Dynastic Human Capital, Inequality and Intergenerational Mobility*

by

Adrian Adermon,# Mikael Lindahl¤ and Mårten Palme§

This version: 23 March 2016

First version: September 2014

Abstract: We study the importance of the extended family – or the dynasty – for persistence in human capital and inequality across generations. We use data including the entire Swedish population, linking four generations. This data structure enables us to - in addition to parents, grandparent and great grandparents - identify aunts and uncles and their spouses, as well as parents' cousins and parents' aunts and uncles. We use three different indicators of social status – years of schooling, average earnings and an index of occupational status. Our results suggest that both the inclusion of several indicators as well as the dynastic "group effect" are important, and that traditional intergenerational persistence estimates misses about half the persistence across generations estimated by the extended model.

* We are grateful for comments given on seminars at the Department of Economics as well as Statistics at Uppsala University, Helsinki School of Economics, IFAU, ESPE, UCLS, the Family and Education workshop in Hawaii in January 2015 and ISER Workshop on Family Economics at University of Essex in September 2015. Mikael Lindahl is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Torsten and Ragnar Söderberg Foundation, the Scientific Council of Sweden and the European Research Council [ERC starting grant 241161]. Mårten Palme gratefully acknowledges financial support from the Swedish Council of Social Research.

Department of Economics, Uppsala University. E-mail: Adrian.Adermon@nek.uu.se.

¤ Department of Economics, Gothenburg University. E-mail: Mikael.Lindahl@economics.gu.se.

§ Department of Economics, Stockholm University. E-mail: Marten.Palme@ne.su.se.

1. Introduction

It is well known that group effects are of key importance for how inequality is upheld across generations. Following the social capital theory (see Coleman, 1988) and the strain theory (see Merton, 1938) studies have shown the importance of social class. There is also a large empirical literature following e.g. Myrdal (1944) on the importance of race and ethnicity in this context. However, applied work in economics on intergenerational mobility is almost exclusively constrained to the relation between parents and children. A third group of potential importance for an individual's life outcomes is the extended family, or the dynasty, including also more distant relatives than those in the nuclear family.

In this paper we use a modeling framework inspired by Borjas' well-known ethnic capital model (see Borjas, 1992) for intergenerational mobility. However, instead of including the average outcome from the individual's ethnic group, we use the outcome from relatives, or "dynastic human capital". Similarly as in Borjas for ethnic background, the argument is that "the dynasty" acts as an externality in the production of human capital of the next generation, which then depends on both parental inputs and the average of the inputs from the dynastic members in the parental generation, i.e., the siblings, cousins etc. of the parents. The implication is then that if this externality is strong, inequality across generations will persist for a long time and hence convergence between individuals with different human capital backgrounds will be slower than what is predicted using the association in human capital between children and parents.

We combine several Swedish registers with information on educational attainment, labor earnings, and occupation for the period 1968-2009. We use imputed measures of years of schooling as a measure of educational outcome for about 650,000 individuals in the child generation. The Swedish multi-generation register and the fact that the entire Swedish population is included in the data allow us to link dynasties up to siblings and cousins of parents, as well as (through marriage and cohabiting records) their spouses and, in turn, their siblings and cousins and so on.[1] This makes it possible to form dynasties that are genetically linked or linked through marriage/cohabitation.

[1] Note that Hällsten (2014) previously have used the Swedish multigenerational registry to link first and second cousins in the child generation.

By using three different indicators of social status – years of schooling, average earnings between age 30 and 60 as well as an index of occupational status - we are also able to address the concern raised by Clark (2014) that parental earnings or educational attainments, used separately, is not sufficiently informative for predicting long-term persistence. This means that we are able to disentangle the additional persistence coming from including dynasties and from different indicators of social class, respectively.

Our results strongly suggest that there is additional information in the outcomes from the extended family and that only using parental outcomes severely underestimates the long term persistence in human capital outcomes across generations. Only considering parental years of schooling in a regression for child years of schooling yields a coefficient estimate on 0.26. Using the average of the dynastic members' years of schooling in the parental generation moves this estimate by 61 percent to 0.43. If this model is extended to simultaneously taking into account information on schooling, income and social stratification, the estimates increase even further by 20 percent to 0.51. Overall, taking dynastic human capital into account results in an estimate of intergenerational mobility in human capital that is twice as slow as conventional estimates from data including parents and children suggest.

Our paper relates to a number of recent studies (see Clark, 2014; Lindahl et al., 2015; Stuhler, 2014; or Solon, 2015, for an overview) that have shown that the long term intergenerational persistence in human capital is much greater than what can be learned from data covering parents and children only. This study extends the existing literature by, in the same analytical framework, disentangle the effects of generations earlier than the parental generation (vertical influence), close and/or distant relatives that are still active (horizontal influence) as well as different measures of human capital on the estimated intergenerational persistence. As a result of the very large data set, the precision of our estimates are substantially improved compared to those presented in previous studies. Compared to the various papers by Gregory Clark, summarized in Clark (2014), we are also able to actually link people across four generations, not only sharing the same surname. This is of particular importance since a surname is likely to proxy for other characteristics important for human capital accumulation, such as race or ethnicity (Chetty et al., 2014).

The paper proceeds as follows. In Section 2 we discuss the empirical specifications using the standard and extended models incorporating dynastic capital, also in a latent variable framework. In Section 3 we introduce the data set, discuss the construction of variables, and present some descriptive statistics. In Section 4 we present the results of the importance of dynastic human capital and the underlying mechanisms. Section 5 concludes the paper.

2. Empirical Specification

2.1 The Standard Model

The standard way to measure how strongly an outcome is transmitted from parents to children is to use the framework of a simple markovian regression model:

$$y_{jt} = \alpha + \beta y_{jt-1} + \varepsilon_{jt}, \tag{1}$$

where *y* represents the outcome under study, *j* is an index for family including children (t) and parents (*t-1*); possibly also controlling for basic characteristics such as birth cohort and gender. Assuming that we can measure the outcome perfectly for both generations, the resulting estimate of β will be an unbiased measure of the strength of the association between the outcome for parents and children. For instance, if *y* represents the logarithm of lifetime earnings, an estimate of β will be an unbiased estimate of the intergenerational earnings elasticity. If *y* is standardized to have the same variance, an estimate of β can also be interpreted as an intergenerational correlation (see e.g. Black and Devereux, 2010).

There is a large literature that estimates β for different outcomes such as earnings, income, education and social class, using data for many countries and time periods. The key challenge has been to accurately measure the outcome of interest (se e.g. Solon, 1992; Haider and Solon, 2006; Böhlmark and Lindquist, 2006; Nybom and Stuhler, 2016). Measurement error "free" estimates for β typically range between 0.2 and 0.6 for lifetime earnings and years of schooling, depending on the outcome studied and the country of origin of the data set used in the study.

