014)	0.101*** (0.022)	0.441*** (0.014)	0.108** (0.035)
Observations	15		15	
Pseudo R-squared	0.619		0.490	

Notes: The results for grade and rank persistence come from a constrained linear regression, which allows for different trends before and after a structural break, while ensuring that there is no jump in the level.
The after-break constant refers to the change in the y-intercept.
*** p<0.01, ** p<0.05, * p<0.10. Robust standard errors in parentheses.

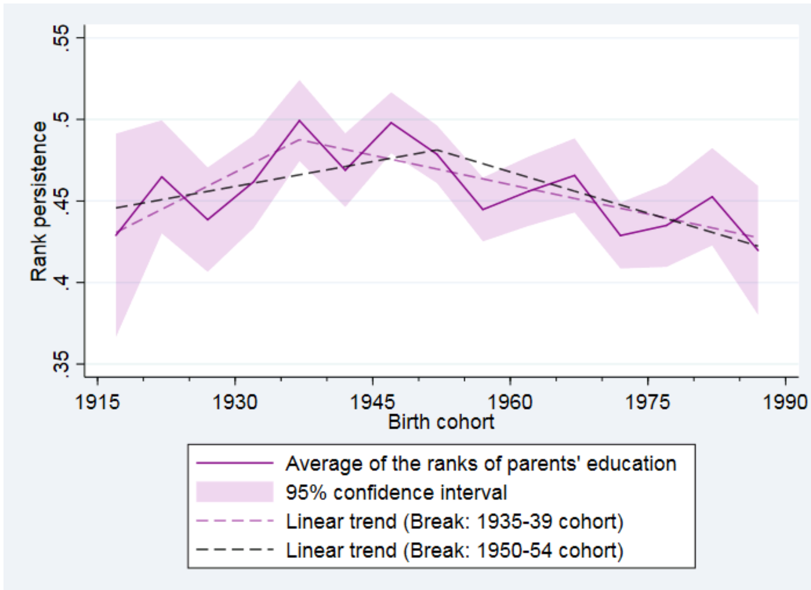
a linear trend.

Both sets of results suggest that rank persistence increased at the beginning of the century and is now decreasing, either by 0.060 or 0.084 over 50 years. This means that it would take between 130 and 170 years for rank persistence to halve.

4.1.3 Limitations and sensitivity analysis

Coding of minimum years of schooling As compulsory schooling laws have changed, assigning the same number of years of education for the "no diploma" category throughout the whole period might be inappropriate. To check the robustness

Figure 3: Rank persistence, 1915-90



Notes: Rank persistence corresponds to the coefficient in a regression of child's educational rank on their parents' educational rank.

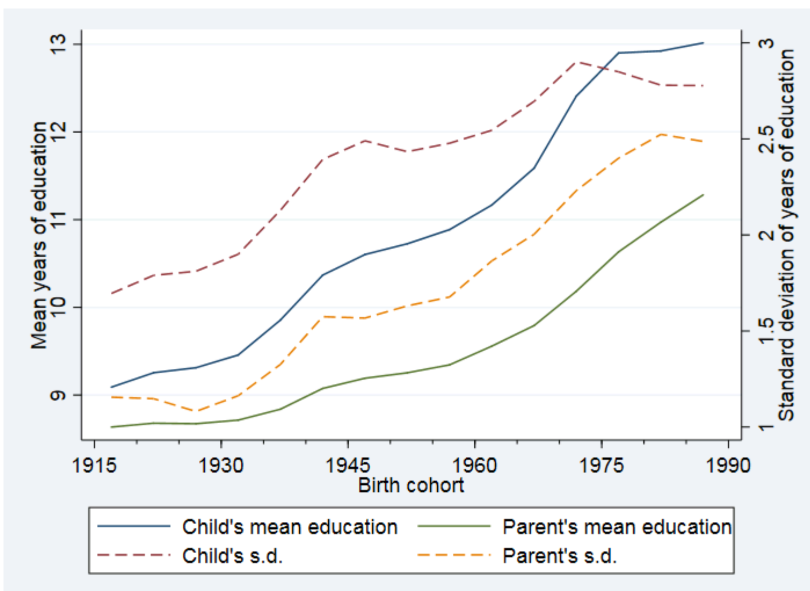
of the results, I assign 6 years of education for children and parents born before 1923 who have no diploma, and 7 years for those born between 1923 and 1953, instead of 8 years. Rank persistence is not affected by construction. Standardised persistence is not much affected. However, grade persistence is sensitive to the change, averaging 0.54 instead of 0.65. The structural break also occurs earlier, at the fourth cohort (1930-1934) or fifth cohort (1935-39). The coefficient on the linear trend after the break becomes -0.0102, three times smaller than before (in absolute value). These results highlight the sensitivity of the regression coefficient to different coding assumptions, giving a practical justification for preferring the correlation coefficient as an indicator.

