CESifo WORKING PAPERS

7615
2019

April 2019

Dynastic human capital, inequality and intergenerational mobility

Adrian Adermon, Mikael Lindahl, Mårten Palme

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7615*
Category 5: Economics of Education

Dynastic human capital, inequality and intergenerational mobility

Abstract

We study the importance of the extended family – the dynasty – for the persistence in inequality across generations. We use data including the entire Swedish population, linking four generations. This data structure enables us to identify parents' siblings and cousins, their spouses, and the spouses' siblings. Using various human capital measures, we show that traditional parent-child estimates of intergenerational persistence miss almost one-third of the persistence found at the dynasty level. To assess the importance of genetic links, we use a sample of adoptees. We then find that the importance of the extended family relative to the parents increases.

JEL-Codes: I240, J620.

Keywords: intergenerational mobility, extended family, dynasty, human capital.

Adrian Adermon
Institute for Evaluation of Labor Market and Education Policy (IFAU)
Uppsala / Sweden
Adrian.Adermon@ifau.uu.se

Mikael Lindahl
Department of Economics
University of Gothenburg
Gothenburg / Sweden
Mikael.Lindahl@economics.gu.se

Mårten Palme
Department of Economics
Stockholm University
Stockholm / Sweden
Marten.Palme@ne.su.se

April 11 2019
We are grateful for comments from Gordon Dahl, Markus Jäntti, Chris Karbownik, Marco Manacorda, Magne Mogstad, Martin Nybom, Jan Stuhler, Kelly Vosters, Olof Åslund and from participants at seminars at the Department of Economics as well as the Department of Statistics at Uppsala University, Helsinki School of Economics, Lund University, University of Copenhagen, University of Gothenburg, Stockholm University, Hong Kong University, Singapore Management University, IFAU, ESPE, UCLS, the Family and Education workshop in Hawaii, the CESifo Economics of Education Conference in Munich, ISER Workshop on Family Economics at University of Essex, the NBER Cohort Studies Meeting in Los Angeles, the Intergenerational Transmission of Economic Status: Exposure, Heritability and Opportunity conference in Madrid and the Zurich Workshop in Economics. Mikael Lindahl is currently the Torsten Söderberg Research Professor at Handelshögskolan, University of Gothenburg. He also acknowledges financial support from Jan Wallanders and Tom Hedelius Stiftelse, Tore Browaldh Stiftelse and the European Research Council [ERC starting grant 241161]. Mårten Palme gratefully acknowledges financial support from the Swedish Council of Social Research. Adrian Adermon gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Research Foundation.

1 Introduction

Recent research has highlighted intergenerational mobility as a way of avoiding perpetuation of inequality and promoting economic growth and social progress (see e.g. Chetty et al., 2014, 2016). A series of recent studies have shown a growing concern that the traditional parent-child regression model does not give an exhaustive description of social mobility and that such models underestimate the long-term persistence of social positions across generations. For example, Braun and Stuhler (2018), Lindahl et al. (2015), Long and Ferrie (2018) and Mare (2011) highlight the contributions from generations preceding the parental one, while Jaeger (2012) examines influences from other relatives. The results based on surnames in Guell, Rodriguez-Mora and Telmer (2014) and, in particular, in Clark (2014) point to much higher long-term intergenerational social persistence than what is implied by estimates from AR(1) models.[5] Others are more skeptical to the new results and argue that the influence of ancestors might be spurious (Solon, 2018) and that group-level estimates, such as those in Clark (2014), are distinct from estimates of the traditional child-parent parameter (Chetty, 2014; Solon, 2018).

In this study we use extraordinary Swedish administrative data enabling us to construct (overlapping) family trees spanning four generations, and to observe several measures of human capital and social status for the individuals. This makes it possible to identify the extended family, or the *dynasty*, based on actual family relations instead of surnames as in Clark (2014) and Guell, Rodriguez-Mora and Telmer (2014).[6] Our main contribution is to estimate the degree of intergenerational mobility where we – in addition to the outcomes of the parents – also consider the extended family in the parental generation (aunts/uncles and their spouses as well as parents' cousins and their spouses). We propose a method to decompose the traditional parent-child mobility measure into one component reflecting how the extended family moves across generations, and one component measuring the individual's own within-dynasty mobility. We compare our estimates to those from models including outcomes from the grandparent and the great grandparent generations. We also study what parts of the extended families that contribute to the additional persistence. Finally, we reformulate our various models in a latent variable framework using three different indicators for "social status" (years of schooling, lifetime family income, and an index of occupational-based social stratification) for the parental and ancestor generations, in the spirit of Vosters (2018) and Vosters and Nybom (2017) using the method outlined in Lubotsky and Wittenberg (2006).

Our data combines several population-wide Swedish registers. We use GPA in the last year of compulsory schooling and years of schooling as measures of educational outcomes for up to 575,000 individuals in the child generation. The Multigenerational register is used to link all family connections, requiring identification of the great grandparents of each individual in the child generation. Data from several censuses with information on educational attainment, labor earnings, and occupation for the period 1968-2009 are used to construct outcomes for everyone in the parent as well as grandparent and great grandparent generations. The fact that the entire Swedish population is included in the data allows us to link dynasties up to parents'

[5] Olivetti and Paserman (2015) use a similar strategy but instead of surnames they use first names to create pseudo links between fathers and sons as well as fathers and daughters.
[6] See also Aaronson and Mazumder (2008), who used data from several censuses (which lacked information on intergenerational family links) and state-level information to form intergenerational estimates at the state-of-birth and birth cohort level to compare trends in intergenerational mobility.

siblings and cousins, the siblings' and cousins' spouses (through marriage and cohabiting records), and the siblings of the aunts and uncles.

Our results show that estimation of between-dynasty intergenerational models generates a much higher degree of intergenerational persistence in social position than the standard individual intergenerational models: individual persistence across generations in the *Years of schooling* measure increases by 50 percent, from 0.28 for the simple parent-child model to 0.42 when we use the between-dynasty persistence measure, where the latter figure can be interpreted as a one standard deviation (SD) higher years of schooling of the *average* parent in the dynasty being associated with 0.42 SD higher years of schooling of a child. Using GPA as the outcome measure gives a similar picture, with the intergenerational correlation increasing from 0.36 to 0.52. The decomposition analysis reveals that more than 50 percent of the total persistence in *GPA* and *Years of schooling* can be attributed to between-dynasty persistence. When we estimate the models in terms of "social status" using the three indicators mentioned above, the between-dynasty persistence estimates increase to 0.48 and 0.59, for years of schooling and GPA respectively.

Results from sequentially adding parts of the dynasty show that each part of the extended family makes a highly significant contribution to overall persistence. However, they also show that the parents and their siblings capture most of the persistence (around 80 percent). The multigenerational estimations show that even the great grandparent generation makes a small, but statistically significant contribution in the AR(3) model. The persistence contribution of the grandparent generation is, however, more important: in the AR(2) model about 20 percent of the total persistence can be attributed to the grandparent generation when we correct for measurement error using the latent variable framework.

Previous research has repeatedly shown that that group-level effects is a key mechanism behind persistence in socio-economic positions across generations. Following the social capital theory (see Coleman, 1988) and the strain theory (see Merton, 1938), several studies have shown the importance of social class. There is also a large empirical literature on the importance of race and ethnicity (Borjas, 1992; Hertz, 2008; Torche and Corvalan, 2018). However, the group most closely connected to the parents' socio-economic position, and therefore most closely related to the traditional parent-child measure of social mobility, is the extended family, the group we focus on in this paper. We also extend the previous literature including grandparents and great grandparents by comparing the importance of these generations with the influence of the extended family in the parental generation. In addition, we estimate new multigenerational models within the latent variable framework, hence correcting for measurement error bias and interpreting the estimates in terms of "social status" transmission.

To study the mechanisms behind our results we use a sample of adoptees. This enables us to remove the genetic family links and isolate environmental mechanisms. The intergenerational associations between the human capital outcomes for adopted children and parents are about 25-30% of the estimated associations between non-adopted children and their parents, quite similar to other adoption studies for Sweden (Holmlund, Lindahl and Plug, 2011). When we extend the adoption design and estimate extended family models for adopted children, we find that both the adoptive parents and other extended family members contribute to the human capital outcome of the adopted child. As expected, because genetic factors are likely to be more important for the child-parent association than for the association between the child and aunts and uncles etc., the contribution of the extended family becomes relatively more important compared to the contribution of parents when the genetic link is removed.

The paper proceeds as follows. In Section 2, we discuss empirical specifications using the standard and extended models incorporating dynastic capital, a latent variable framework, and adopted children and their parents. 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 on the importance of dynastic human capital and look at the separate contributions of the dynastic members and the ancestors. In section 5 we present results using adopted children and their genetically unrelated adoptive parents. In Section 6, we provide some sensitivity analyses and discuss external validity. Section 7 concludes the paper.

2 Empirical Specifications

2.1 The individual AR(1) model

Most empirical studies on intergenerational mobility use the simple individual AR(1) model (see e.g. Solon, 1999):

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

where y represents the outcome under study; i is an index for the child; j is an index for the nuclear family including children (t) and parents (t-1). Assuming that we can measure the outcome perfectly for both generations, the OLS estimate of β is an unbiased measure of the strength of the linear association between the outcome for parents and children. For instance, if y represents years of schooling or the logarithm of lifetime earnings, β will estimate the intergenerational schooling coefficient or the intergenerational earnings elasticity. If y is standardized to have the same variance, an estimate of β (or the square root of the R^2 in this regression) can also be interpreted as an intergenerational correlation (see, e.g., Solon, 1992).

2.2 The extended family and dynasty model

A limitation of the individual AR(1) model is that it only measures the influence from the parents and ignores potential influences from other relatives such as grandparents or aunts/uncles (the extended family). If outcomes for the extended family are predictive of a child's success, a random individual born into a high-SES family is likely to have better outcomes than what is measured by the model in Equation (1). Previous papers have therefore extended this model by adding measures of outcomes for aunts/uncles and/or grandparents.