2.2 A Model including Dynastic Human Capital

If the outcomes of dynastic members outside the immediate family are important for a child's success – i.e., if there exist positive group effects from members of the extended family - a random individual born into a high income family is more likely to have superior outcomes than what can be inferred from an estimate of equation (1). Reinterpreting Borjas' ethnic capital model yields the following simple extension of equation (1):

$$y_{jdt} = \alpha' + \beta' y_{jdt-1} + \delta \bar{y}_{dt-1} + \varepsilon'_{jdt}, \qquad (2)$$

where d is a sub-index for dynasty and $\bar{y}_{dt-1}$ is an average of y_{jdt-1} over the members of the dynasty in the parent's generation (parents, siblings of parents, cousins of parents, etc).[2]

Averaging over the members in each dynasty d, we, instead of equation (2), get:

$$\bar{y}_{dt} = \alpha' + (\beta' + \delta)\bar{y}_{dt-1} + \bar{\varepsilon}'_{dt},$$

or

$$\bar{y}_{dt} = \alpha' + \gamma \bar{y}_{dt-1} + \bar{\varepsilon}'_{dt}. \qquad (3)$$

If we estimate (3), and use weights equal to the number of dynasty members, we would get an identical estimate as if we estimate

$$y_{jdt} = \alpha' + \gamma \bar{y}_{dt-1} + \varepsilon'_{jdt} \qquad (3')$$

at the individual level.

β in equation (1) captures the intergenerational persistence in inequality between families (j) whereas γ in equation (3) captures the intergenerational persistence in inequality between dynasties (d), where the latter parameter can be decomposed as the sum of the parts transmitted from parents and the dynasty, respectively.

The parameters β and γ will differ if other members than the parent of the dynasty d have an impact on the outcome of the child in dynasty d.[3] However, the magnitude of the difference between β and γ depends on the importance of within dynasty variation relative to the overall variation in y_{jdt-1}. Borjas (1992) shows that estimating (1), when the true model is (2), results in $plim\ \hat{\beta} = \beta' + \delta(1-k) < \beta' + \delta = \gamma$, where 0<$k$<1 is the fraction of within dynasty variation in y_{idt-1}. Hence, the difference between β and γ is increasing (decreasing) in the degree of within (between) dynasty variation in y_{jdt-1} and β and γ are equal only if there is no within dynasty variation in y_{jdt-1}. Note also that estimating equations (3) and (3')

[2] When we later calculate dynasty averages from the data and estimate equation (2) we do not include the parent among the dynastic member for which we calculate the dynastic average. Hence, we would replace $\bar{y}_{dt-1}$ in equation (2) by $\bar{y}_{jdt-1}$. However, this would still give the same equation (3) as we have now, where $\bar{y}_{dt-1}$ include the parent and all other dynastic member in the parental generation. Hence, this does not matter for any of the reasoning laid out here in section 2

[3] One might compare this to the individual and social returns to education, where the social returns is the sum of the individual returns and external effects within the aggregate unit (see e.g. Acemoglu and Angrist, 2001).

directly is equivalent to regressing equation (1) using a full set of dynasty fixed effects as instruments for y_{jdt-1} (see Solon, 1992).

We also note that it requires stronger assumptions to arrive at consistent estimates of β' and δ than for γ. The reason is that if y_{jdt-1} is measured with error, β' will be biased downwards. Since y_{jdt-1} and $\bar{y}_{dt-1}$ will be positive correlated, δ will be biased upwards. This is still the case if the number of dynastic members is very large, i.e., so that measurement error in $\bar{y}_{dt-1}$ approaches zero. However, if equation (3) is estimated using OLS, the bias in the estimate of γ (i.e., the sum of β' and δ) is approaching zero in the number of dynastic members that $\bar{y}_{dt-1}$ is averaged over.[4] As we will see below, this issue is less problematic when we use several proxy variables as measuring the human capital of the dynasty members.

2.3 Extension to a Model allowing for Mis-measurement of an Underlying Latent Variable

To understand the impact of measurement errors on the estimates in a model with a latent variable, we use the simple model laid out in Clark (2014) (see the Appendix or Clark and Cummins, 2014). This simple model of intergenerational transmission is that for family *j* some outcome *y* evolves as $y_{jt} = x_{jt} + u_{jt}$ and $x_{jt} = bx_{jt-1} + e_{jt}$, where x is some underlying latent variable such as human capital or (as in Clark, 2014) “social status”, which evolves as an AR(1) across generations (*t*); u is a generation specific error in measuring *x* (uncorrelated with *x* and with u for other generations); and *e* is an error term, uncorrelated with *u* and with *x* in the previous generation. We call b the long-term intergenerational persistence in human capital.

We can use this simple model to illustrate what parameters we are estimating in a situation where we allow for multiple (*h*) proxies for *x* and where we recognize that individuals can be grouped into dynasties (*d*). We can then write the model as:

$$y_{hjdt} = \rho_h x_{jdt} + u_{hjdt} \tag{4}$$

$$x_{jdt} = bx_{jdt-1} + e_{jdt} \tag{5}$$

[4] See Borjas (1992, page 141) for the case when the true model is equation (1) but when model (2) is estimated.

By adding the coefficient ρ_h in the proxy equation (4), we allow each *y* to vary in their scale and in how well they proxy for *x*. Let us for illustrative purposes normalize and set $\rho_h = 1$, so that the measurement errors in (4) are classical. In the estimations we will set $\rho_1 = 1$, but allow for $\rho_{h>1} \neq 1$, where h=1 is the proxy for which scale we interpret b.

A regression of y_{jdt} on y_{jdt-1} (i.e., model (1)), for any of the *h* proxies, then gives:

$$\text{plim}\hat{\beta}_{OLS} = \frac{cov(y_{jdt}, y_{jdt-1})}{var(y_{jdt-1})} = b\theta_h \leq b$$

where $\theta_h = \frac{var(x_{jdt})}{var(y_{hjdt})}$ assuming variances are constant across *t*'s. This is the standard formula for measurement error bias.[5] By instead taking averages of *y* over *h* in this regression we have that the variance of the average of *y* (across *h*'s) approaches the variance of *x* (since $var(\bar{u}_{jdt}) < var(u_{hjdt})$ where we note that now $\theta = \frac{var(x_{jdt})}{var(\bar{y}_{hjdt})}$), so that $\theta \to 1$ and $\hat{\beta}_{OLS} \to b$. Taking averages over multiple measures is explicitly stated as a solution to arrive at an unbiased estimate of *b* in Clark and Cummins (2014), even though it has yet to be tested empirically.

Clark (2014) argues that by taking group averages (they use surnames) one can eliminate the importance of generation specific effects because the average of u_{hjdt}, for a given *h*, will then approaches zero, and hence $\theta \to 1$. However, as pointed out by Chetty et al. (2014: Online Appendix, p. 9), using averages over any group will estimate intergenerational persistence between groups, which is a different parameter than the intergenerational persistence between families.

In this paper we deal with this difference in three different ways. First, we average over individuals in a group, which, by definition, only includes individuals in the immediate and extended family, the dynasty (instead of using proxies such as surnames or other groupings that are less relevant for our purpose). Second, we utilize three different outcomes: educational attainments, earnings and the score from the so called CAMSIS index for occupation-based social stratification (all based on data from high quality registers). The simplest approach would be to just take averages over these outcomes. However, since they are measured on different scales as is likely to vary in how well they proxy for *x*, we instead

[5] Note that if we allow for $\rho_h \neq 1$, we get that $\text{plim}\hat{\beta}_{OLS} = \rho_h{}^2 b\theta_h$ which can be greater than b, unless $\rho_h \leq 1$.

use the more efficient method suggested by Lubotsky and Wittenberg (2006) (see below). Third, we explicitly recognize that we then estimate different parameters.