Missing values Observations with missing values for own education (105 observations, 0.08% of the full sample) or for both father and mother's education (3,534 observations, 2.9% of the sample) are excluded. There is some evidence that information is not missing randomly: mean years of own education for individuals whose parents' education is missing is about 0.5 years lower, and statistically significant. Moreover, quantile regressions show that grade and standardised persistence is systematically lower for respondents with less education, while the pattern for rank persistence is less clear. This implies that the estimated $\hat{\beta}$ and $\hat{r}$ would be biased upwards. These biases are likely to be small given that less than 3% of the full sample is missing.

4.1.4 Historical discussion

All three measures suggest that intergenerational persistence has fallen in the second part of the twentieth century, but to different extents.

Figure 4: Means and standard deviations of years of education, by cohort



Notes: Education is divided into six categories, ranging from 8 to 17 years of education. Sample weights have been applied.

The evolution of $\hat{\beta}$ and $\hat{r}$ can be linked to the democratisation of education. Recall that the two measures are linked by $\hat{r} = \frac{\hat{\sigma}_{t-1}}{\hat{\sigma}_t}\hat{\beta}$. The rise in the average years of education, documented in figure 4, can be decomposed into an increase at the bottom of the distribution (which is likely to decrease $\hat{\beta}$ and $\hat{\sigma}$) and an increase at the top of the distribution (which is likely to increase $\hat{\sigma}$). The overall impact is an increase in $\frac{\hat{\sigma}_{t-1}}{\hat{\sigma}_t}$, which has counterbalanced the fall in $\hat{\beta}$, leaving $\hat{r}$ unchanged.

In the case of rank persistence, the break in the trend occurs later than for grade persistence. To affect $\hat{\rho}$, increasing the number of years of compulsory schooling is not sufficient: policies need to be explicitly targeted at promoting relative educational mobility. The timing of the break reflects the start of the reforms aimed at creating a comprehensive educational system. The French system was divided into strictly separated "orders" until its progressive unification after the War, starting with the Berthoin reform of 1959 (Prost, 1968). The law raised the age of compulsory schooling until 16 for all children born after 1953. It also provided a push towards the creation of comprehensive lower secondary schools, which was reinforced in 1963 and completed in 1975 with the introduction of the *collège unique* (La documentation française, 2001). These measures promoted educational equality and are associated with the fall in intergenerational educational rank persistence.

4.1.5 Comparison with previous studies

Fabre and Moullet (2004) estimate intergenerational mobility for individuals born between 1945 and 1973, using *FQP* 1993. They regress children's years of education on the most educated parent's estimated years of education, and find that the regression coefficient for individuals born before 1953 is 0.64, as compared to 0.52 for individuals born after 1953. Their results are similar to the ones found here: intergenerational education seems to have decreased in the post-war.

Ben-Halima, Chusseau and Hellier (2014) argue that using years of education masks the disparities in skills and incomes between individuals who have attended universities and those who went to *Grandes Écoles*. The authors first determine the wage value attached to each education level, using Mincerian equations, and use these to estimate intergenerational mobility. Using data from *FQP*, they find an increase in the coefficient on total father's influence from 0.383 to 0.479 between

1993 and 2003. The direct impact of father's education increases from 0.310 to 0.363. Contrary to the results presented here, Ben-Halima, Chusseau and Hellier find an increase in persistence over time.

The different results can be reconciled when considering the time scale and the object of study. Firstly, Ben-Halima, Chusseau and Hellier (2014) analyse changes over 10 years, as opposed to 75 years here. Secondly, instead of years of education, they focus on the earning potential associated with different schooling choices. If intergenerational persistence is heterogeneous, it is possible that the two measures evolve in different directions, especially if income inequality rises. A practical advantage of focusing on years of education is the availability of data over a longer period, and the possibility to compare results with other countries.