If we include outcomes for other family members of the parental generation to the standard model in Equation (1) we get:

$$y_{idt} = \alpha + \beta_1 y^{p}_{dt-1} + \beta_2 y^{sp}_{dt-1} + \beta_3 y^{cp}_{dt-1} + \cdots + \beta_K y^{K}_{dt-1} + \varepsilon_{idt}, \tag{2}$$

where i is an index for the child; d is an index for the dynasty, including children (t) and members in the parental generation (t-1), whose relationship with the child is denoted by a superscript, where p denotes a parent; sp parent's sibling, and; cp parent's cousin, up to K relatives, or dynasty categories, all drawn from the parental generation ($k = p, sp, cp, \dots, K$).

Under the assumption of equal variance for the outcomes in the parental generation, i.e., that $\sigma^2_{y^k_{dt-1}} = \sigma^2_{y^{k\prime}_{dt-1}}$ for all dynasty categories k and k', we have that plim $\hat{\gamma} = \beta_1 + \beta_2 + \beta_3 + \cdots + \beta_K$, where $\hat{\gamma}$ is the coefficient estimate from:

$$y_{idt} = \alpha' + \gamma \bar{y}_{dt-1} + \varepsilon'_{idt}, \tag{3}$$

where $\bar{y}_{dt-1} = \sum_{k \in \{p,sp,\dots,K\}} y^k_{dt-1}/K$, i.e., the average of the K members of dynasty d in the parents' generation.[7] Since all k:s are from the same generation, equal variances for the outcomes is not a strong assumption.[8] However, because the number of individuals in each dynasty category varies (e.g., the number of parent's siblings differ from the number of parent's cousins), we standardize the average of the outcome for members in each dynasty category to have mean zero and standard deviation equal to one, since we want to give equal weight to each dynasty category. An OLS estimate of γ should therefore be interpreted as the change in the dependent variable (always scaled in standard deviation units) for a child associated with, on average, a standard deviation unit higher outcome for the members in the dynasty in the parental generation, where each of the dynasty categories are weighted equally. If all dynasty members are parents (or if non-parents are similar as parents in their influence on nephews) this means that we can interpret an OLS estimate of γ as a one standard deviation (SD) increase in years of schooling of the *average* parent in the dynasty being associated with a γ increase in years of schooling or GPA of a child.

If we have a balanced panel, so that all members in the parental generation have at least one child, WLS (with weights equal to the number of children in each dynasty) using y aggregated within each dynasty in the child generation as a dependent variable, i.e., $\bar{y}_{dt} = \alpha' + \gamma \bar{y}_{dt-1} + \bar{\varepsilon}_{dt}$, gives us an identical estimate of γ as the one obtained by OLS in Equation (3).

Equation (3) captures dynastic persistence in one parameter, γ, which, if we have a balanced panel, measures intergenerational persistence in inequality *between* dynasties (d), as opposed to β in Equation (1) which captures the intergenerational persistence in inequality between nuclear families (j). Hence, these parameters capture intergenerational persistence at different levels. We argue that knowledge about persistence at both these levels is necessary in order to understand inequality transmission across generations. However, because β and γ are different parameters, stemming from models formulated using different units of aggregation, inferring the importance of the dynasty by directly comparing estimates of these two parameters can be misleading. Therefore, we next propose an alternative approach to infer the importance of the dynasty for social mobility.

---

[7] Lichtenberg (1990) shows that if the true model is $y = \beta_1 x_1 + \beta_2 x_2 + u$, but we estimate $y = \gamma(x_1 + x_2) + u$, we have that the probability limit of an OLS estimate of γ is plim$\hat{\gamma} = w \cdot \beta_1 + (1 - w) \cdot \beta_2$. If x_1 and x_2 have the same variance (but still allowed to be correlated), we then have $w = 1/2$, so that plim$\hat{\gamma} = (\beta_1 + \beta_2)/2$. This result can be generalized to more than two variables.

[8] If we have unequal variances, but a balanced panel of children and external family members, Equation (2) can still be rewritten as $y_{idt} = \alpha' + \gamma \bar{y}_{dt-1} + \varepsilon'_{idt}$. This result now holds because by observing the complete kinship tree of children and parents, the order of the k's in (2) will vary symmetrically across i's. This approach easily generalizes to wider dynasty definitions. In fact, this result holds for any panel with a symmetric structure of coefficients and variables across i's.

2.3 The relationship between the dynasty model and the AR(1) model

To understand the relationship between the parameters β and γ, one can use the result from panel data analysis that a pooled OLS estimator can be written as a variance-weighted average of the within- and between-unit estimators (see e.g. Green, 1990, p. 471-472). In an intergenerational setting with a balanced panel, the traditional child-parent estimator, $\hat{\beta}$, can be decomposed as a variance-weighted average of the between- and within-group parameter estimates (Borjas, 1992; Hertz, 2008; Torche and Corvalan, 2018). In our specific setting, where the group level is the dynasty, we have that

$$\hat{\beta} = \left(1 - \hat{k}_B\right)\hat{\beta}_w + \hat{k}_B\hat{\gamma}, \tag{4}$$

where the weight, $\hat{k}_B$, equals the estimated fraction of the variance in the outcome of the parental generation that is due to between-dynasty variation, i.e., $\hat{k}_B = \hat{\sigma}_B^2/\hat{\sigma}^2$. This implies that an OLS estimate of β in Equation (1) is equal to a weighted average of the within-dynasty estimate $\hat{\beta}_w$ and the between-dynasty estimate $\hat{\gamma}$, where the weights depend on the degree of within- and between dynasty variation for outcome of the parental generation. We can calculate the fraction of $\hat{\beta}$ that is due to between-dynasty variation in the data as $\hat{k}_B\hat{\gamma}/\hat{\beta}$. The fraction of $\hat{\beta}$ that can be attributed to within-dynasty variation in the data is then just $1 - \hat{k}_B\hat{\gamma}/\hat{\beta}$.[9]

Equation (4) makes it clear that although β and γ are different parameters, stemming from models formulated using different units of aggregation, they have an exact relation. By decomposing the traditional social mobility parameter into parts attributed to within- and between variation, using Equation (4), we can directly characterize the importance of the dynasty for a *given* social mobility in a society. The decomposition shows what share of the parent-child persistence that can be attributed to being member of a particular family dynasty and how much that can be attributed to persistence within the dynasty.

Some extreme cases may be of particular interest for the interpretation of the decomposition. One of these is when all persistence parameters are equal, i.e., if $\beta = \beta_W = \gamma$. This case would imply that the formation of dynasties would not impose additional persistence across generations and is easily tested for using a Hausman-Wu type test. If there were no persistence between dynasties – or, say, that the dynasties were formed randomly - the between parameter would be zero and all parent-child persistence would be attributed to parent-child persistence within the dynasty. The other extreme case would be if *all* persistence would be attributed the dynasty, i.e., within the dynasty, parents would not matter for the child's position and β_W would equal zero.

Equation (2) is a generalized version of a model with spillover effects from members in some group captured into one variable added to model (1), as specified in Borjas (1992) analyzing intergenerational spillovers from ethnical group members. Reinterpreting Borjas' model to our setting yields the following extension of Equation (1):

[9] An estimate of the fraction of within-dynasty variation can also be obtained by estimating β_w directly if we include dynasty fixed-effects in the model and then multiply the coefficient estimate by $(1 - \hat{k}_B)/\hat{\beta}$. Given the complexity of our data structure with dynasties overlapping (see footnote 23), we instead estimate the fraction of within-dynasty variation as described in the text.

$$y_{ijdt} = \alpha' + \beta' y_{jdt-1} + \delta \bar{y}_{d_{-j}t-1} + \varepsilon'_{jdt}, \tag{5}$$

where $\bar{y}_{d_{-j}t-1}$ is an average of y_{jdt-1} over the members of the dynasty in the parents' generation, excluding the parents. Borjas (1992) shows that estimating (1) by OLS, when Equation (5) is the true model, results in $plim\ \hat{\beta} = \beta' + \delta k_B \leq \beta' + \delta = \gamma$.[10] Hence, $\hat{\beta}$ is a lower bound estimate of γ, where the difference between β and γ is increasing (decreasing) in the relative degree of within (between) dynasty variation in y_{jdt-1}. Only if there is no within-dynasty variation in y_{jdt-1}, so that $k_B = 1$, would $\hat{\beta}$ be an unbiased estimate of γ.[11] In our setting the spillovers would come from all relatives in the parental generation, excluding the parent. We will estimate Equation (5) as well.

2.4 The relationship to multi-generational models

The framework above captures associations with grand- (and great grand-) parents indirectly through the association of outcomes of children with outcomes of their parent's siblings and cousins. Although it utilizes multigenerational data to form these horizontal extended family links, it requires outcome data on two generations only. This is a great advantage since data sets on outcomes for multiple generations are not easy to find and are more susceptible for measurement problems.[12]

There is a literature that has estimated multigenerational models including outcomes for parents and grandparents, and some cases even great grandparents. The specifications estimated are of the following type:

$$y_{ijt} = \alpha' + \rho_1 y_{jt-1} + \rho_2 y_{jt-2} + \rho_2 y_{jt-3} + \varepsilon_{it}, \tag{6}$$

where *y* represents the outcome under study; *i* is an index for the child; *j* is an index for the vertically extended family consisting of the children (*t*), parents (*t-1*), grandparents (*t-2*) and great grandparents (*t-3*).

Equation (6) can be extended to the dynasty framework as:

$$y_{idt} = \alpha' + \gamma_1 \bar{y}_{dt-1} + \gamma_2 \bar{y}_{dt-2} + \gamma_2 y_{dt-3} + \bar{\varepsilon}_{dt}, \tag{7}$$

where $\bar{y}_{dt-1} = \sum_{k \in \{p,sp,....K\}} y^k_{dt-1}/K$, i.e., an average over the *K* members (i.e., dynasty categories) of the dynasty d in the parents' generation and $\bar{y}_{dt-2} = \sum_{q \in \{gp,sgp\}} y^q_{dt-2}/Q$ is an average of y^{gp}_{dt-2}, and y^{sgp}_{dt-2}, i.e., an average of the outcomes for the Q=2 dynastic categories

[10] In Borjas' model, the spillover term was defined as the average earnings in the ethnical group in the parental generation. Whether or not we exclude the parents from the spillover term does not matter for these formulas in large samples, as long as the sample is balanced (or that the non-parents have similar characteristics as the parents) since then the numerator in $\hat{k}_B = \hat{\sigma}^2_B/\hat{\sigma}^2$ will be identical.