Just as in equation (2) we now extend this model by allowing for individuals in generation t-1, other than the parent, to influence the human capital of the child. This gives a modified equation (5) as

$$x_{jdt} = b' x_{jdt-1} + c\bar{x}_{dt-1} + e_{jdt}, \qquad (6)$$

which is an hybrid of the model of transmission of ethnic capital between generations in Borjas (1992), adapted to dynastic capital, and the modelling framework in Clark (2014).

Taking averages over individuals within dynasties we can rewrite (4) and (6) as:

$$\bar{y}_{hdt} = \rho_h \bar{x}_{dt} + \bar{u}_{hdt} \qquad (7)$$

$$\bar{x}_{dt} = (b' + c)\bar{x}_{dt-1} + \bar{e}_{dt} \qquad (8)$$

A regression of $\bar{y}_{dt}$ on $\bar{y}_{dt-1}$ (i.e., equation (3)) for any of the h's (again normalizing ρ_h to one) then gives:

$$\text{plim } \hat{\gamma}_{OLS} = \frac{cov(\bar{y}_{dt}, \bar{y}_{dt-1})}{var(\bar{y}_{dt-1})} = (b' + c)\theta' \leq b' + c$$

where $\theta' = \frac{v(\bar{x}_{dt})}{v(\bar{y}_{dt})}$ assuming all variances are constant across time.

By taking averages of *y* over *h* in this regression we have again that $\theta' \rightarrow 1$ and that $\hat{\gamma}_{OLS} \rightarrow b' + c$. Note that for small d, $\theta' < 1$ is likely but that we expect $\theta < \theta'$ since we already averaged across dynastic members.

3. Data and Descriptive Statistics

3.1 Data and key variables

Our data set is compiled from different Swedish registers using the individual identification number. The Swedish Multi-generation register, covering the full population, enables us to link parents to children for all those children born 1932 or later, provided that the child and the parents have been registered as living in Sweden at some point after 1961. We have been able to link dynasties up to siblings and cousins of parents, as well as (through marriage and cohabiting records) their spouses and their siblings and cousins and so on. This makes it possible to (infinitely) form dynasties that are genetically linked and/or linked through assortative mating.

To maximize the number of dynasty members we use predicted years of schooling instead of actual ones, by using information on 9th grade GPA, an academic high school track indicator (and gender and birth year) for a sample of individuals at least 30 years of age. This is our main outcome variable for the child generation. Using actual years of schooling give qualitatively similar results, but a much smaller sample size. The final estimation sample covers close to 650,000 children.

We further compile data from registers (and censuses for earlier years) that contain information on education, income and occupation. The education information is available in the 1970 census and in yearly registers between 1985 and 2009. Income data is drawn from tax registers and is available for the years 1968, 1971, 1973, 1976, 1979, 1982, and every year between 1985 and 2009. Occupation information is available from censuses every fifth year between 1970 and 1990. To be included in the estimations we therefore also implicitly require that at least one parent and one cousin to parent have to have survived and still worked in 1970.

For the parental and other ancestor generations we use, in addition to years of schooling, two outcomes: family income and the so called CAMSIS index for occupation-based social stratification. We calculate average log income in the following way: We use data for all available years for each individual between ages 30-60; we take logs and residualize by adjusting for both cohort and year fixed effects; we then take the average of the residuals (and add the constant) for each individual.

The CAMSIS measure of social distance uses occupations of spouses to create social distance table (Prandy and Lambert, 2003). Hence, the closer is the occupation of marriage partners, the higher is the CAMSIS score. It is based on estimation of occupational scores (0-

100) that best fit the observed data and is a commonly used measure of occupational status in sociology.[6]

The main independent variables are constructed by taking averages of non-missing observations within each category of relatives (parents, grandparents, aunts and uncles, etc.) For example, if for one child we observe years of schooling for three of their four grandparents, the grandparental years of schooling variable will be the average of those three, excluding the fourth. To construct the dynastic variables, we then average these group averages in the same way. This ensures that each *category* of relatives is given the same relative weight in the dynasty variables regardless of how many individuals we observe in that category. If we instead were to construct the dynasty variables by averaging directly across all relatives, we would implicitly be giving a disproportionate large weight to, e.g., parents' cousins, simply because the average family has many of them.

Table 1 shows descriptive statistics for the data set. The first three columns show average number of years of schooling, average residualized log family income and the CAMSIS index by category in the dynasties. The fourth column shows the average number of observation used for calculating the averages corresponding to the category in the dynasty. In effect, we only require one non-missing observation for each category of relatives in order for a child to be included in the main regressions.[7] The figures show that we have data on almost all grandparents (3.869) and about half of the great grandparents (4.107). The fifth column shows average birth years and, finally, the last column shows the number of observations we have for each category.

The standard deviation for years of schooling for great grandparents is very low. This is because a high fraction (about 80-85%) of the individuals in that generation only had primary school education. It is therefore very beneficial to be able to measure income and occupation for these individuals, so as to incorporate variation in socioeconomic status across the whole distribution.

---

[6] It also takes so-called "pseudo-diagonals" into account (e.g., doctor-nurse couples). See http://www.camsis.stir.ac.uk/Data/Sweden90.html. The Swedish scores are estimated by Lambert and Bihagen (2007).

[7] As assortative mating is high in Sweden, this is unlikely to not pose any serious problems in intergenerational estimations (Holmlund, 2008).

Table 1. Summary statistics.

	Years of Schooling	Log Income (residualzied)	Social stratification	Observations/ child	Birth year	Observations
Child	12.46 (1.40)				1988.15 (4.76)	647250
Parents	11.65 (1.70)	-0.04 (0.50)	46.95 (9.82)	1.99 (0.07)	1960.95 (4.70)	647250
Aunts/uncles	11.69 (1.46)	-0.06 (0.41)	46.74 (8.48)	4.51 (2.34)	1960.98 (4.76)	647250
Spouses of aunts/uncles	11.82 (1.53)	-0.01 (0.40)	47.32 (9.01)	3.73 (2.04)	1961.24 (6.39)	639303
Parents' Cousins	12.17 (1.20)	-0.02 (0.35)	46.01 (7.29)	10.20 (7.67)	1967.92 (4.06)	647250
Grandparents	9.29 (1.72)	-0.15 (0.38)	45.85 (7.74)	3.87 (0.43)	1934.20 (5.81)	647233
Great grandparents	7.33 (0.80)	-0.20 (0.63)	40.92 (8.62)	4.12 (2.01)	1908.57 (5.81)	645109
Parents' aunts/uncles	9.92 (1.84)	-0.14 (0.40)	46.89 (8.26)	5.22 (3.67)	1942.16 (4.29)	646995
Dynasty	11.83 (1.05)	-0.04 (0.26)	46.74 (5.45)		1962.78 (3.88)	647250

Means with standard deviations in parentheses.