4.2 International comparisons

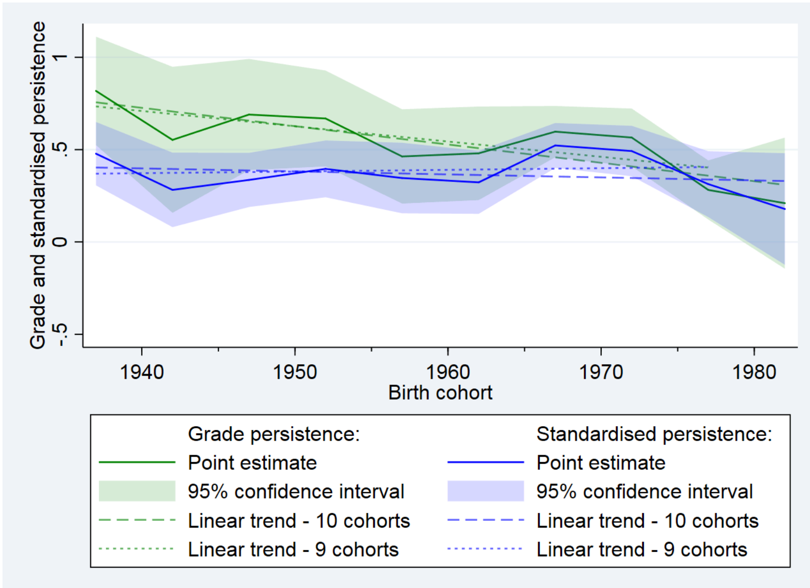
To make my results directly comparable to Hertz et al. (2007), I focus now on cohorts born between 1935 and 1984.

4.2.1 Results

Figure 5 shows the estimates of grade and standardised persistence based on the ESS data, and figure 2 on page 10 those based on the *FQP* surveys. For the 1935-84 period, the two samples yield qualitatively similar results: grade persistence has declined, whereas the decline in standardised persistence has been much smaller. This is consistent with the experience of most other European countries evidenced by Hertz et al. (2007), except for Great Britain (see table 3).

There are significant quantitative differences between the two sets of results: the results based on the ESS surveys show significantly lower levels of persistence. The levels of persistence found for France using the ESS data would put the country on par with Sweden, which is renowned for its high levels of equality and mobility (Hertz et al., 2007). However, the *FQP* data (with six categories) suggests France is closest to Ireland.

Figure 5: Grade and standardised persistence, 1935-84



Notes: Grade persistence corresponds to the coefficient in a regression of the child's years of education on the average of their parents' years of education. Standardised persistence corresponds to the correlation coefficient.
Source: European Social Survey (2016).

4.2.2 Sensitivity analysis

The difference between the two sets of results is mainly driven by life-cycle and attenuation biases in the ESS sample. A life-cycle bias seems to be present in the ESS sample, where persistence is extremely low for the two last cohorts. When excluding the last cohort of individuals aged 20-25, $\hat{\beta}$ and $\hat{r}$ rise to 0.57 and 0.39 respectively, closer to the results based on the *FQP* surveys. The second source of difference could come from classical measurement error. When using four categories instead of six for the *FQP* sample, the average regression coefficient for the 1935-84 period falls from 0.63 to 0.56, and the average correlation coefficient from 0.45 to 0.40. This is similar to the results based on the first nine cohorts of the ESS sample.

Hertz et al. (2007) do not provide additional details on the categorisation of

Table 3: Grade and standardised persistence for a sample of European countries

Country	Average of 5-year cohorts		Linear trend (per 5-year)	
	$\hat{\beta}$	$\hat{r}$	$\hat{\beta}$	$\hat{r}$
France†	0.65	0.45	-0.0219**	-0.0041**
FQP (1915-90)	(0.028)	(0.009)	(0.0028)	(0.0018)
France†	0.63	0.45	-0.0294**	-0.0033
FQP (1935-84)	(0.029)	(0.007)	(0.0030)	(0.0029)
France	0.53	0.37	-0.0414**	0.0022
ESS	(0.058)	(0.032)	(0.0160)	(0.0138)
Italy‡	0.67	0.54	-0.0339**	-0.0026
	(0.043)	(0.008)	(0.0048)	(0.0041)
Sweden‡	0.58	0.40	-0.0582**	-0.0050
	(0.076)	(0.024)	(0.0205)	(0.0098)
Ireland‡	0.70	0.46	-0.0700**	-0.0281
	(0.066)	(0.031)	(0.0169)	(0.0156)
Great Britain‡	0.71	0.31	0.0306	0.0283*
	(0.075)	(0.035)	(0.0321)	(0.0136)

Notes: $\hat{\beta}$ is the coefficient from the regression of child's years of education on the average years of education of the parents. $\hat{r}$ is the correlation coefficient. All reported values for the linear trend exclude cohorts of individuals aged less than 25.
** p<0.05, * p<0.10. Robust standard errors in parentheses.
† Excludes 20-24 years old.
‡ Source: Hertz et al. (2007).

education, making precise comparisons difficult. Using the two sets of results as a lower and upper bound, I find that France is close to the regional average for Western capitalist economies. Based on the correlation coefficient, which is the most stable measure, France was significantly more mobile than Italy, and less mobile than Denmark, Finland, Northern Ireland and Norway, on average. These results are consistent with Torul and Oztunali (2017), who divide Europe into four main groups and place France with the "Rest of Europe", along with Austria, Belgium, Germany and Ireland amongst others. This group is characterised by steadily declining grade and standardised persistence, for individuals born between 1940 and 1985.