[11] The relationship between Equation (4) and the results in Borjas (1992) follows directly from $\hat{\beta} = (1 - \hat{k}_B)\hat{\beta}_w + \hat{k}_B\hat{\gamma} \cong$[if $\beta' = \hat{\beta}_w$; $\hat{\gamma} = \beta' + \delta$]$= \beta' + \hat{k}_B\delta$.

[12] Outcomes for multiple generations can often be difficult to compare. An example is the change from a schooling system that limits the available slots to higher than compulsory schooling, generating a skewed and narrow years of schooling distribution with a high fraction of individuals having only the minimum required years of schooling, to a schooling system with a higher fraction of students at many levels.

in grandparents' generation (grandparents and grandparent's siblings). As we do not have data on dynasty members in the first generation we simply add y_{dt-3} for the great grandparent generation.

As pointed out by Solon (2018), omitted group effects can be picked up by outcomes of ancestors in Equation (6). By including control for dynasty averages, as in Equation (7), we are able to control for the relevant group effects and investigate how this changes conclusions regarding the importance of multigenerational effects.[13]

Again, working with standardized variables (separately for the parent and grandparent generation), γ_1 will capture the sum of the association with the outcomes for the members in the parental generation and γ_2 the sum of the association with the outcomes for the members in the grandparental generation. In an extended analysis we estimate both AR(2) and AR(3) multigenerational models.

2.5 The relationship to a latent variable model

Following Clark (2014), the underlying intergenerational relation can be formulated as a latent variable relation, i.e.,

$$y^*_{ijdt} = \alpha'' + \beta'' y^*_{jdt-1} + \varepsilon''_{ijdt}, \qquad (8)$$

where $y^*_{ijdt} = \bar{y}_{dt}$ and $y^*_{ijdt-1} = \bar{y}_{dt-1}$, i.e., the dynasty averages in the corresponding generations are assumed to fully capture the latent variables. The observed individual variables can be written as the sum of the dynasty averages and an individual specific part: $y_{ijdt} = \bar{y}_{dt} + \eta_{ijdt}$ and $y_{jdt-1} = \bar{y}_{dt-1} + \eta_{jdt-1}$. As pointed out by Torche and Corvalan (2018), if η_{ijdt} and η_{jdt-1} are iid, the latent variable model in equation (8) can be used to estimate the parameter β in the individual model.[14] However, if η_{ijdt} and η_{jdt-1} are correlated, which would be the case if the parental deviation from the dynasty average contains information for the expected outcome in the child generation, the latent variable model would estimate another parameter compared to the individual model.

One way to test the independence assumption is to use the decomposition in Equation (4). If an estimate of the within parameter, β_w, can be shown to be significantly different from zero, we can reject the iid-assumption of the latent variable model (8). Hence, the model with dynasty averages across generations will estimate a different intergenerational parameter than the standard intergenerational model.[15]

[13] Our approach is related to directly adding parental characteristics (other than y_{jt-1}) to Equation (6), as is done in Warren and Hauser, 1997, and Braun and Stuhler, 2018, with the purpose of investigating whether this produces smaller estimates of the grandparental parameter ρ_2. An advantage with our approach is that we add characteristics in both generations (through the LW approach) which hence will decrease measurement error of some underlying characteristic for both generations.

[14] This is very similar to the simple Wald model for treating errors in variables (see e.g. Johnston, 1984).

[15] The issue that an estimate of an intergenerational group-level parameter is distinct from an estimate from the traditional child-parent parameter was emphasized by Chetty et al. (2014) and Solon, (2018) in their critique of Clark's work using surname averages when estimating intergenerational models.

2.6 Excluding the genetic link using adoptees in extended family models

We use adoption families to understand the role of genetic links in intergenerational transmissions in the dynasty model.[16] The model used in most studies using adoptees has the following form:

$$y_{ijzt}^{ac} = \alpha + \theta y_{jt-1}^{ap} + \varepsilon_{it}, \qquad (9)$$

where i is an index for the adopted child ac born in family $z \neq j$ and reared in family j by the adoptive parent indicated by the super-index ap.

The parameter θ can be interpreted as showing the strength of the intergenerational transmission of y due to (post-birth) environmental factors for the sample of adoptees conditional on the following three assumptions:[17]

1) Adoptees are conditionally randomly assigned to adoptive families.
2) The adoption should have taken place close to birth. If this is not the case, the postnatal pre-adoption environment (e.g., the quality of the nursery home) is uncorrelated with the post adoption environment (or has no influence on the outcome of the adopted child).
3) The biological parents have no contact with the adopted child post adoption.

In this paper we use data on foreign-born children adopted by Swedish parents before 12 months of age. Assumption 3 should hold in a sample of foreign-born adoptees. By imposing the age at adoption restriction we expect Assumption 2 to hold approximately. Finally, as we will show in Section 5.1, we are able to assess to what extent the critical Assumption 1 holds empirically.

The results from Equation (9) exclude the genetic transmission channels between parents and children. If the adoptees are comparable to a representative sample of the population of children, we can compare estimates of θ in Equation (9) with estimates of β in Equation (1) to learn about the fraction of the overall intergenerational association that is due to environmental factors.

If one is prepared to assume that, conditional on parents' outcomes, there is no genetic family link between the outcomes of the other extended family members and the child's outcome, the association of other extended family members and the child will increase relative to the association between the parent and the child, using the population of children. However, this assumption is unlikely to hold because the associations between the child and extended family members, conditional on the parents, can be due to genetic factors as i) we only observe proxies for some underlying trait ("human capital"), meaning that extended family members can pick up some of the genetic link, unaccounted for by the outcome measure used for parents, and ii) we only observe the outcomes (the phenotypes) and not the genotypes, meaning that there can still be information about the genetic background in analyzed outcomes for grandparents and aunts/uncles (even if we could observe all relevant outcomes).

[16] Models using adoptees has become an established method in social science (see Sacerdote, 2011, for a survey).
[17] See Björklund, Lindahl and Plug (2006).

If we reformulate the extended family model with adopted children and their genetically unrelated parents and other family members in the parental generation, Equation (2) can be rewritten as:

$$y_{idt}^{ac} = \alpha + \theta_1 y_{dt-1}^{ap} + \theta_2 y_{dt-1}^{sap} + \theta_3 y_{dt-1}^{cap} + \cdots + \theta_K y_{dt-1}^{K} + \varepsilon_{idt}, \qquad (10)$$

where *i* is an index for the adopted child *ac*; *ap* denotes an adoptive parent; *sap* the adoptive parent's sibling, and; *cap* the adoptive parent's cousin, up to K relatives of the adopted parent. Note that the parameters now are labelled as θ_k instead of β_k to make it clear that the parameters in Equation (10) now is purged from genetic factors explaining the association between the parental and child generations.

This gives the dynastic model for adoptees as

$$y_{idt} = \alpha' + \theta \bar{y}_{dt-1}^{ap} + \varepsilon'_{idt}, \qquad (11)$$

where $\bar{y}_{dt-1}^{ap}$ is an average of y_{dt-1}^{k}, i.e., over the *k* members of the dynasty *d* in the adoptive parents' generation ($k = ap, sap, cap, \ldots, K$), and plim $\widehat{\theta} = \theta_1 + \theta_2 + \theta_3 + \cdots + \theta_K$ assuming equal variances across groups. The parameter θ measures the degree of dynastic inequality due to environmental factors.

2.7 Measurement error bias and a latent variable framework

Measurement errors will create different forms of biases in the models estimated in this paper. First, if the true model is an AR(1) one (as in Equation (1)), but the model that is estimated contains outcomes for other relatives, any mismeasurement of parents' outcome can be picked up by the added outcomes for other relatives, inducing a spurious correlation with child outcomes (see Solon, 2018). Second, in our application, comparisons of estimates for parents and dynasty averages can be problematic since the latter measure will be expected to have less measurement errors simply by being an average across individuals. Hence, measurement error will lead to an estimate for parents being more attenuated than an estimate for the dynasty average. Third, in the multigenerational models, the data on outcomes for great grandparents, and to some degree also for grandparents, are of lower quality, leading to possible larger downward bias in ancestors' contribution compared to parents.[18]

We address these problems by using a latent variable approach proposed by Lubotsky and Wittenberg (2006). This approach produces a coefficient that equals the coefficient on an optimally weighted linear combination of multiple proxies. In the same spirit as Vosters and Nybom (2017) we re-estimate all models discussed above in a latent variable framework. We use three different indicators for "social status": years of schooling, lifetime family income and an index of occupational-based social stratification. The approach suggested by Lubotsky and Wittenberg (LW) will also correct for classical measurement bias. Hence, the difference between the standard OLS and the LW-adjusted estimates can be due to both changed interpretation in terms of latent variables and elimination of classical measurement error bias.

[18] The data for these ancestors are still based on administrative sources. However, for earnings and occupation, the information is only available at later ages, and for years of schooling, there is less variation in the data because fewer individuals continued to post-compulsory education levels.

The LW approach proceeds by first regressing the outcome variable on the full set of k proxy variables x_{ki}, $y_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i$. The coefficients from this regression are then combined to give the coefficient on the latent variable as the linear combination $\hat{\beta}^* = \sum_k \frac{\mathrm{cov}(y,x_k)}{\mathrm{cov}(y,x_1)} \hat{\beta}_k$.[19] This coefficient is scaled to be directly comparable to x_1, which is always *Years of schooling* in our regressions.[20]

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 biological (and adoptive) 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 January 1, 1961.[21] For the purpose of this paper, we require i) that we can identify cousins in the parental generation, i.e., that we are able to identify families through four generations; and ii) that we are able to measure human capital outcomes in the parent and child generations.