Tables 2a and 2b show correlations between the three main variables years of schooling, average residualized log family income and the CAMSIS index. In Table 2a, where we use the parent as the unit of observation, we observe the highest correlation between years of schooling and social stratification, whereas the two correlations with log income are smaller. Although these three variables clearly contain common information, they certainly also capture different things. In Table 2b, where we use the dynasty as the unit of observation, the pattern is similar although all three correlations increase.

Table 2a. Correlation matrix. Parents.

	Years of schooling	Log income	Social stratification
Years of schooling	1		
Log income	0.300	1	
Social stratification	0.521	0.306	1

Table 2b. Correlation matrix. Dynasties.

	Years of schooling	Log income	Social stratification
Years of schooling	1		
Log income	0.426	1	
Social stratification	0.618	0.392	1

3.2 Incorporating multiple proxies

Multiple proxies for the human capital of dynasty members can be incorporated in several ways. The simplest one is to take the average of the standardized variables. This weights the contribution of each variable equally. An alternative is to weight each variable by its contribution in explaining the child's outcome. To do this, we use a framework proposed by Lubotsky and Wittenberg (2006).

We can view the three outcome measures for the different parts of the family as a set of proxy variables for a single latent variable that is transmitted across generations. If the true model we want to estimate is $q = bx + e$, where q is the outcome variable (child's years of schooling in the estimations below) and *x* is a latent independent variable only observed through a set of *H* proxy variables, i.e., $y_h = \rho_h x + u_h$, where *h=1, 2……H*. Lubotsky and Wittenberg (2006) show that the most efficient way to use the information in the proxies is to estimate the regression $q = \sum_{h=1}^{H} \varphi_h y_h + v$, and then take the weighted average $\beta^* = \sum_{h=1}^{H} \frac{cov(q,y_h)}{cov(q,y_1)} \varphi_h$ which is used as an estimate of b. Note that for y_1, which can be any of the proxy variables, we need to normalize and set $\rho_1 = 1$. In our main results, we use parents' schooling as y_1, so that all weighted average coefficients can be interpreted relative to the coefficient on parents' schooling. Note that the key assumption is that each y_h have no impact on q, conditional on x.

Also note that if $\rho_h = 1$ for all *h*, we have that $\beta^* = \sum_{h=1}^{H} \varphi_h$. This means that we revert back to a standard measurement error model where each proxy is scaled in the same way and where we have assumed that the covariance between *y* and each proxy is identical. As we will see below, this is important for the interpretation of the coefficient estimates in our multivariate regression estimations, where we include an outcome for each relative separately in a regression, as the sum of these estimates then can be interpreted as an estimate of b.

4. Results

4.1 The Importance of Dynastic Human Capital

Table 3 shows the first set of results. As for all results shown in this paper, the dependent variable is years of schooling predicted from ninth grade GPA scores and choice of high school academic track for the individual in the child generation. The table contains three panels. Panel A shows the results when we use years of schooling in the independent variables; Panel B when we use income; and Panel C when we use the CAMSIS occupational social stratification index. Each panel shows the results from the same specifications. The estimates shown in the first column corresponds to Equation (1) in Section 2 and those in the fourth column to Equation (2). The result in Column (2) corresponds to Equation (3') and that in Column (3) where we have, following the convention in estimation of group effects, excluded the child's parents from the calculation of the dynasty's educational attainments.

The first column in Panel A shows that one year of additional parental education is reflected in 0.264 extra years of schooling for the child. This result is very similar to what has typically been obtained in previous studies on Swedish data (see e.g. Björklund et al., 2006). Turning to the second column, we see that an additional year of schooling for the average of the dynasty is associated with 0.426 extra schooling for the child. Comparing the result in the first column with that in the second, following the discussion of Equations (1) and (3'), we can reject that schooling within the dynasty, in addition to the parents, has no effect on the child's educational attainments. This is confirmed by the results in Column (3) – where we estimate the separate group effect from the dynasty educational attainments – and in Column (4), where we include both independent variables.[8]

[8] Note that the decrease in R^2 between columns 1 and 2 is because in column 2, we only use the between dynasty variation in the outcomes.

Table 3. Results from OLS regressions of child's predicted years of schooling on three different measures of parental and dynastic human capital: Years of Schooling, Income and Occupational Status Index.

Panel A: Years of Schooling

	(1)	(2)	(3)	(4)
Parents' schooling	0.264			0.218
	(0.001)			(0.001)
Dynasty schooling		0.426		
		(0.002)		
Dynasty schooling, excl parents			0.333	0.172
			(0.002)	(0.002)
R^2	0.145	0.142	0.105	0.157
N	647,250	647,250	647,250	647,250

Panel B: Income

	(1)	(2)	(3)	(4)
Parents' income	0.643			0.566
	(0.003)			(0.004)
Dynasty income		1.267		
		(0.006)		
Dynasty income, excl parents			0.853	0.573
			(0.007)	(0.007)
R^2	0.100	0.101	0.072	0.110
N	647,250	647,250	647,250	647,250

P**anel C: Occupation Status**

	(1)	(2)	(3)	(4)
Parents' occupation	0.036			0.031
	(0.000)			(0.000)
Dynasty occupation		0.067		
		(0.000)		
Dynasty occupation, excl parents			0.047	0.031
			(0.000)	(0.000)
R^2	0.108	0.111	0.081	0.121
N	647,250	647,250	647,250	647,250

Note: Robust standard errors in parentheses.

By comparing the results in Columns 1 and 2 one can get an estimate of the magnitude of the downward bias in the persistence of inequality across generations induced by ignoring the dynasty effects. This comparison suggests that we underestimate the effect by 38 percent. This is very similar to the share of the sum of the coefficients in Column 4, which suggests 44 percent can be attributed to the dynasty. Hence, dynastic human capital appears only slightly less important than parents' human capital.

The results for income and occupational status, shown in Panels B and C, respectively, are qualitatively the same as those for education shown in Panel A. However, the magnitude of the share attributed to the dynasty is substantially larger for both and, between themselves, remarkably similar in magnitude. Both suggest that about 50 percent of the persistence can be attributed to the dynasty. However, it should be noted that measurement bias in parent's outcome likely result in (as discussed in Section 3) an overestimate of the dynastic human capital contribution. On the other hand, incorporating all human capital proxies simultaneously might lead to an increase in the overall contribution of the dynasty.

Table 4 shows the results when we add *years of schooling* and *earnings* to the same specification and Table 5 where we, in addition to this, have added the *index for occupational social stratification*. The lower panels of these tables show the results when we have used the Lubotsky-Wittenberg method for combining indicators.

From the results in Table 4 it is evident that adding *income* significantly contributes to the estimated persistence. Comparing the estimate for parents in Column 1 of the lower panels, we see that the estimate increases from 0.264 to 0.322, when we add income, and to 0.344, when we also add occupation in Table 5, which corresponds to a 22 percent increase. This increase is likely driven by both classical measurement errors in years of schooling and by years of schooling being an imperfect proxy for human capital.

If we turn to the estimates for the dynasty in Column 2 of the lower panels, we see that the estimate increases from 0.426 to 0.489, when we add dynastic income, and to 0.512, when we also add dynastic occupation. In column 4 we see that this increase is driven by the contribution of the parents, since the dynasty parts (excluding parents) is pretty much unchanged compared to the estimate from only using years of schooling. This result supports the claim that bias from measurement error is much more of a problem when estimating the association between outcomes for children and parents, compared to the association between children's and dynasty averages.