5 Persistence across three generations

Although the data is not available to study trends in three-generation mobility, this analysis informs the discussion of trends in two-generation mobility. Stuhler (2012) and Lindahl et al. (2015) have highlighted that exponentiating the two-generation measures is inappropriate, as $(\beta_1)^2$ is not necessarily equal to β_2, where the subscript indicates the number of generations. Lindahl et al. (2015) present a simplified version of the Becker and Tomes (1979, 1986) model, focusing on the production function of human capital

$$H_t^i = \phi H_{t-1}^i + \tau E_t^i + \nu_t^i \tag{3}$$

$$E_t^i = \delta_t + \lambda E_{t-1}^i + \epsilon_t^i \tag{4}$$

where H_t^i is human capital of the ith family in the tth generation, E_t^i is the endowment, and ϵ_t^i and ν_t^i are random error terms. Stuhler (2012) show that the extrapolation error is 0 only if either $\lambda = 0$, $\tau = 0$, or if errors are zero. That is, either the endowment must be irrelevant in the production function of human capital, or it cannot be transferred across generations. These conditions are unlikely to be satisfied: iterating the two-generation measures of persistence could lead to an underestimation of long-run persistence.

5.1 Estimating grandparents' education

5.1.1 Two-sample two-stage least squares estimation

The *FQP* surveys contain information about the respondent's grandparents' occupation, but not education. To overcome this limitation, I use a two-sample two-stage least squares (TS2SLS) approach, which is frequently used in studies of intergenerational mobility (see Jerrim, Choi and Simancas, 2016 for a review).

The data comes from the 1985, 1993 and 2003 waves of the *FQP* surveys, as they share a common classification for occupation. The main sample consists of the 2003 survey, whereas the auxiliary sample is based on fathers from the 1985 and 1993 surveys.

Figure 6: Relationship between main and auxiliary samples
for TS2SLS estimation

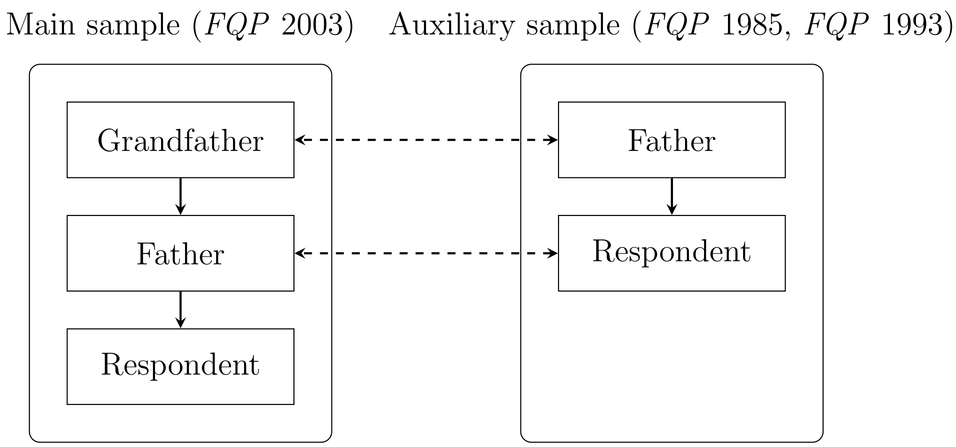


Figure 6 represents the relationship between the two samples. The key assumption is that fathers from the auxiliary sample are drawn from the same population as grandfathers from the main sample. Both samples are restricted to respondents born in Metropolitan France and aged between 25 and 65 at the time of the survey.

The auxiliary sample is further restricted to respondents who have at least one child. The age constraint excludes respondents born before 1921.

The main sample only includes respondents whose parents were born in Metropolitan France. Since all respondents in the auxiliary sample were born after 1921, fathers in the main sample also need to be born after 1921. I drop the 1954-58 cohort and all previous cohorts because of the high proportion of fathers born before 1921 (16.0% of all fathers for the 1954-58 cohort, compared to 5.3% for the 1959-63 cohort).

In the first stage, using data from the auxiliary sample, I regress father's education S^i_{t-1} on a categorical variable describing his occupation, for each five-year birth cohort starting from 1921. Here, $\boldsymbol{Z}^i_{t-1}$ is a vector of occupations, $\boldsymbol{\gamma}$ a vector of coefficients.

$$S^i_{t-1,aux} = \boldsymbol{Z}^i_{t-1,aux}\boldsymbol{\gamma} + \varepsilon^i_{t-1,aux} \tag{5}$$

The estimated coefficients $\hat{\boldsymbol{\gamma}}$ are used to predict grandfather's years of education in the main sample.

$$\hat{S}^i_{t-2,main} = \boldsymbol{Z}^i_{t-2,main}\hat{\boldsymbol{\gamma}} \tag{6}$$

In the second stage, for each birth cohort, I regress child's education on the grandfather's predicted years of education.

$$S^i_{t,main} = \beta_2 \hat{S}^i_{t-2,main} + v^i_{t,main} \tag{7}$$

I repeat the analysis based on ranks to estimate the long-term rank persistence $\hat{\rho}_2$. I also compute the long-term correlation coefficient $\hat{r}_2$.