We proceed as follows. We restrict the individuals in the child generation to be born no earlier than 1972, which is the earliest birth cohort where a comprehensive human capital indicator is available in Swedish registry data: The Grade Point Average (*GPA)* at the end of compulsory school (age 16) is constructed from the national grade 9 registers using grades in all compulsory subjects. To capture later education for the child generation we also use *years of schooling* as an outcome variable, constructed using information from national educational registers. The last year for which we observe data on GPA and educational attainment is in 2009. Hence, the child generation birth cohorts are 1972-1993 for GPA and 1972-1983 for years of schooling (to give the individuals enough time to finish tertiary education). Because grandparents of these child cohorts must be born 1932 at the earliest to be present in our data, the years-of-schooling sample is less representative and much smaller.

We then link these children to their parents and other relatives using personal identification numbers. We construct vertical family links up to great grandparents, and are hence able to identify cousins in the parent generation and second cousins in the child generation. Marriage and cohabiting registers further extend horizontal links. This means that we have linked dynasties up to siblings and cousins of parents, as well as (through marriage and cohabiting

[19] In practice, we estimate the LW coefficient using residuals from regressions of each variable on the full set of birth year controls.

[20] Regarding separately estimating the coefficients in the multivariate regression models (Equations (2) and (3)) and in the multigenerational regression models (Equations (6) and (7)) we acknowledge that this is more difficult. For instance, in Equation (6), the outcome for the parents is a mediating variable in the relation between grandparents' and child's outcomes. We therefore view estimates from these regression models as being more suggestive regarding their separate contributions.

[21] For the non-adopted children, we always use the biological ancestors of the child, regardless of whether these are the parents that raise the children.

records) the siblings' and cousins' spouses, as well as the siblings of the siblings' spouses.[22] In principle, it would be possible to (almost infinitely) extend the size of the dynasties, where individuals are genetically linked and/or linked through assortative mating. In practice, we add additional relatives until they no longer provide any explanatory power.

We further compile data from registers (and censuses for earlier years) that contain information on education, income and occupation for the parental and other ancestor generations. The education information is available in the 1970 census and in yearly registers between 1985 and 2009.[23] Income data are drawn from tax registers and are 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 require that at least one of each type of relative in the parental generation must have survived and still be working in 1970.

For the parental and other ancestor generations we use, in addition to years of schooling, two outcomes: log family income and the so-called CAMSIS index for occupation-based social stratification. The income measure we use is calculated as the sum of gross labor earnings, income from businesses and unemployment benefits. Average log *family income* is calculated in the following way: we use income data for all available years for each individual between ages 30-60; we take logs and residualize by adjusting for both birth cohort and income year fixed effects; we then take the average of the residuals for each individual. Lastly, we take averages among parents (if we only observe one parent, we use that observation). The Cambridge Social Interaction and Stratification (CAMSIS) measure of social distance uses occupations of spouses to create an index (0-100) of *social stratification*. The basic idea is that individuals who are similar in terms of social status are more likely to marry each other.[24] While there are many occupation-based social classifications, the CAMSIS scale has two advantages for our purposes – first, unlike categorical classifications of social class schemes (e.g., Erikson et al., 1979), it is continuous; second, unlike the Socio-Economic Index of occupational status (ISEI) and similar measures (Ganzeboom et al., 1992), it does not rely on income or education in its construction. Hence, CAMSIS provides independent information beyond that contained in our schooling and income variables.[25]

The main independent variables are constructed by taking averages of non-missing observations within each category of relatives (parents, grandparents, aunts/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

[22] We are not the first to link four generations using the Swedish Multigenerational registry (see e.g.. Hällsten 2014, and Persson and Rossin-Slater, 2018).

[23] Both for the child and parental generations, we construct years of schooling as follows: seven for (old) primary school, nine for (new) compulsory schooling, 9.5 for (old) post-primary school (*realskola*), 11 for short high school, 12 for long high school, 14 for short university, 15.5 for long university and 19 for a PhD. For the child generation we mainly use the latest educational register available, which is for 2009. If education for the individual is missing in 2009, we use 2008, and so forth.

[24] The CAMSIS score is constructed by analyzing a frequency cross-table of husbands' and wives' occupations. This table maps out the space of social distances, and from this, it is possible to locate each occupation along an index of social status or stratification. We use the Swedish CAMSIS scale based on data for 2001-2007 prepared by Erik Bihagen and Paul Lambert, available at http://www.camsis.stir.ac.uk/Data/Sweden90.html.

[25] Lambert and Bihagen (2014) compare a large set of occupation-based social classifications, showing that most measures tend to be relatively highly correlated with each other, and that CAMSIS performs relatively well in predicting unemployment and health.

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 directly averaging across all relatives, we would implicitly be giving a disproportionately large weight to, e.g., parents' cousins, simply because there are relatively many of these in the average family. In the estimations we always standardize these dynastic category averages to have mean zero and standard deviation one.

3.2 Descriptive statistics

Table 1 shows descriptive statistics for the data set, focusing on the sample where we use GPA as the child outcome.[26] The first three columns show averages of the number of years of schooling, average residualized log family income and the social stratification index (from 0 to 100) by category in the dynasties. Note that we here show means and standard deviations for the original variables, whereas in the result tables below we always show estimates using standardized variables. The fourth column shows the average number of observations 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 for a child to be included in the main regressions.

The means and standard deviation for GPA for the child generation is shown at the top of Table 1. The original scores were transformed to percentile ranked by birth cohort, in order to take into account changes in the grading over time. If we correlate GPA with years of schooling for the subsample where both measures are available for the same individuals, we get a correlation of 0.62.

The means and the standard deviations of years of schooling (Column 2) is similar within generations, but differs a lot across generations. The reason is that the schooling systems have pretty much remained constant within, but not between, generations. Income and the occupation index is here already transformed into measures that are calculated within birth cohorts, hence showing no trend across generations.

Summary statistics for all categories (except the children) are based on averages over various numbers of individuals, with more observations used for more distant dynasty categories. Hence, the standard deviations of these averaged variables will be lower. This shows why it is important to standardize the averages calculated for each dynasty groups, if we want to make the estimates comparable.

TABLE 1 ABOUT HERE

Table 2 shows correlations between the three main variables years of schooling, log family income and the social stratification index. In Panel A, 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, as

[26] We show descriptive statistics for the data set where we use years of schooling as the outcome variable for the child generation in Appendix Table A1 and descriptive statistics for the sample of adopted children in Appendix Table A6.

correlations ranges between 0.25 and 0.52. In Panel B, where we use the dynasty as the unit of observation, the pattern is similar although all three correlations increase.

TABLE 2 ABOUT HERE

4 Results

4.1 The associations between outcomes of children, parents and the dynasty

Table 3 shows the first set of results. The dependent variable is either GPA or years of schooling for the individual in the child generation, both standardized to have mean zero and standard deviation one in the respective sample. The table contains four columns for each outcome, where we either use years of schooling, the logarithm of family income, the social stratification index or the LW-weighted index of all three measures, as a measure of socio-economic outcome for the members in the parental generation. The LW weights are scaled to be comparable to *Years of schooling* of the parents - the LW estimates can thus be compared to those in the first and fifth columns, respectively. All right-hand side variables are standardized to have zero mean and unit standard deviation within each dynasty category (i.e., parents; parents' siblings; etc). We use linear and quadratic controls for average year of birth for each type of relative included and year of birth indicators for the children in all specifications.

Each row in Table 3 shows the results from separate regressions. The estimates shown in the first row correspond to Equation (1), which is similar to the traditional specification for the estimation of the parent-child transmission coefficient, although we here use the average of the socio-economic outcome measures of the parents. Each row then sequentially uses broader and broader definitions of the dynasty, i.e. corresponding to different dynasty definitions in Equation (3). In the specification of the second row we take averages over the measure for the parents, parents' siblings and their spouses; in the third row the average also includes cousins and their spouses; and in the fourth row we add outcomes for the siblings of the spouses of the parents' siblings. We always report robust standard errors.[27]

TABLE 3 ABOUT HERE

The first row, Columns 1 through 3, show that a one standard deviation unit higher parental outcome is reflected in a between 0.26 and 0.36 of a standard deviation higher GPA relative to the mean, depending on the measure of parental outcome used. Columns 5 through 7 show that the corresponding result for *Years of schooling* is 0.22 and 0.28. The result for *Years of schooling* is similar to what has typically been obtained in previous studies on Swedish data.[28] When we use the LW weighted averages as outcome, we find the estimates to increase to 0.46 for GPA and to 0.39 for *Years of schooling*, an increase by about 29-37% relative to the

[27] However, even though the errors are likely to be correlated within dynasties we do not report cluster-robust standard errors. The reason is that the child is the unit of analysis and dynasties in the parental generation will therefore overlap (e.g., an aunt of one child can be the mother of another child). Since we have the population of individuals, and the clusters are relatively small, imposing various cluster definitions is of very little importance for the precision of our estimates.

[28] See e.g. Holmlund et al. (2011) for evidence on the relationship between years of schooling of children and parents.

estimates reported in Columns 1 and 5 of the first row. This increase is higher than what was found in Vosters and Nybom (2017) for intergenerational income regressions.[29]

Turning to the results shown in the second, third and fourth rows, we can see that the estimates, as expected, increase as the dynasty definition becomes broader. Note that the dynasty measures are based on averages over standardized variables for each dynasty category. As explained in Section 2, this implies equivalence of the coefficient estimate in terms of approximately being equal to the sum of the estimates from a multivariate regression with outcomes for the separate dynasty categories included as independent variables. The estimates should thus be interpreted as the change in the dependent variable associated with a standard deviation unit higher parental generation outcome, where each of the dynasty categories are weighted equally. If we for instance focus on the estimate in the second row of first column, which equals 0.449, it should be interpreted as the average of the standard deviation unit higher years of schooling for parents, parents' siblings and parent's siblings' spouses, where each of these three dynasty groups are weighted equally, being reflected in half a standard deviation higher GPA of the child.