To conclude, we arrive at a decomposition of the dynasty transmission of inequality in human capital that to about two-thirds is due to the parents and about one third due to the dynasty. Overall, a child that is born into a dynasty where the members have on average one additional year of "true" years of schooling will have, on average, half a year more education. Adding more than one proxy for social status also significantly contributes to the estimated persistence across generations.

Table 4. Results from OLS regressions of child's predicted years of schooling on parent and dynasty schooling and income.

	(1)	(2)	(3)	(4)
Parents' schooling	0.228			0.190
	(0.001)			(0.001)
Parents' income	0.433			0.385
	(0.003)			(0.003)
Dynasty schooling		0.354		
		(0.002)		
Dynasty income		0.694		
		(0.007)		
Dynasty schooling, excl parents			0.293	0.133
			(0.002)	(0.002)
Dynasty income, excl parents			0.418	0.214
			(0.007)	(0.007)
Parents	0.322			0.274
	(0.001)			(0.001)
Dynasty		0.489		
		(0.002)		
Dynastic excl par			0.366	0.171
			(0.002)	(0.002)
R^2	0.166	0.155	0.109	0.177
N	647,250	647,250	647,250	647,250

Note: Standard errors in parentheses. Dependent variable is child's predicted years of schooling

Table 5. Results from OLS regressions of child's predicted years of schooling on parent and dynasty schooling, income and CAMSIS index for occupational status.

	(1)	(2)	(3)	(4)
Parents' schooling	0.193			0.163
	(0.001)			(0.001)
Parents' income	0.392			0.354
	(0.003)			(0.003)
Parents' occupation	0.013			0.011
	(0.000)			(0.000)
Dynasty schooling		0.299		
		(0.002)		
Dynasty income		0.641		
		(0.007)		
Dynasty occupation		0.018		
		(0.000)		
Dynasty schooling, excl parents			0.253	0.107
			(0.002)	(0.002)
Dynasty income, excl parents			0.382	0.193
			(0.007)	(0.007)
Dynasty occupation, excl parents			0.014	0.005
			(0.000)	(0.000)
Parents	0.344			0.294
	(0.001)			(0.001)
Dynasty		0.512		
		(0.002)		
Dynastic excl par			0.386	0.166
			(0.002)	(0.002)
R^2	0.172	0.157	0.111	0.181
N	647250	647250	647250	647250

Note: Standard errors in parentheses. Dependent variable is child's predicted years of schooling

4.2 Decomposing the Dynasty Effect

In Tables 6a-c we disentangle the importance of different parts of the extended family for the education outcome in the fourth generation, using schooling, income and occupational rank, respectively, as proxy variables for the human capital of the ancestors. Column 1 shows the results from the first column in Table 3, Panel A, as a reference. We first include aunts/uncles, spouses of aunts/uncles and parents' cousins – horizontally related to the parents (Columns 2-4). We then extend the baseline specification by including outcomes from generations preceding the parental one, i.e., grandparents and great grandparents – vertically related to the parents (Columns 5-6). Finally, we add outcomes for all dynasty members (Column 7).

Table 6a. Results from regression models including successively additional components of the dynasties. Schooling.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents	0.264	0.225	0.220	0.214	0.241	0.239	0.207
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Aunts and uncles		0.103	0.090	0.083			0.079
		(0.001)	(0.001)	(0.001)			(0.001)
Spouses of aunts/uncles			0.038	0.035			0.031
			(0.001)	(0.001)			(0.001)
Parents' cousins				0.057			0.045
				(0.001)			(0.002)
Grandparents					0.048	0.048	0.015
					(0.001)	(0.001)	(0.001)
Great grandparents						0.003	-0.011
						(0.002)	(0.002)
Parents' aunts/uncles							0.004
							(0.001)
Sum of coefficient estimates	0.264	0.328	0.348	0.389	0.289	0.290	0.370
Total effect, LW method	0.264	0.285	0.289	0.295	0.263	0.261	0.287
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Only parents	0.264	0.225	0.220	0.214	0.241	0.239	0.207
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Excluding parents		0.060	0.069	0.081	0.022	0.021	0.080
		(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
R^2	0.145	0.154	0.156	0.158	0.150	0.151	0.160
N	647,250	647,250	647,250	647,250	647,250	647,250	647,250

Standard errors in parentheses
Dependent variable is child's predicted years of schooling

The results shown in Table 6a reveal highly significant estimates for all parts of the dynasty sequentially added to the baseline specification. The magnitude of the aunt/uncle coefficient is almost half the size of the estimate for parents. It can also be noted that it is well above twice as large as the corresponding one for the spouses of aunts and uncles, which in turn is smaller than the estimate for parents' cousins, indicating the importance of genetic proximity for this outcome.[9]

[9] These results can be compared to Jaeger (2012) that used data for the US on almost 17,000 children (the Wisconsin Longitudinal Survey) and regressed models of child's years of schooling on parent's education, SES

Taken together, the results in Column 4, confirming the one shown in Column 4 of Panel A of Table 3, indicate that changing all the extended family members' years of schooling by one year on average change the child's predicted years of schooling by about 0.4 years. Almost 60 percent of this effect (0.214/0.389) can be attributed to the change in parent's years of schooling and 40 percent to other dynasty members in the parental generation.[10]

Turning to the results on the association between vertically related relatives, it is evident from Column 5 that our results supports those obtained in Lindahl et al. (2015) that the traditional AR(1) model can be rejected as a description of intergenerational association in educational attainments in the sense that the coefficient estimate for grandparents turned out significantly different from zero. The results in Column 5 show that education for great-grandparents education is only marginally significant in the AR(3) model, even if the precision is quite good in this very large sample.

Column 7 shows the results when all dynasty members under study are included in the specification. As expected, the coefficient estimates decrease with both the vertical and the horizontal distance to the individual. Years of schooling of aunts and uncles have a strong effect, almost 40 percent of the corresponding effects of parents, on the child's years of schooling.

The lower panel of Table 6a reports the total effect of dynastic members' schooling measured by the sum of the coefficients for the dynasty members as well as the by using the Lubotsky-Wittenberg (LW) method. As explained in Section 3.2, although the numerical difference between these two measures is simply due to differences in weighting of the underlying coefficients, the interpretation of the results from each of them are very different. The coefficient sum measures the association between the child's schooling and that of a randomly chosen dynasty member, i.e. the latent schooling for the dynasty. The LW estimate

---

and income, as well as aunts and uncles education, SES and income. They found that, conditional on parents' outcomes, only aunts and uncles education were statistically significantly associated with child's years of schooling. The coefficient estimates were less than one-third of the ones for parents. When they added controls for grandparent's education, SES and income, these variables were statistically insignificant and hence the results for aunts and uncles were unchanged.

[10] This result is naturally very much in line with what we found in column 4 of Panel A of Table 3, where parents' schooling had a coefficient of 0.218 and dynastic schooling of 0.172, summing to 0.390, since the dynastic average used in Table 3 was calculated based on aunts and uncles, spouses of aunts and uncles and parent's cousins schooling.

measures the association between the child's schooling and the latent schooling of his or her parents as proxied by schooling of the other dynasty members.[11]

As expected from the reasoning above, the coefficient sum and the LW estimates, the total effect is numerically smaller than the sum of the estimates. The quite large difference between the estimates – 0.370 versus 0.287 – suggests a strong correlation in educational attainments between the dynasty members.