5.1.2 Inconsistency of TS2SLS

In this section, I attempt to estimate the size of the inconsistency of TS2SLS estimates in the case of three-generation educational mobility. This provides a basis for scaling the results in the next section. Throughout the discussion, OLS coefficients are assumed to be consistent estimates of persistence, against which TS2SLS estimates can be compared. This is a descriptive statement: OLS coefficients might not reflect the causal impact of parents' schooling on children's schooling, but will accurately describe the association between the two observed variables that are measured without error.

The TS2SLS estimates will be consistent if either the R-squared of the equation used to predict the grandfather's education equals one, or if the grandfather's occupation has no direct effect upon grandchild's education (Jerrim, Choi and Simancas, 2016). The data shows that the first condition is not satisfied, and the second condition is highly unlikely to hold.

To estimate the size of the inconsistency, I use the fact that male respondents in the auxiliary sample can be matched to fathers in the main sample. In the first stage, I regress education on occupation for each five-year birth cohort based on the auxiliary sample. Father's education is then predicted in the main sample. In the second stage, I calculate the three measures of two-generation persistence, and compare them to OLS estimates. I repeat this procedure separately for the four cohorts, and report the average over all four cohorts in table 4.

Table 4: Estimates of the inconsistency of TS2SLS compared to OLS

Average over four cohorts	$\hat{\beta}$	$\hat{r}$	$\hat{\rho}$
TS2SLS estimates	0.615	0.343	0.293
OLS estimates	0.421	0.374	0.317
Inconsistency of TS2SLS	0.194	-0.031	-0.024
Percentage	46.0%	-8.1%	-7.6%

Notes: Father's education is predicted based on occupation and birth cohort. Child's years of education is regressed on father's years of education (predicted for TS2SLS, observed for OLS) to obtain $\hat{\beta}$. $\hat{r}$ is the correlation coefficient, and $\hat{\rho}$ the rank-rank slope.

Theory predicts that intergenerational earnings elasticities are upwardly inconsistent (Blanden, 2013). TS2SLS estimates are about 30 per cent higher than OLS coefficients according to Björklund and Jäntti (1997), or between 13 and 45 per cent higher (in the absence of life-cycle bias) according to Jerrim, Choi and Simancas (2016). The results shown in table 4 confirm the upward inconsistency of $\hat{\beta}_{TS2SLS}$ in the case of educational grade persistence, by 46% on average.

The inconsistency in $\hat{r}_{TS2SLS}$ and $\hat{\rho}_{TS2SLS}$ could go in either direction. Björklund and Jäntti (1997) find that TS2SLS estimates of intergenerational earnings correlation are about 30 per cent higher than OLS estimates, whereas Jerrim, Choi and Simancas (2016) find they are between 3 and 18 per cent lower than OLS estimates. Here, I find that the correlation coefficient is downwardly inconsistent by about 8%. Rank persistence also tends to be underestimated by $\hat{\rho}_{TS2SLS}$ by about 8%.

A limitation of TS2SLS is the loss of precision caused by the fact that information is not observed directly but predicted, so the imputed regressors are subject to sampling error. The usual OLS standard errors for the second stage will be biased. Murphy and Topel (1985) and Inoue and Solon (2010) propose a procedure to obtain the correct covariance matrix in the second stage, whereas Björklund and Jäntti (1997) rely on bootstrapping to compute appropriate standard errors.

5.2 Estimates of long-run persistence

5.2.1 First stage: predicting grandfather's education

Based on the auxiliary sample, I regress father's education on father's occupation separately for each birth cohort. The most frequent of the 29 possible occupations are, in order, farmer, unskilled industrial worker, and craftsman, for which the results are reported in table 5. To these are added the results for the professions, as it is one of the occupations with the highest education. I report the results for the first cohort from the auxiliary sample (born in 1921-25) and the 1956-60 cohort, which is the second-to-last cohort (only one father in the main sample is born after 1960). The R-squared of the first stage lies between 0.372 and 0.493 for all cohorts, averaging 0.445 across cohorts. The F-statistic is at least 10 for all cohorts, implying that the relevance condition is satisfied. These coefficients are used to predict grandparent's education in the main sample.