If we compare the estimates in the first and fourth rows – the traditional parent-child regression with the broadest dynasty definition - we see that the point estimates increase with between 46 and 78 percent. Results are very similar for the two measures of child outcome: GPA and *Years of schooling*, but the percentage increase is higher using *Income* and *Occupation* (67-78%), compared to using *Years of schooling* (46-48%), as parental measure.

When we use the LW weighted index, the estimates increase less, by 25-28%, and become 0.592 for GPA and 0.484 for years of schooling.[30] This is expected since the LW approach already to a high degree adjusts for measurement error bias (see the discussion in section 2.7), which is decreased as dynasty averages are calculated over more individuals. We also see that the increase in the estimate from applying the LW approach, compared to the standard AR(1) estimate, as well as the increase in the estimate from broadening the dynasty definition, compared to only using outcome for the parents, both are sizable.

4.2 Decomposing the standard intergenerational child-parent estimate into parts due to within and between variation

It is evident from Table 3 that the between-dynasty estimates are always larger than the traditional child-parent estimates, and that they are substantially larger using broad dynasty definitions. As shown in Equation (4), the traditional parent-child persistence parameter can be decomposed into a weighted average of the within- and between-dynasty persistence parameters. This decomposition shows what share of the parent-child persistence that can be attributed to being a member of a particular family dynasty and how much that can be attributed to persistence heterogeneity within the dynasty.

[29] This can simply be explained by the different variables that we focus on. Years of schooling is transformed from seven levels of education, and as such there is no variation within these levels (for a given birth cohort). We also use both men and women, and their results increases more for women. We also use an occupational index that is different.

[30] Braun and Stuhler (2018) use a latent variable model and three generations of data for Germany and find estimates that range between 0.494 and 0.699 using schooling as outcome variable. Hence, our estimate of 0.484 using the LW approach is at the lower end of that range.

Table 4 shows the results from such decomposition. Panels A, B and C, show estimates using years of schooling, log family income and the social stratification index, respectively. Finally, panel D shows the results using the LW weighted index of these three variables, as outcomes for the parents. For each panel and outcome, we show estimates for the three dynasty definitions from the second, third and fourth (from less to more broader dynasties) rows of Table 3.

TABLE 4 ABOUT HERE

In each panel, the first two rows show estimates of the between-dynasty parameter (taken from the second to fourth row of Table 3) and the between-dynasty weight. In the third row, we take the product of these two estimates and divide by the overall standard child-parent estimate, to produce an estimate of the fraction of the standard child-parent estimate that is due to between-dynasty variation in the data. These are the key estimates from each panel, and for each outcome, as this answers the question of how much of the inequality in the parental generation that is transmitted to the child generation that is determined by the dynasty the child is born into. For reference, we also show estimates of the standard child-parent estimate at the top of each panel.

The fraction of intergenerational persistence attributed to between-dynasty persistence depends on the definition of the dynasty. A tighter definition of the dynasty would attribute a larger share of the parent-child persistence to the between-dynasty component. An extreme case would be if the dynasty coincides with the nuclear family so that *all* parent child persistence would be attributed to the dynasty. Conversely, if the dynasty is extended to the entire population, all persistence would be attributed to the within component. Following this line of reasoning, it is not surprising that the estimates of the between component in Table 4 (third row) are decreasing as the definition of the dynasty is widened across columns. This result is driven by the estimated between weights (second row) which, by construction, decrease with a widening of the dynasty.

The results in Table 4 also reveal that the fraction of the persistence attributed to the dynasty varies with the variable used to measure the position in the parental generation. If we confine ourselves to the widest definition of the dynasty in panels A through C, we see that using *Years of Schooling* for the parental generation combined with GPA as outcome measure, attributes the largest fraction, 0.60, to between persistence, while the parental generation *Income* measure combined with *Years of Schooling* as outcome, gives the smallest share at 0.45. Although this range is quite large, the results taken together suggest that a large part of the observed parent-child persistence should be attributed to the dynasty.

In the last panel we show LW estimates scaled to years of schooling (so as to be comparable to the estimates in panel A). Using this measure, the estimated persistence attributed to the between component decreases from 0.60 to about 0.46 for the broadest dynasty definition. However, the latter estimate implies that still almost half of the variation of the standard social mobility estimate is due to between-dynasty variation. This figure can be compared to studies of other groups, such as ethnicity (Borjas, 1992, Torche and Corvalan, 2018), which generate much smaller estimates of the variation of the standard social mobility estimate that is due to between-dynasty variation (estimates range from 0.00 to 0.19).[31]

In the fourth row of each panel we show the within estimate calculated from Equation (4). This estimate is approximately equal to the parental coefficient in a regression including dynasty fixed effects. As expected, the within-estimates increases as the dynasty definitions become broader. Using the broadest definitions (Column 3), we see that the within-estimates are about 65-75% of the standard child-parent estimates shown at the top of each panel.

As explained in Section 2.4, a way of testing if the latent variable model using dynasty-means would estimate the parameter in the individual parent-child model is to test if the within parameter β_w is equal to zero. If we can reject this, we have shown that the outcomes of the parents, conditional on being member of a particular dynasty, indeed has predictive power for the outcome of the child, which implies that the latent variable model estimates another parameter than the individual model.

As we already noted, it is very apparent that β_w-estimates are significantly different from zero. It is also apparent that $\hat{\beta}_w < \hat{\gamma}$ in all specifications implying that there is indeed less persistence within the family conditional on the dynasty compared to between dynasties. Indeed, this is a part of the explanation to why Clark (2014) get such a large discrepancy between his group-level estimates and the ones based on individual data collected from other studies.

4.3 Results from models with outcomes from relatives entered separately

In Table 5 we disentangle the association between GPA in the child generation and years of schooling for different dynastic categories of the extended family by sequentially adding variables inversely related to the family distance to the parents, estimating different version of Equation (2). Column 1 of Panel A shows the results from the first row of the first column in Table 3, as a reference. Then, we sequentially include *Years of schooling* of parents' siblings, spouses of aunts/uncles, parents' cousins, spouses of parents' cousins and siblings of spouses of aunts/uncles. The last column shows the results from a model that decomposes the overall dynasty average into schooling for parents and for the other dynasty members, respectively. In panel B, we show LW weighted coefficient estimates, from a regression that in addition to schooling also includes income and occupation as proxy variables for human capital for relatives.

TABLE 5 ABOUT HERE

The results shown in Panel A of Table 5 reveal highly significant estimates for all parts of the dynasty. The magnitude of the aunt/uncle coefficient is about one-third of the estimate for

[31] In Torche and Corvalan (2018) the variation of the standard social mobility estimate that is due to between-group variation is 0.19 for 12 race/ethnical groups and 0.00 for 5 European race/ethnical groups using data from NLSY, 1979. In Borjas (1992), for education, the estimates for ethnic groups are 0.14 using data from GSS and 0.12 using data from NLSY, 1987 (calculated from columns 1-3 of Table III). These estimates are based on the reported estimates of $\hat{\beta}$, $\hat{\beta}_w$ and $\hat{\gamma}$, which, in combination with Equation (4), can be used to calculate the variation of the standard social mobility estimate that is due to between-dynasty variation as $\frac{\hat{\gamma}}{\hat{\beta}} \cdot \frac{\hat{\beta}-\hat{\beta}_w}{\hat{\gamma}-\hat{\beta}_w}$.

parents (Column 2), but more than twice as large as the corresponding one for the spouses of aunts and uncles (Column 3). Years of schooling for parent's cousins are roughly half as important as parent's siblings schooling (Columns 4 and 5). However, even schooling of the siblings of spouses of aunt/uncles enters statistically significant, conditional on schooling for all the other dynasty categories, with an estimate less than one-tenth of the one for schooling of parents. We note that controlling for all relatives, the estimate for parents' schooling decreases but is still quite large (0.28).[32] In the last column, where we estimate Equation (5), we find that parents schooling is more strongly associated with child's schooling, than what is the case for the dynasty average, but that the dynasty average is large as well, contributing about two-fifth of the sum of the two estimates.

Panel B of Table 5 shows the results when we have used the Lubotsky-Wittenberg method for combining indicators. The estimates are always interpretable in terms of years of schooling for each relative. We find that the estimate for parental years of schooling (in the first row in each panel) always increases, but that the estimates for years of schooling for other relatives are basically unchanged. This is expected since parent's years of schooling is an imperfect proxy for parent's "social status", which is partly captured by the years of schooling of other relatives. Hence, the contribution of the dynasty average decreases somewhat to about one-third of the sum of the two estimates. Still, using high quality schooling measures from administrative data sources *and* correcting for mismeasurement of some underlying latent variables, still leaves a very significant contribution from the extended family, other than the parents.[33]

We also report the sum of the coefficients in the bottom of both panels of Table 5. As expected (see Section 2.2), these sums are very similar to estimates of the dynastic persistence parameters in Table 3. In addition, we note that the coefficient estimate for parent's years of schooling in the first row of Column 6 (0.284) falls somewhere in between the estimate in Column 1 (0.359) and the calculated within estimate in the fourth row of Column 3 in Table 4 (0.243). This is expected since the calculated within estimate in Table 4 (from Equation (4)) is comparable to a child parent model with dynasty fixed effects, and hence control for additional features of the dynasty not captured by the dynasty categories included in the estimates shown in Table 5.

4.4 Results from the multigenerational models

So far, we have concentrated on the horizontal dimension in the parental generation. However, there is a large (mostly) recent literature that has analyzed multigenerational associations, i.e., looking at the vertical dimension by adding outcomes of grandparents (and, in some cases, even great grandparents) to the standard parent-child model. In most estimations of such AR(2) models, grandparents have been found to provide additional

[32] 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 the child's years of schooling on parents' education, SES and income, as well as aunts' and uncles' educations, SES and income. They found that, conditional on parents' outcomes, only aunts' and uncles' education was statistically significantly associated with the child's years of schooling. The coefficient estimates were less than one-third of those for parents.