We now turn to results from using the two other proxies for human capital of ancestors. Results are presented in Table 6b (income) and Table 6c (occupation). An interesting difference compared to what we observed for schooling is that, for both income and occupation, there is a positive and statistically significant association between great grandparents and great grandchildren in the AR(3) models.[12] We also see stronger estimate for grandparents: about one-half in relation to the estimate for parents (Column 5) for income and occupation, whereas for schooling the estimate is about one-fifth of the size.

The background to this difference is likely to be the weakness of the schooling measure for grandparents and great grandparents. Most of these individuals finished education in a time when quite few people attained schooling levels above primary school. Hence, there is very little variation for a large part of the population, as opposed to the income and occupation measures.

[11] For the LW approach we have set $\rho_k = 1$ for parents schooling (but allow $\rho_k \neq 1$ for the other relatives) and hence, assumed that schooling of the parents is the best proxy for the latent parental schooling variable, each of the other relatives will get lower weights, where the weight of each relative is determined by the strength in the association with child's schooling. If we instead set ρ_k=1 for all proxies, we say that each relative is an equal proxy for the latent variable, (which then would be a representative relative). So if we are interested in the association between child's schooling and a representative dynastic member (so latent schooling for the dynasty), we should not weight, whereas if we are interested in the association between child's schooling and the schooling of dynastic member k (so latent schooling for dynastic member k), we should (as it is more efficient..) use the LW approach and hence set $\rho_k = 1$.

[12] Using data on occupation over three generations, Long and Ferrie (2013) for the US and Braun and Stuhler (2015) for Germany, find grandparent's occupation to be associated with grandchild's occupation, conditional on parent's occupation. Braun and Stuhler do however find evidence against that this grandparental effect can be given a causal interpretation.

Table 6b. Results from regression models including successively additional components of the dynasties. Income.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents	0.643	0.582	0.569	0.561	0.587	0.581	0.517
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Aunts and uncles		0.306	0.277	0.267			0.221
		(0.004)	(0.004)	(0.004)			(0.004)
Spouses of aunts/uncles			0.143	0.138			0.118
			(0.004)	(0.004)			(0.004)
Parents' cousins				0.164			0.108
				(0.005)			(0.005)
Grandparents					0.288	0.285	0.208
					(0.005)	(0.005)	(0.005)
Great grandparents						0.016	0.005
						(0.003)	(0.003)
Parents' aunts/uncles							0.091
							(0.004)
Sumo of estimates	0.643	0.888	0.989	1.129	0.875	0.882	1.268
Total effect	0.643	0.736	0.757	0.778	0.712	0.706	0.805
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Only parents	0.643	0.582	0.569	0.561	0.587	0.581	0.517
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Excluding parents		0.154	0.187	0.217	0.125	0.125	0.288
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
R^2	0.100	0.107	0.109	0.111	0.109	0.111	0.121
N	647250	647250	647250	647250	647250	647250	647250

Standard errors in parentheses
Dependent variable is child's predicted years of schooling

Table 6c. Results from regression models including successively additional components of the dynasties. Occupational status.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents	0.036	0.032	0.031	0.030	0.031	0.031	0.027
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Aunts and uncles		0.016	0.015	0.014			0.011
		(0.000)	(0.000)	(0.000)			(0.000)
Spouses of aunts/uncles			0.010	0.010			0.008
			(0.000)	(0.000)			(0.000)
Parents' cousins				0.007			0.004
				(0.000)			(0.000)
Grandparents					0.017	0.017	0.011
					(0.000)	(0.000)	(0.000)
Great grandparents						0.001	-0.000
						(0.000)	(0.000)
Parents' aunts/uncles							0.005
							(0.000)

Sum of estimates	0.036	0.048	0.056	0.061	0.048	0.049	0.066
Total effect	0.036	0.041	0.044	0.045	0.039	0.039	0.045
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Only parents	0.036	0.032	0.031	0.030	0.031	0.031	0.027
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Excluding parents		0.009	0.013	0.014	0.008	0.008	0.017
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.108	0.117	0.121	0.122	0.120	0.121	0.131
N	647,250	647,250	647,250	647,250	647,250	647,250	647,250

Standard errors in parentheses
Dependent variable is child's predicted years of schooling

In Appendix Table A1, we also show results from models where we have included schooling, income and occupation simultaneously. Although all three variables still contribute, we see that the coefficient estimate for *Occupation* decreases the most when we condition on the other human capital proxies. The proxies for great grandparents are no longer positive, when we condition on the three proxies for parents and grandparents. However, the proxies for grandparents are still positive and highly significant.

When we estimate these models using the LW-method (and interpret the size of the association in terms of years of schooling, i.e., we set $\rho_k = 1$ for schooling) we find that the total effect increases by 25% to 0.370, which is entirely driven by the increased contribution from parents. This is in line with the measurement error story discussed in Section 4.1. Finally, we decompose the separate contributions of *Schooling*, *Income* and *Occupation*. This exercise shows that all three proxies are important for the total effect, but that *Schooling* explains about 55%, *Income* 26% and *Occupation* 19%.

4.3. Upper Bound Estimates of Intergenerational persistence using IV

The traditional AR(1) model will give a consistent estimate of the long term association in human capital across generations under two key assumptions. (1) All influence for human capital accumulation is transmitted through the parents. This means that the process of human capital transmission have a memory of only one generation and that no other relatives in the parental generation have influence on the individual outcome. (2) The measure of human capital used for the parental generation is exhaustive for the outcome in the child generation. That is, as pointed out by e.g. Clark (2014), no dimensions of the human capital concept are ignored by the measure in use.

If the first assumption is maintained, we would be able to relax the second one by using the outcomes of previous generations and/or those of other relatives in the parental generation as instrumental variables. The resulting parameter estimate can be interpreted as the long term intergenerational persistence purged from potential attenuation bias from errors of measurement. The estimate can be seen as an upper bound of the true intergenerational persistence in the sense that if the exclusion restrictions do not hold, it will result in an upward asymptotic bias.[13]

If the exclusion restriction do not hold we implicitly reject the AR(1) model and admit and influence of dynastic human capital. This means that we can form a *joint* hypothesis for a Hausman test for the IV model against traditional OLS model for intergenerational persistence in human capital: if we reject the OLS model it could either be that parental human capital is not adequately measured *or* that the exclusion restrictions do not hold and there is a direct influence of the dynasty outcomes. Either way, the traditional model will give estimates with a downward asymptotic bias.

We present the IV results for *Schooling* in Table 7. The corresponding results *Income* and *Occupation*, which are qualitatively very similar to those shown in Table 7, are presented in the Appendix in Tables A2 and A3. The upper panel of Table 7 shows the results from the first stage. The schooling associations for the relatives are jointly highly significant, but that the contribution to the explained variance in the schooling of parents vary greatly across relatives. Note that the aunts and uncles are the siblings of the parents, so that in Column 1, we report a sibling correlation of about 0.5 (which is similar to other estimates in the literature, see Björklund and Jäntti, 2011).