Table 5: Father's education, by occupation and birth cohort

Father's education	1921-25	1956-60
Farmer	8.43	8.70
	(0.06)	(0.03)
Craftsman	8.75	9.62
	(0.12)	(0.11)
Professions	14.80	16.02
	(1.01)	(0.43)
Unskilled industrial worker	8.38	8.57
	(0.09)	(0.06)
Observations	1,374	3,762
R-squared	0.372	0.429

Notes: Based on the auxiliary sample, father's education is regressed on a categorical variable for occupation (29 categories) separately for each five-year birth cohort. Results are reported for four occupations, and for the first and second-to-last cohorts. Robust standard errors in parentheses. All results are statistically significant at the 1% level.

5.2.2 Second stage: estimates of persistence across three generations

Table 6 shows the second-stage results. The first column shows the estimates for grade persistence $\hat{\beta}$, the second column reports standardised persistence $\hat{r}$, and the third column rank persistence $\hat{\rho}$.

Table 6: Estimates of long-run educational persistence over three generations

Average over four cohorts	$\hat{\beta}$	$\hat{r}$	$\hat{\rho}$
Grandfather to father (TS2SLS)	0.75	0.40	0.29
Scaled estimate	*0.51*	*0.44*	*0.32*
Average R-squared	0.16	0.16	0.11
Total observations	8,859	8,859	8,859
Father to child (OLS)	0.42	0.38	0.32
Average R-squared	0.14	0.14	0.12
Total observations	9,024	9,024	9,024
Grandfather to child (TS2SLS)	0.44	0.21	0.15
Scaled estimate	*0.30*	*0.23*	*0.16*
Average R-squared	0.04	0.04	0.04
Total observations	8,859	8,859	8,859
Predicted grandfather to child	0.22	0.17	0.10
Observed (scaled) - Predicted	0.09	0.06	0.06
Percentage difference	39%	38%	62%

Notes: Grandfathers' education is estimated based on occupation and father's birth cohort.
The scaled estimates come from dividing the TS2SLS estimate by 1.46 for $\hat{\beta}$ and 0.92 for $\hat{r}$ and $\hat{\rho}$, reflecting the estimated inconsistency from table 4.
Predicted grandfather-to-child persistence equals the scaled grandfather-to-father coefficient multiplied by the father-to-child coefficient.
The number of observations is lower for TS2SLS estimates due to missing information about grandfather's occupation. Repeating the analysis for father-to-child persistence using the same subsample does not alter the coefficients.

The top panel shows TS2SLS estimates of grandfather-to-father persistence. Results are scaled in the second row, based on the previous analysis of the inconsistency of TS2SLS. The second panel shows OLS estimates of father-to-child persistence. The third panel presents TS2SLS and scaled estimates of grandfather-to-child persistence.

Grandfather-to-child persistence seems to be slightly lower than in Germany, although this depends on assumptions about the size of the inconsistency of TS2SLS. The scaled $\hat{\beta}$ is 0.30 compared to 0.354 for Germany, and the scaled $\hat{r}$ is 0.23 compared to 0.251 for Germany (Braun and Stuhler, 2018).

The last panel presents predicted grade persistence based on the exponentiation fallacy. Since the coefficients are not stable over time, the prediction is based on multiplying the scaled grandfather-to-father persistence by the father-to-child persistence. Results show that iteration would underestimate long-term grade and standardised persistence by about 40%, which is consistent with Stuhler (2012), Lindahl et al. (2015) and Braun and Stuhler (2018). The results for rank persistence suggest a 62% underestimation.

5.3 Direct impact of grandparents' education

The previous analysis has shown that long-run persistence is underestimated by iteration of two-generation coefficients. Here, I attempt to decompose the different family influences on child's education. I regress different measures of children's educational outcomes on information about their father, mother, paternal grandfather and maternal grandfather. Grandmothers are left out as data on grandfathers and grandmothers is not simultaneously available. I pool all four cohorts to boost the sample size and improve the precision of the results. Cohort fixed effects have been applied, and are statistically significant, especially for the regressions based on years of education. Table 7 shows the results of the decomposition.

Results reported based on years of education are shown in the first two columns, those based on standardised education in the middle two columns, and those based on ranks are in the last two columns. For each measure, the first column regresses child's educational outcome on father's outcome, while the second column adds mother's and grandfathers' education.

Three main findings emerge. Firstly, the addition of additional family members only slightly improves the fit of the model, increasing the R-squared by 15-31%.