[33] In Appendix Table A2 we show that for the smaller sample where we use years of schooling as outcome variable for the children the results remain very similar. The results are also very similar if we use GPA as the outcome variable in the smaller schooling sample (see Appendix Table A3).

information, although coefficient estimates are often quite small and less likely to be detected in smaller samples.[34] These results suggest that the parent-child association sometimes yields misleading results regarding long-run intergenerational persistence (i.e., a parent-child estimate overestimates the degree of social mobility for descendants to the child), using various outcomes such as schooling, income, occupation, and wealth.

Table 6 shows results from various multigenerational regressions using GPA as outcome for the child and years of schooling as the outcome for ancestors (as before all variables are standardized to have mean zero and standard deviation equal to one). Panel A shows estimates from Equation (6) using the parents, grandparents and great grandparents, and Panel B from Equation (7) using the average of the dynastic members in the parental, grandparental and great grandparental generations, respectively. Column 1-4 use actual years of schooling of ancestors and Columns 5-8 use the LW approach, incorporating income and occupational rank of the ancestors. Again, for each generation, the LW estimates are scaled to be comparable to the estimates for *Years of schooling*.

TABLE 6 ABOUT HERE

From the results in Panel A, Columns 1-4, we first see that grandparents' schooling (Column 2) is strongly related to GPA of the child: an additional standard deviation (SD) of years of schooling of grandparents is, on average, associated with one-fifth of a SD GPA for children. In Column 3, we see that grandparents' schooling is associated with child's GPA, even conditional on parents' schooling. The conditional estimate for grandparents is still sizable, and we also note that the estimate for parents' schooling only decreases by about 10% when we control for grandparents' schooling. If we compare the estimates for grandparents' schooling in columns 2 and 3, viewing parents' schooling as a mediating variable, we would conclude that parents' schooling explains 63% of the association between grandparents' schooling and child's GPA ((0.208-0.077)/0.208).[35] The results in Column 4 show that the estimate for great-grandparents' education is, although small, statistically significant in the AR(3) model.

A problem with interpreting the estimates in Columns 1-4 of Panel A, and with almost all multigenerational estimates in the literature, is that measurement error can lead to biased estimates from the outcomes for grandparents in the AR(2) model and great grandparents in the AR(3) model. There are two counteracting effects. First, the noise due to measurement error in parents' outcomes could be picked up by grandparents' outcomes leading to an upward bias of the estimates of the contribution from grandparents. On the other hand, measurement errors are likely more prevalent for outcomes based on data collected further back in time, and hence, the estimate for grandparents might be more attenuated.[36]

[34] See Anderson et al. (2018) and Solon (2018) for two surveys.

[35] This is very much in line with Anderson et al. (2018) who review 69 analyses and find that, on average, 30% of the child-grandparent association remains when information on parents is included. Hence, parents explain 70% of the child-grandparent association.

[36] Although we are using high-quality administrative data, the individuals in the various generations have gone through different schooling systems (e.g., with different years of compulsory school and differential access to higher education institutions) and for great grandparents, (and to some, but less, degree grandparents), the distribution of years of schooling is right-skewed due to fewer individuals pursuing years of schooling above the compulsory level.

To handle this potential problem, we estimate the multigenerational models using the latent variable LW method to combine the three human capital proxies for individuals in each ancestor generation. Results are shown in Columns 5-8 of Panel A. Interestingly, they are very similar for grandparents in the AR(2) model (Column 7), whereas the estimate for parents and grandparents in the unconditional models increase by 29% and 21%, respectively. Parents' "social status" now explains 72% of the association between grandparents' and child's GPA.

Following Ferrie et al. (2016) the last row of Panel A shows prediction from the multigenerational model for descendants 10 generations ahead. Using the estimates from the AR(3) model in Column 4, we find that an AR(1) model would have to produce an estimate equal to 0.49, in order to predict the same intergenerational persistence after 10 generations as the AR(3) model. This is 37% higher than the actual estimate of 0.36 as shown in the first row of Column 1. Using the LW approach we find the AR(3) model to produce a prediction that is 21% higher than the AR(1) model in the first row of Column 5. Hence, the additional information from ancestors is still large enough to be economically meaningful. Ferrie et al. (2016), using multigenerational data linked across censuses during the 20th century, found intergenerational education persistence to be underestimated by 20%.[37]

Panel B shows results from Equation (7) using the average of the generation-specific dynasty members. These models therefore control for dynasty specific group effects (measured at the parental dynasty level) and arguable controls for any remaining measurement error bias by averaging over many individuals. Interestingly, years of schooling of the grandparent's dynasty members (constituting of the grandparents and the aunts/uncles of the parents) do indeed still contribute to the GPA of the children, even conditional on the average of all the dynasty members' years of schooling in the parental generation. As can be seen in Columns 5-8 of Panel B, this is also the case if we use the LW approach. If we compare the estimates for the dynasty in the grandparents' generation, viewing the dynasty in the parental generation as a mediating variable, we would conclude that the dynasty in the parental generation explains around 80% of the association between grandparents' schooling and child's GPA.

If we estimate multigenerational schooling associations using years of schooling, instead of GPA, as human capital measure for the children, we get much weaker associations with grandparents' years of schooling (we always use variables standardized to have mean zero and standard deviation one in the sample to facilitate comparison). Results are shown in Appendix Table A4. The estimate for grandparents in the AR(2) model is about 10% of the estimate for parents (Column 3 of Panel A). This can be compared to the estimates reported above in Table 6, where the relationship was about 24% of the parental one. Parents' schooling explains between 76-86% of the association between grandparents' and child's schooling, and for the dynasty averages, the dynasty at the parental generation explains everything. However, this discrepancy appears partly due to the different measures and partly due to the much more selected sample we have to rely on for these results.[38] This can be seen in Appendix Table A5, where we report estimates for the smaller sample using GPA as child

[37] They deal with measurement error bias by using measures of educational attainment of the ancestors reported at two points in time for the same individual.

[38] As explained above, since we have to allow about a decade longer for children to accurately measure their years of schooling, we lose about five-sixth (83%) of the sample. These children that are included in this smaller sample are only those from families that (over these generations) give birth at young ages. Otherwise we will not be able to link four generations of data.

outcome. We then find the association to be 17% of the parental one and that parents' schooling now explains between 66-75% of the association between grandparents' schooling and child's GPA. These figures are in between those from the large GPA sample and the much smaller schooling sample using child's schooling as outcome. Hence, we conclude that the multigenerational associations would have been weaker if we had been able to use years of schooling as child outcome for the larger more representative sample.[39]

5 Results from the Models using Adoptees

5.1 Sample restrictions, descriptive statistics and tests of quasi-randomization of adoptees

We now turn to the results from the sample of adoptees. The purpose of this extension is to investigate the strength of the intergenerational associations in a sample where there is no genetic family links. To be able to interpret these estimates as showing that the family environment and/or the extended family environment matters for child's GPA, we need i) to restrict the sample to those children adopted at an early age and ii) the adoptees in our sample to be quasi-randomly assigned. To increase the probability of meeting these requirements, we have restricted the sample to international adoptees and to those adopted before their first birthday. The adoption age is calculated as the difference between the immigration date and the birth date. Both dates are obtained from Swedish administrative registers.[40]As the years of schooling sample becomes very small for the sample of adoptees, we only estimate models using GPA as an outcome variable for the children.

We show summary statistics for the sample of adoptees in Appendix Table A6. If we compare the figures with those in the population (Table 1), we see that the adoptive parents are more educated, have higher income and score higher on the social stratification index but that the children's GPA is very similar. The adoptive parents are also born earlier whereas the adopted children are, on average, similarly aged to the population of children.

Following Sacerdote (2007), Fagereng, Mogstad and Ronning (2018), Holmlund, Lindahl and Plug (2011) and Lundborg, Nordin and Rooth (2018) we start by investigating the quasi-randomization assumption for foreign-born adoptees by regressing variables determined prior to adoption (gender and adoption-age of the adoptees) on measures for the parents (in our case years of schooling, income and social stratification). In addition, we show results for the dynasty measures. The results from these tests are shown Appendix Table A7. The outcome variables are *Child gender* (Panel A) and *Age of adoption* in months (Panel B). We show results without and with controls for region-of-birth fixed effects. We cannot reject that the children are randomly placed in families. For child gender the parent and dynasty measures are always statistically insignificant. For adoption age, most estimates are statistically insignificant, and those that are significant are only marginally so and the magnitudes of

[39] Comparing the standardized estimates to those reported for years of schooling in Lindahl et al. (2015) (0.096-0.110), who used data on three (and four) generations for individuals from the city of Malmö, Sweden, the estimates for years of schooling for grandparents in the AR(2) model for are now much more precisely estimated and smaller, using GPA as child outcome, and much smaller, using years of schooling as child outcome. However, they share the conclusion that the AR(1) model can be rejected.

[40] To facilitate comparison with our other results, the variables are standardized using the means and standard deviations from the full population sample.

estimates are in all cases extremely small. One year of higher years of schooling of parents (about 0.5 SD) is associated with one-tenth of a month (3 days) lower adoption age.[41] We conclude that we cannot reject that international adoptees in Sweden (adopted at infancy) during this time are in effect quasi-randomly assigned.[42]

5.2 Results

The results from the sample of adoptees are shown in Table 7. Panel A shows estimates using dynastic averages (Equation 10); Panel B shows estimates from multivariate models (Equation 9); and, finally, Panel C shows results from the multigenerational models. Columns 1-4 show the baseline GPA-schooling estimates, and Columns 5-8 the GPA-schooling estimates using the LW approach including *Years of schooling, Income* and *Occupational rank*. Again, the estimates are scaled to be comparable to the *Years of schooling* estimates.