The second stage results are reported in the lower panel. Using relatives in the parental generation (sibling, cousins) generate IV estimates above 0.410, whereas using ancestors (parents and grandparents) generate slightly tighter upper bounds of 0.380. Although, as shows by the *P*-values reported from the Hausman tests, we can reject the estimates from the traditional OLS model from Table 3 in all specifications, we conclude that these are lower than the ones reported in Lindahl et al. (2015) and much below the estimates of about 0.7 reported in Clark (2014).

[13] As long as the outcomes for these other relatives have a positive impact on grandchild's outcome, conditional on parent's outcome.

Table 7: Schooling IV regressions. First stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Aunts and uncles	0.502	0.427	0.399			0.304
	(0.001)	(0.002)	(0.002)			(0.002)
Spouses of aunts/uncles		0.155	0.144			0.103
		(0.002)	(0.002)			(0.001)
Parents' cousins			0.176			0.087
			(0.002)			(0.002)
Grandparents				0.404	0.393	0.228
				(0.001)	(0.001)	(0.001)
Great grandparents					0.055	-0.000
					(0.003)	(0.003)
Parents' aunts/uncles						0.038
						(0.001)
R^2	0.240	0.255	0.268	0.223	0.226	0.315
N	647,250	647,250	647,250	647,250	647,250	647,250

Standard errors in parentheses. Dependent variable: Parents' average years of schooling.

Schooling IV regressions. Second stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Parents	0.410	0.434	0.441	0.374	0.380	0.418
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Hausman test, P-value	0.000	0.000	0.000	0.000	0.000	0.000
R^2	0.115	0.104	0.101	0.128	0.126	0.112
N	647,250	647,250	647,250	647,250	647,250	647,250

Standard errors in parentheses. Dependent variable: Child's predicted years of schooling.

5. Conclusions

The results obtained in this study unambiguously suggest that it is not sufficient to use data on only two consecutive generations when studying intergenerational persistence in human capital outcomes. Results from such studies underestimate the true long term persistence. We investigate two potential mechanisms behind this result: (i) that two generation studies ignore the association to other close relatives than the parents - both horizontally and vertically related; (ii) that traditional studies are restricted to only one measure of human capital outcome in the parental generation.

To study the first mechanism, we obtain two sets of results: (a) from models where we add an additional component measuring the average human capital of the dynasty; (b) from models where we add separate measures of human capital outcomes for the different parts of the dynasty under study. Our results from the first set of results suggest that we miss around 40 percent of the persistence when omitting the educational attainment outcomes of the dynasty. If we use income or the social stratification measure the downward bias is much larger. The most interesting contribution from the second set of results is that the most important part to the omitted persistence is attributed to the relatives horizontally related to the parents, i.e., aunts/uncles and cousins of the parents.

To study the second mechanism we obtain three sets of results: (a) from models where we combine human capital outcomes from the dynasty using the Lubotsky-Wittenberg (LW) method; (b) from models where we combine the results from the three different human capital outcome measures using the LW method; (c) from IV models where we use human capital outcomes from the dynasty as instrumental variables.

All three sets of results show significant increases in the intergeneration persistence compared to those obtained from the traditional AR(1) model. In the first set of results we estimate an increased persistence from 0.26 to 0.29 (9 percent) when we use the educational outcomes from all dynasty members as indicators in the LW procedure. The second set of results show that the intergenerational persistence estimate increase from 0.26 to 0.34 (or 30 percent), when we include income and the social stratification index in addition to educational attainment in the LW procedure. The dynasty persistence estimate increases from 0.43 to 0.51 (or 30 percent). Finally, the IV model when including educational attainments of all dynasty members as instrumental variables gives an estimate on 0.42, which is a 58 percent increase.

The 0.51 estimate of intergenerational persistence, presented in Table 5 and obtained using dynasty means and the LW procedure, is, as expected, the estimate reflecting the strongest intergenerational persistence. This estimate is interesting, not only because it indicates a substantial downward bias of traditional intergenerational persistence estimates, but also as it is substantially lower than the influential estimate if Clark (2014) between 0.70 and 0.80.[14]

Our results contradicts the interpretation of the Becker and Tomes classical results of a weak intergenerational persistence of economic advantage and inequality, but supports the general message of the Becker-Tomes model of the family as an important engine in formation of human capital. The most novel contribution of our paper *vis-à-vis* the previous empirical literature on intergenerational mobility is that we show the importance of group effects in this

[14] Using an alternative approach and data on occupation and education for Germany, Braun and Stuhler (2015) also find that the latent variable "social status" is transmitted much more persistent across generations than suggested by the parent-child correlation, but not as high as suggested in Clark (2014).

research area. The results show that there are important externalities of human capital formation within the extended family, or the dynasty as we label it in this paper. Although we find a weaker long term intergenerational persistence than suggested by Clark (2014), our main results supports the general framework of Clark in the sense that they highlight the fact that it is not sufficient to look at the outcomes of the parents only when studying intergenerational persistence in socio-economic positions.

Appendix

Table A1: Results from regression models including successively additional components of the dynasties., Schooling, Income, and Occupational status.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Schooling							
Parents	0.193	0.166	0.164	0.160	0.178	0.177	0.156
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Aunts and uncles		0.065	0.060	0.055			0.054
		(0.001)	(0.001)	(0.002)			(0.002)
Spouses of aunts/uncles			0.017	0.015			0.013
			(0.001)	(0.001)			(0.001)
Parents' cousins				0.041			0.031
				(0.002)			(0.002)
Grandparents					0.018	0.018	-0.001
					(0.001)	(0.001)	(0.001)
Great grandparents						-0.002	-0.011
						(0.002)	(0.002)
Parents' aunts/uncles							0.000
							(0.001)
Income							
Parents	0.392	0.358	0.354	0.350	0.381	0.379	0.346
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Aunts and uncles		0.126	0.117	0.111			0.107
		(0.004)	(0.004)	(0.004)			(0.004)
Spouses of aunts/uncles			0.025	0.023			0.027
			(0.004)	(0.004)			(0.004)
Parents' cousins				0.057			0.051
				(0.005)			(0.005)
Grandparents					0.040	0.042	0.004
					(0.005)	(0.005)	(0.005)
Great grandparents						-0.010	-0.011
						(0.003)	(0.003)
Parents' aunts/uncles							0.011
							(0.005)
Social stratification							
Parents	0.013	0.011	0.011	0.011	0.011	0.011	0.010
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Aunts and uncles		0.004	0.002	0.002			0.002
		(0.000)	(0.000)	(0.000)			(0.000)
Spouses of aunts/uncles			0.003	0.003			0.002
			(0.000)	(0.000)			(0.000)
Parents' cousins				0.000			0.000
				(0.000)			(0.000)
Grandparents					0.005	0.005	0.002
					(0.000)	(0.000)	(0.000)
Great grandparents						0.000	-0.000
						(0.000)	(0.000)
Parents' aunts/uncles							0.001
							(0.000)