Table 7: Estimates of the association in educational outcomes between child and different family members

Child	Years of education		Standardised score		Rank	
Father	0.421***	0.253***	0.376***	0.225***	0.318***	0.181***
	(0.012)	(0.017)	(0.011)	(0.015)	(0.009)	(0.012)
Mother		0.258***		0.197***		0.192***
		(0.018)		(0.014)		(0.012)
Paternal grandfather		0.075***		0.036***		0.036***
		(0.028)		(0.014)		(0.010)
Maternal grandfather		0.053*		0.029**		0.036***
		(0.028)		(0.013)		(0.010)
Constant	7.164***	5.372***	-0.000	0.035	28.785***	27.216***
	(0.125)	(0.267)	(0.020)	(0.023)	(0.665)	(0.771)
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,024	7,237	9,024	7,237	9,024	7,237
R-squared	0.198	0.227	0.141	0.171	0.123	0.161

Notes: Grandfathers' education is predicted based on occupation and parent's birth cohort.
Naive robust standard errors are reported in parentheses, which have not been corrected for the bias in the covariance matrix due to the predicted nature of grandfathers' education.
The constant in the fourth column is not zero because of missing observations, as variables were standardised prior to the estimation.

However, the coefficient on father's education falls dramatically, by 66-76%. This is suggestive of the prevalence of assortative mating, and guards us against interpreting these coefficients causally.

Secondly, the coefficients on father and mother are not significantly different, suggesting that looking at mother-child relationships is a viable alternative to the traditional focus on father-child associations, at least for recent cohorts.

Thirdly, the size of the coefficients on grandfathers' education are around 20% of that of parents. It is about 20-30% in the specification based on years of education, but these coefficients are likely to be biased given the upward inconsistency of $\hat{\beta}_{TS2SLS}$. The ratio is about 15-16% for standardised education, and 19-20% in the rank specification, and both are likely to be downwardly inconsistent. Note that grandmothers' education is not controlled for in the regression, so the coefficients on grandfathers also includes the impact of assortative mating and the indirect impact of grandmothers' education.

The evidence presented suggests that two-generation measures of persistence cannot accurately predict long-term intergenerational persistence. The relevance of grandparents could be the result of regression to the mean, or could be driven by a causal mechanism such as dynastic endowment, where E_t depends on $\bar{E}$, or follows an AR(2) process instead of a simple AR(1) process (Lindahl et al., 2015). The analysis here is purely descriptive and cannot distinguish between the statistical and the causal explanation.

6 Conclusion

This paper has studied recent trends in two-generation educational mobility, as well as patterns of long-run mobility across three generations.

The two-generation results suggest that persistence has declined. Grade persistence showed no trend before 1940 but has thereafter halved, while the intergenerational correlation coefficient fell by 15% over the period. Finally, the rank-rank slope coefficient has risen for individuals born before 1950 and fallen for those born after.

The three-generation results show that long-run mobility would be overestimated by exponentiating the two-generation coefficients. Grandparents' educational outcomes are related to their grandchildren's, even after controlling for parents' education. The coefficient on grandparents' education is approximately 20% of the size of the coefficient on parents' education.

Further research could focus on sibling correlations as an alternative indicator of intergenerational educational mobility. Solon (1999) and Jäntti and Jenkins (2015) show that the sibling correlation is $r_{sib} = r^2 + s$ where r is the intergenerational educational correlation coefficient and s represent the correlation of all other shared factors that are unrelated to parental education. *FQP* 2003 and 2015 contain detailed information on sibling's education and occupation, making the study of sibling correlations in education over time a natural extension of this paper.

References

Becker, Gary, and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87(6): 1153.

Becker, Gary, and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3): S1–S39.

Ben-Halima, Bassem, Nathalie Chusseau, and Joel Hellier. 2014. "Skill premia and intergenerational education mobility: The French case." *Economics of Education Review*, 39(Supplement C): 50–64.

Björklund, Anders, and Markus Jäntti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *The American Economic Review*, 87(5): 1009–1018.

Black, Sandra, and Paul Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*. Vol. 4B, ed. David Card and Orley Ashenfelter, Chapter 16, 1487–1541. Elsevier.

Blanden, Jo. 2013. "Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology." *Journal of Economic Surveys*, 27(1): 38–73.

Braun, Sebastian Till, and Jan Stuhler. 2018. "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany." *The Economic Journal*, 128(609): 576–611.

Breda, Thomas, Elyès Jouini, and Clotilde Napp. 2018. "Societal Inequalities Amplify Gender Gaps in Math." *Science*, 359(6381): 1219–1220.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014*a*. "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, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014*b*. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review*, 104(5): 141–47.

Code de l'éducation. 2013. "Article L111-1. Version en vigueur au 10 juillet 2013." (Accessed 2 April 2018).

De Graaf, Douglas. 2018. "Perdre une année dans ses études, une crainte très française." *Le Monde*. Available at: `http://www.lemonde.fr/campus/article/2018/03/26/perdre-une-annee-dans-ses-etudes-une-crainte-tres-francaise_5276681_4401467.html` (Accessed 27 March 2018).