The first column shows the standard child-parent estimate using the adoption sample. The estimate is positive and statistically significant but the magnitude is only about one quarter of the size of the child-parent estimate for the population (0.359). The child-parent estimate for adoptees is fairly similar to earlier estimates in the adoption literature using education outcomes (see Holmlund, Lindahl and Plug, 2011). This reflects a strong influence of genetic factors for school achievement and other education outcomes.[43]

The estimates in Columns 2-4 of Panel A show that dynasty averages clearly give larger estimates than for parents, although the increase seems to be driven by the narrow dynasty average including parents' siblings and their spouses, and not the more distant extended family members. This result is confirmed in Panel B, where we separately include selected extended family members. Aunts'/uncles' schooling is statistically significantly related to children's GPA (Column 2), and their spouses seem to possibly matter as well. When we separately include the parents and the dynasty average (net of parents) in Column 4, we see that the schooling of the other extended family members combined appear to be (at least) as

[41] For older adoptees (not adopted within 12 months) we do see evidence of systematic placement, likely because those who wanted to adopt quicker could do so by adopting older children, which probably also meant that these adoption families were of higher SES on average.

[42] Sacerdote (2007) and Fagereng et al (2018) found Korean adoptees to be quasi-randomly assigned to adoptive families using these types of quasi-randomization tests. They argue that Korean adoptees were assigned quasi-randomly because the adoption agencies handling the adoptions (Holt) worked on a first-come first-serve basis, conditional on a family having fulfilled various (age-, marriage) restrictions and interviews (see Fagereng, Mogstad and Ronning, 2018). The families in their sample could not specify any characteristics of the child that they wanted (except if being open to an older child). In Sweden, the families can (after clearance by the Swedish social authorities) contact adoption agencies for different countries where to various degree specific characteristics of the child (i.e., age, gender) can be specified (see Lundborg, Nordin and Rooth, 2018, and SOU, 2003). Even though we can control for most of these characteristics, the excess demand for *infant* adoptees is probably what made the adoptions conditionally quasi-random in Sweden, since it is very costly for a family to decline a child in terms of waiting time. Holmlund, Lindahl and Plug (2008, 2011) and Lundborg, Nordin and Rooth (2018) investigate quasi-randomness for children adopted by Swedish parents children and born abroad mainly during the 1970s. Both these papers find some evidence of selection using similar tests as we do in this paper, but conclude that magnitudes of the estimates are every small. Interestingly, both papers find that Korean born children are selectively placed in Swedish families.

[43] Haegeland et al. (2010) is one of very few papers that have regressed adopted children's grade scores on years of schooling of parents. They use data on Korean-born children adopted by Norwegian parents and find the resulting estimates to be about one-third of those using non-adopted children. For surveys of other approaches to estimate intergenerational causal effects, such as twin-parents fixed effects and IV approaches, see Björklund and Salvanes (2011), Black and Devereux (2010) and Holmlund, Lindahl and Plug (2011).

important as the schooling of parents for the GPA of the adopted children. In the last panel, investigating multigenerational associations, the sample seems too small to generate precision enough to detect any associations beyond the parents.

Turning to Panel B, and the LW estimates, we see that the difference between estimates for parents and the dynasty averages decreases, just as was the case for the population estimates (in Table 3). As is evident from Panel 2, however, the schooling of aunts/uncles, and possible their spouses, still associate with adopted children's GPA, conditional on the years of schooling of the adopted parents. In the last column, we see further evidence of this, although the combined contribution of the other extended family members is no longer larger than the contribution of the adoptive parent. However, the relative contribution for the extended family, compared to the parents, is clearly larger for adoptees than for the population estimates (from one-third to almost one-half compared to the results reported in Column 7 of Table 5). In Panel C, we do see a statistically significant association between the adopted children and the parents of the adoptive parents, although when the outcomes for the latter are included as control the association with grandparents becomes insignificant.

To summarize the results for adoptees, we can conclude that both the parents and the other extended family members do matter for the GPA of the adopted child, and hence, that also the environment of the nuclear as well as the extended family seems to matter. As expected, given that we expect genetic factors be more important for the child-parent association than for the child-dynasty (net of parents) association, the contribution of the extended family becomes relatively more important compared to the contribution of parents, in this sample.

TABLE 7 ABOUT HERE

6 Sensitivity analysis and external validity concerns

The intergenerational persistence estimates of Clark (2014), using dynasties linked with surnames, have been criticized for ignoring the potential effect of neighborhood, race and ethnicity, for which surnames are indicative and would therefore result in an overestimate of true persistence across families (see e.g. Chetty et al., 2014, and Solon, 2018). Solon (2018) argues that a positive association with grandparents in AR(2) models can be driven by omitted group effects such as neighborhood, race, ethnicity etc. This is potentially an issue also for the analysis in this paper.

There are, however, at least two reasons to why we think that this critique is less relevant in our context. First, since we require great-grandparents to be identified in the data for all our estimates, our sample consist of children whose ancestors have been living in Sweden for at least four generations. At that time, Sweden was very ethnically homogenous and it is therefore unlikely that group effects based on race or ethnicity would affect our results. Second, as shown by Lindahl (2011) neighborhoods are of limited importance in explaining variation in outcomes such as earnings, schooling and student achievement in Sweden.

Nevertheless, to test if residential location can explain our large dynasty estimates, we have added various regional fixed effects to the baseline models. The results of this exercise are shown in Appendix Table A8. Comparing the baseline results in Column 1 to those from models including parish fixed effects (Column 2) and municipality fixed effects (Column 3), it is obvious that the baseline results are very robust, suggesting that the main results are not driven by dynasty members growing up in the same regions.

The results in this paper use administrative data that has made it possible to define broad extended family links, which is likely not possible in data sets for all other countries. This makes our results particularly intriguing, but at the same time opens up for criticism for lack of external validity. What can our results teach us about intergenerational mobility in other countries? One way to examine this important issue is to treat Swedish regions as separate entities, and estimate the same intergenerational models for each region. We then relate the intergenerational estimates across regions. We ask the following questions: what is the relation between the traditional intergenerational persistence parameter and i) the intergenerational dynasty persistence parameter and ii) the share of the traditional intergenerational persistence parameter that can be attributed to between-dynasty variation?

The results from our county-level estimates are shown in Figures 1 and 2.[44] The standard child-parent estimates are measured on the horizontal axis in both figures and the dynasty ones on the vertical axis. Figure 1 reveals a strong positive relationship between the child-parent and the dynasty parameters. The estimate (standard error) is 0.869 (0.180). This suggests that countries with higher intergenerational persistence will have higher intergenerational dynastic persistence and that we cannot reject a 1:1 relation between the two measures. Figure 2 shows the relation between the child-parent estimates and the fraction of this estimate that is due to between-dynasty variation. As can be seen in the figure, there is no apparent relation (estimate (standard error) is 0.296 (0.728)). Although this estimate is imprecise, we cannot reject that our share of between-variation can be extrapolated for other countries and there is no evidence of a negative relationship. Hence, a higher intergenerational persistence at the individual level would not be predicted to result in a higher degree of between-variation. These results are reassuring from the point of view of ability to extrapolate the results in this paper to those for other countries, where only the parent-child model can be estimated.

7 Conclusions

There are several previous studies using various group-level estimators – such as Borjas (1993) on immigrant groups, Sharkey (2008) on neighborhoods or Clark (2014) on groups with common surnames. Relating to that literature and the observation, starting with Hodge (1966), that different family members seem to have independent associations in outcomes, we propose a new measure of intergenerational social mobility based on the between extended family, or dynasty, mobility across generations. We argue that the dynasty is the group most closely connected to an individual's social position in a society. We show how the traditional parent-child measure of social mobility can be decomposed in our proposed between-dynasty component and one within-dynasty component. We also show a framework for how to study the contribution of each part of the extended family to the overall dynasty-persistence and how the measure relates to multigenerational mobility.

Our results show that estimates of between-dynasty mobility reflect a much stronger social persistence than the estimates from the traditional parent-child regression model. For *GPA*

[44] We define our regions using mother's county of residence from the 1985 census. Sweden had 24 counties in 1985.

and *Years of schooling*, the human capital measures we use for the child generation, the coefficient estimates increase by almost 50 percent using Years of Schooling for the parents (from 0.36 to 0.52 and 0.28 to 0.42, respectively). Our results also suggest that about 50 percent of the parent-child correlation in educational attainments can be attributed to between-dynasty persistence. We find that most of the additional persistence of the extended family can be measured by the outcomes of the siblings of the parent, i.e., uncles and aunts, although both their spouses and parents' cousins contribute as well. Furthermore, the results reveal that the dynasty model gives a stronger persistence than when we include outcomes from grandparents and great grandparents.

The previous empirical literature on intergenerational mobility is dominated by two main measures. The first is based on parent-child regressions and the second on sibling correlations (see e.g. Solon, 1999, and Björklund and Jäntti, 2012). In this paper, we propose a measure based on between extended-family mobility. This parameter measures the share of the advantage/disadvantage of the extended family that is expected to remain in the next generation. Compared to the measure based on parent-child regressions, it reduces the attenuating effect of only using the individual outcome in the parental generation, if the extended family is important. In contrast to the measures based on sibling correlations it does not include the direct effects of shared home environment, neighborhood and, in most cases, elementary school quality. In our view, it adds a vital piece of information to our understanding of the different aspects of persistence in economic positions across generations.

References

Aaronson, D. and B. Mazumder (2008) "Intergenerational Economic Mobility in the United States, 1940 to 2000" *Journal of Human Resources* 43(1), 139–172.

Anderson, L. R., P. Sheppard, C. W. S. Monden, "Grandparent Effects on Educational Outcomes: A Systematic Review," *Sociological Science* 5: 114-142.

Björklund, A., and M. Jäntti (2012). "How Important is family background for labor-economic outcomes", *Labour Economics,* 19, 465–474.

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.

Björklund, A. and K. Salvanes (2011). "Education and Family Background: Mechanisms and Policies", in *Handbook of Economics of Education*, vol. 3, eds. E. A. Hanushek, S. Machin and L. Woessmann, 201-247. Amsterdam: Elsevier.

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.

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

Borjas, G. J. “The Intergenerational Mobility of Immigrants.” *Journal of Labor Economics*, vol. 11, no. 1, 1993, pp. 113–135.

Braun, S. and J. Stuhler (2018). “The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany”, *The Economic Journal* 128(609) 576-611.

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.

Chetty, R., N. Hendren and L. Katz (2016). “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment”, *American Economic Review* 106(4): 855–902.