Total effect	0.344	0.362	0.366	0.370	0.343	0.340	0.363
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Only parents	0.344	0.300	0.295	0.289	0.320	0.318	0.283
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Excluding parents		0.062	0.070	0.081	0.023	0.023	0.080
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Schooling	0.193	0.204	0.206	0.210	0.186	0.185	0.201
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Income	0.086	0.092	0.092	0.093	0.087	0.087	0.093
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Social stratification	0.065	0.066	0.067	0.066	0.069	0.069	0.069
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R2	0.172	0.180	0.181	0.182	0.176	0.176	0.184
Obs	647250	647250	647250	647250	647250	647250	647250

Standard errors in parentheses
Dependent variable is child's predicted years of schooling

Table A2: IV regressions. Income. First stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Aunts and uncles	0.296	0.265	0.257			0.215
	(0.002)	(0.002)	(0.002)			(0.002)
Spouses of aunts/uncles		0.126	0.123			0.101
		(0.002)	(0.002)			(0.002)
Parents' cousins			0.095			0.062
			(0.002)			(0.002)
Grandparents				0.285	0.281	0.204
				(0.002)	(0.002)	(0.002)
Great grandparents					0.014	0.005
					(0.001)	(0.001)
Parents' aunts/uncles						0.050
						(0.002)
r2	0.080	0.089	0.094	0.070	0.073	0.124
N	647250	647250	647250	647250	647250	647250

Standard errors in parentheses
Dependent variable is parents' average log income

IV regressions. Income. Second stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Parents	1.599	1.616	1.656	1.692	1.745	1.717
	(0.015)	(0.014)	(0.013)	(0.016)	(0.016)	(0.011)
r2	.	.	.	.	.	.
N	647250	647250	647250	647250	647250	647250

Standard errors in parentheses. Dependent variable is child's predicted years of schooling

Table A3: IV regressions. Occupational stautus. First stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Aunts and uncles	0.305	0.281	0.273			0.196
	(0.001)	(0.001)	(0.001)			(0.001)
Spouses of aunts/uncles		0.165	0.160			0.112
		(0.001)	(0.001)			(0.001)
Parents' cousins			0.098			0.043
			(0.002)			(0.002)
Grandparents				0.397	0.386	0.269
				(0.002)	(0.002)	(0.002)
Great grandparents					0.044	0.020
					(0.002)	(0.002)
Parents' aunts/uncles						0.063
						(0.001)
r2	0.133	0.158	0.163	0.164	0.166	0.213
N	647250	647250	647250	647250	647250	647250

Standard errors in parentheses
Dependent variable is parents' average social stratification score

IV regressions. Occupational stautus. Second stage results.

	(1)	(2)	(3)	(4)	(5)	(6)
Parents	0.085	0.085	0.086	0.078	0.079	0.083
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
r2	.	.	.	0.027	0.022	0.006
N	647250	647250	647250	647250	647250	647250

Standard errors in parentheses
Dependent variable is child's predicted years of schooling

References

Acemoglu, D., and J. Angrist (2001). "How large are human-capital externalities? Evidence from compulsory-schooling laws", *NBER Macroeconomics Annual 2000, Volume 15*. MIT Press, 9-74.

Becker, G. S. and N. Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", *Journal of Political Economy* 87(6), 1153–1189.

Björklund, A., M. Lindahl, and E. Plug (2006). "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data", *Quarterly Journal of Economics*, 121(3), 999–1028.

Black, S. and P. Devereux (2010). "Recent Developments in Intergenerational Mobility", in *Handbook of Labor Economics*, vol. 4B, eds. O. Ashenfelter and D. Card, 1487-1541. Amsterdam: Elsevier.

Böhlmark, A., and M. J. Lindquist (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.

Borjas, G. J. (1992). "Ethnic Capital and Intergenerational Mobility" *The Quarterly Journal of Economics*, *107*(1), 123-150.

Boserup, S. H., W. Kopczuk and C. Thustrup Kreiner (2014). "Intergenerational Wealth Mobility: Evidence from Danish Wealth Records of Three Generations." Working Paper, University of Copenhagen.

Braun, S. and J. Stuhler (2014) "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany" Mimeo, November 2014, Universidad Carlos III de Madrid.

Chan, T. W. and V. Boliver (2013). "The Grandparents Effect in Social Mobility: Evidence from British Birth Cohort Studies", *American Sociological Review* 78(4), 662–679.

Chetty, R., N. Hendren, P. Kline and E. Saez (2014). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", *Quarterly Journal of Economics* 129(4): 1553–1623.

Clark, G. (2014). *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton, NJ: Princeton University Press.

Clark, G. and N. Cummins (2014) "Intergenerational Wealth Mobility in England, 1958–2012: Surnames and Social Mobility", *Economic Journal*, 125(582), 61–85.

Coleman, J. S. (1988): "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, 94, pp. S95–S120.

Haider, S. and G. Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings", *American Economic Review* 96(4), 1308–1320.

Holmlund, H. (2008). “Intergenerational Mobility and Assortative Mating: Effects of an Educational Reform”, Centre for the Economics of Education, CEE DP 91.

Hällsten, M. (2014) “Inequality across three and four generations in Egalitarian Sweden: 1st and 2nd cousin correlations in socio-economic outcomes”, *Research in Social Stratification and Mobility* 35: 19–33

Jaeger, M. M. (2012). “The Extended Family and Children’s Educational Success”, *American Sociological Review* 77(6), 903-922.

Lambert, P. S., and E. Bihagen, (2007). “Concepts and Measures: Empirical evidence on the interpretation of ESeC and other occupation-based social classifications”, *International Sociological Association*.

Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren (2015) “Long-term intergenerational persistence of human capital: an empirical analysis of four generations” *Journal of Human Resources* 50(1), 1–33.

Long, J. and J. Ferrie (2013). “Intergenerational Occupational Mobility in Great Britain and the United States since 1850”, *American Economic Review* 103(4), 1109–1137.

Lubotsky, D., & Wittenberg, M. (2006). “Interpretation of regressions with multiple proxies. *The Review of Economics and Statistics*, *88*(3), 549-562.

Merton, R. K. (1938): “Social Structure and Anomie”, *American Sociological Review*, 3(5), pp. 672–682.

Myrdal, G. (1944) *An American Dilemma, Volume 2: The Negro Problem and Modern Democracy*. Vol. 2. Transaction Publishers.

Nybom, M., and J. Stuhler (2013). “Interpreting Trends in Intergenerational Income Mobility” Swedish Institute for Social Research Working Paper.

Nybom, M., and J. Stuhler. (2016). “Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation”, *Journal of Human Resources*, Vol. 51(1), 239-268.

Prandy, K. and Lambert, P. S. (2003) “Marriage, social distance and the social space: an alternative derivation and validation of the Cambridge scale”, *Sociology*, 37, 397-411.

Solon, G.. (1999). “Intergenerational Mobility in the Labor Market”, in: O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Vol. 3*. Amsterdam: Elsevier.

Solon, G. (2013). “Theoretical Models of Inequality Transmission across Multiple Generations”, *NBER Working Paper* 18790.

Solon, G. (2015). “What Do We Know So Far about Multigenerational Mobility?”, Mimeo Michigan State University, January 2015

Stuhler, J. (2013). “Mobility Across Multiple Generations: The Iterated Regression Fallacy”, IZA Working Paper 7072.