European Commission / EACEA / Eurydice. 2017. *The Structure of the European Education Systems 2017/18: Schematic Diagrams". Eurydice Facts and Figures*, Luxembourg:Publications Office of the European Union.

European Social Survey. 2016. "ESS Round 2: European Social Survey Round 2 Data (2004)." Data file edition 3.5. NSD - Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC.

Fabre, Alice, and Stéphanie Moullet. 2004. "Externalités de l'éducation et mobilité intergénérationnelle : application au cas français." *Economie & prévision*, 166(5): 19–37.

Goux, Dominique. 2010. "L'enquête *Formation et Qualification Professionnelle* : une source inépuisable pour la recherche en sciences sociales." *Economie et Statistique*, 431(1): 3–11.

Goux, Dominique, and Eric Maurin. 1997. "Meritocracy and Social Heredity in France: Some Aspects and Trends." *European Sociological Review*, 13(2): 159–177.

Haveman, Robert, and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33(4): 1829–1878.

Hertz, Tom, Tamara Jayasunderay, Patrizio Pirainoz, Sibel Selcuk, Nicole Smith, and Alina Verashchaginazz. 2007. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *The B.E. Journal of Economic Analysis & Policy*, 7(2 (Advances)). Article 10.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods." *Journal of Economic Literature*, 49(3): 615–51.

Inoue, Atsushi, and Gary Solon. 2010. "Two-Sample Instrumental Variables Estimators." *The Review of Economics and Statistics*, 92(3): 557–561.

Insee (Institut national de la statistique et des études économiques). 2011*a*. "Formation, qualification professionnelle (FQP) - 1985." Version 2. Distributed by ADISP-CMH.

Insee (Institut national de la statistique et des études économiques). 2011*b*. "Formation, qualification professionnelle (FQP) - 1993." Version 2. Distributed by ADISP-CMH.

Insee (Institut national de la statistique et des études économiques). 2011*c*. "Formation, qualification professionnelle (FQP) - 2003." Version 2. Distributed by ADISP-CMH.

Insee (Institut national de la statistique et des études économiques). 2016. "Formation, qualification professionnelle (FQP) - 1977." Version 2. Distributed by ADISP-CMH.

Insee (Institut national de la statistique et des études économiques). 2017. "Formation, qualification professionnelle (FQP) - 2014-2015." Version 1. Distributed by ADISP-CMH.

Jerrim, John, Álvaro Choi, and Rosa Simancas. 2016. "Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility: how consistent are they?" *Survey Research Methods*, 10(2): 85–101.

Jäntti, Markus, and Stephen Jenkins. 2015. "Income Mobility." In *Handbook of Income Distribution*. Vol. 2, ed. Anthony Atkinson and François Bourguignon, Chapter 10, 807–935. Elsevier.

La documentation française. 2001. "Le collège unique de 1975 aux années 2000 : Mise en place et réformes." Available at: `http://www.ladocumentationfrancaise.fr/dossiers/d000130-le-college-unique-de-1975-aux-annees-2000/mise-en-place-et-reformes` (Accessed 9 April 2018).

Lindahl, Mikael, Mårten Palme, Sofia Sandgren Massih, and Anna Sjögren. 2015. "Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations." *Journal of Human Resources*, 50(1): 1–33.

Maurin, Eric, and Sandra McNally. 2008. "Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students." *Journal of Labor Economics*, 26(1): 1–33.

Murphy, Kevin, and Robert Topel. 1985. "Estimation and Inference in Two-Step Econometric Models." *Journal of Business & Economic Statistics*, 3(4): 370–379.

Prost, Antoine. 1968. *Histoire de l'enseignement en France, 1800-1967.* Paris:Armand Colin.

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*. Vol. 3, ed. Orley C. Ashenfelter and David Card, Chapter 29, 1761–1800. Elsevier.

Stuhler, Jan. 2012. "Mobility Across Multiple Generations: The Iterated Regression Fallacy." Institute for the Study of Labor (IZA) IZA Discussion Papers 7072.

Thévenot, Laurent, and Olivier Monso. 2010. "Les questionnements sur la société française pendant quarante ans d'enquêtes *Formation et Qualification Professionnelle*." *Economie et Statistique*, 431(1): 13–36.

Torul, Orhan, and Oguz Oztunali. 2017. "Intergenerational Educational Mobility in Europe." Unpublished manuscript. Version: 2 December 2017. Latest version available at: `http://www.econ.boun.edu.tr/torul/ieme.pdf` (Accessed 3 December 2017).

Vallet, Louis-André. 2017. "Mobilité entre générations et fluidité sociale en France. Le rôle de l'éducation." *Revue de l'OFCE*, 150: 27–67.