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

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

Erikson, R., J.. Goldthorpe, and L. Portocarero. “Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden.” *The British Journal of Sociology* 30, no. 4 (1979): 415–41.

Fagereng, A., M. Mogstad, and M. Ronning (2018) “Why do wealthy parents have wealthy children?” CESifo Working Paper No. 6955.

Ferrie, J., C. Massey and J. Rothbaum (2016). “Do Grandparents and Great-grandparents Matter? Multigenerational Mobility in the US, 1910-2013”. NBER Working Paper 22635.

Ganzeboom, Harry B. G., Paul M. De Graaf, and Donald J. Treiman. “A Standard International Socio-Economic Index of Occupational Status.” *Social Science Research* 21, no. 1 (March 1, 1992): 1–56.

Griliches, Z. (1979) “Sibling Models and Data in Economics: Beginnings of a Survey”, *Journal of Political Economy* 87(5, Part 2), S37-S64.

Güell, M., J. V. Rodríguez Mora and C. I. Telmer (2015) “The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating,” *The Review of Economic Studies*, 82(2), Pages 693–735

Haegeland, T., L. Kirkebøen, O. Raaum, and K. Salvanes (2010) “Why children of college graduates outperform their schoolmates: A study of cousins and adoptees, IZA Discussion Papers, No. 5369.

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.

Hertz, T. (2008). “A Group-specific Measure of Intergenerational Persistence”. *Economics Letters* 100:415-17.

Hodge, R. W. (1966). “Occupational Mobility as a Probability Process”, *Demography*, *3*(1), 19-34.

Holmlund, H., M. Lindahl, and E. 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.

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

Johnston, J. (1984) “Econometric Methods” Singapore: McGraw-Hill.

Lambert, Paul S., and Erik Bihagen. “Using Occupation-Based Social Classifications.” *Work, Employment & Society* 28, no. 3 (June 1, 2014): 481–94.

Lichtenberg, F. R., (1990). “Aggregation of Variables in Least-Squares Regression”, *The American Statistician* 44, no. 2 (May, 1990), pp. 169-171.

Lindahl, L. (2011). “A comparison of family and neighborhood effects of grades, test scores, educational attainment and income – evidence from Sweden”, *Journal of Economic Inequality* 9, 207-226.

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.

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

Mare, R. D. (2011) “A Multigenerational View of Inequality,” *Demography* 48(1), pp. 1-23.

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

Long, J. and J. Ferrie (2018). “Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the U.S. and Britain, 1850-1910,” *The Economic Journal* 128, F422-F445.

Lundborg, P., M. Nordin and D-O Rooth (2018). “The intergenerational transmission of human capital: the role of skills and health,” *Journal of Population Economics*, 31, pp. 1035-1065.

Olivetti, C., and M. D. Paserman.(2015) “In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940,” *American Economic Review* 105 (8), 2695-2724.

Persson, P. and M. Rossin-Slater (2018) “Family Ruptures, Stress, and the Mental Health of the Next Generation,” *American Economic Review* 108(4-5), 1214-52.

Sacerdote, B. (2007). "How Large are the Effects from Changes in Family Environment? A Study of Korean American Adoptees", *The Quarterly Journal of Economics* 122(1) 119-157.

Sacerdote, B. (2011). "Nature and Nurture Effects On Children's Outcomes: What Have We Learned From Studies of Twins And Adoptees?", in: J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, Vol. 1. Amsterdam: Elsevier.

Sharkey, P. (2008). "The Intergenerational Transmission of Context," *American Journal of Sociology* 113(4), pp. 931-969.

Solon, G. (1992). "Intergenerational Income Mobility in the United States", *American Economic Review* 82 (June), 393–408.

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. (2018). "What Do We Know So Far about Multigenerational Mobility?", *The Economic Journal* 128(612), pp. F340-F352.

Torche, F. and A. Corvalan (2018). "Estimating Intergenerational Mobility With Grouped Data: A Critique of Clark's the Son Also Rises", *Sociological Methods & Research*, 1-25.

Vosters, K. (2018). "Is the simple Law of Mobility really a law? Testing Clark's hypothesis". *The Economic Journal* 128(612), F404-F421.

Vosters, K., and M. Nybom (2017). "Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States", *Journal of Labor Economics*, Vol. 35(3), 869-901.

Warren, J.R. and Hauser, R.M. (1997) "Social Stratification across Three Generations: New Evidence from the Wisconsin Longitudinal Study," *American Sociological Review*, 62(4), 561-572.

Table 1. Summary statistics

	Years of schooling	Log income (residualized)	Social stratification	Observations/ child	Birth year	Observations
	Child generation					
Child (GPA)	46.61				1988.05	575,105
	(27.96)				(4.78)	
	Parent generation					
Parents	11.61	-.05	46.82	1.99	1960.86	575,105
	(1.69)	(.50)	(9.76)	(.07)	(4.66)	
Parents' siblings	11.66	-.06	46.77	4.65	1960.87	575,105
	(1.43)	(.39)	(8.26)	(2.32)	(4.72)	
Spouses of aunts/uncles	11.77	-.01	47.22	3.86	1960.92	575,105
	(1.47)	(.38)	(8.89)	(2.03)	(6.06)	
Parent's cousins	12.14	-.02	46.05	10.35	1967.62	575,105
	(1.17)	(.34)	(7.07)	(7.56)	(3.88)	
Spouses of parents' cousins	12.15	.02	46.15	7.89	1966.14	575,105
	(1.20)	(.32)	(7.68)	(5.92)	(4.09)	
Siblings of spouses of aunts/uncles	11.67	-.05	47.32	7.94	1960.27	575,105
	(1.36)	(.35)	(7.55)	(5.63)	(6.88)	
	Grandparent generation					
Grandparents	9.24	-.16	45.67	3.88	1934.07	574,683
	(1.69)	(.38)	(7.67)	(.40)	(5.72)	
Parents' aunts/uncles	9.87	-.14	46.73	5.32	1942.07	574,683
	(1.80)	(.39)	(8.10)	(3.66)	(4.22)	
	Great grandparent generation					
Great grandparents	7.39	-.17	40.81	4.74	1910.73	397,282
	(.77)	(.60)	(8.26)	(1.92)	(4.54)	

Note: Cells show means with standard deviations in parentheses, except for the last column, which shows number of observations with non-missing data on all variables. The first row shows Grade Point Average for the child in the "years of schooling" column.

Table 2. Correlation matrices for the three different measures of parental and dynastic human capital: Years of schooling, log income, and social stratification index.

	GPA sample (N = 575,105)			Schooling sample (N = 98,052)		
	Years of schooling	Log income	Social stratification	Years of schooling	Log income	Social stratification
	Panel A: Parents					
Years of schooling	1.000			1.000		
Log income	.300	1.000		.248	1.000	
Social stratification	.519	.307	1.000	.482	.293	1.000
	Panel B: Dynasties					
Years of schooling	1.000			1.000		
Log income	.479	1.000		.428	1.000	
Social stratification	.640	.434	1.000	.634	.420	1.000

Note: Panel A shows the correlation matrix between years of schooling, lifetime incomes, and social stratification measures, averaged across parents. Panel B shows the corresponding correlation matrix for averages over parents, uncles and aunts, spouses of aunts/uncles, and parents' cousins, spouses of parents' cousins and siblings of the spouses of aunts/uncles.

Table 3. Results from OLS regressions of the child's GPA and years of schooling on dynasty's schooling, income and occupation. Successively expanding dynasties.

Outcome variable (for child):	Grade Point Average (N = 575,105)				Years of schooling (N = 98,052)			
Dependent variable (for dynasty):	Years of schooling	Log income	Social stratification	LW – scaled to years of schooling	Years of schooling	Log income	Social stratification	LW – scaled to years of schooling
Dynasty level	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(i) Parents	.359	.262	.280	.463	.283	.219	.220	.387
	(.001)	(.001)	(.001)	(.001)	(.003)	(.003)	(.003)	(.004)
(ii) +Parents' siblings and their	.449	.372	.414	.529	.364	.306	.319	.442
Spouses	(.002)	(.002)	(.002)	(.002)	(.004)	(.005)	(.005)	(.005)
(iii) +Parents' cousins and their	.509	.439	.481	.582	.410	.349	.366	.483
Spouses	(.002)	(.002)	(.002)	(.002)	(.005)	(.006)	(.006)	(.005)
(iv) +the siblings of the spouses	.525	.465	.498	.592	.419	.367	.372	.484
of the aunt and uncles	(.002)	(.002)	(.002)	(.002)	(.005)	(.006)	(.006)	(.006)

Note: Each cell shows results from a separate regression. Dependent variable is child's Grade Point Average in columns 1-4 and child's years of schooling in columns 5-8. Parental variables are averages across parents, while dynasty variables are averages over the indicated types of relatives. All regressions include linear and quadratic controls for average years of birth for each included type of relative and birth year indicators for the children. Robust standard errors in parentheses.

Table 4. Between- and within-dynasty decompositions

Dependent variable:	Grade Point Average			Years of schooling		
Dynasty level:	(ii)	(iii)	(iv)	(ii)	(iii)	(iv)
	Panel A: Years of schooling					
		$\hat{\beta} = .359$			$\hat{\beta} = .283$	
Estimated Between parameter ($\hat{\gamma}$)	.449	.509	.525	.364	.410	.419
Estimated Between weight ($\hat{k}_B$)	.628	.447	.411	.587	.395	.364
Fraction of $\hat{\beta}$ due to between variation	.784	.633	.601	.754	.572	.538
($\hat{k}_B\hat{\gamma}/\hat{\beta}$)	(.003)	(.002)	(.002)	(.007)	(.006)	(.006)
Calculated within estimate ($\hat{\beta}_W$)	.208	.238	.243	.169	.200	.206
	(.003)	(.002)	(.002)	(.005)	(.004)	(.004)
	Panel B: Log income					
		$\hat{\beta} = .262$			$\hat{\beta} = .219$	
Estimated Between parameter ($\hat{\gamma}$)	.372	.439	.465	.306	.349	.367
Estimated Between weight ($\hat{k}_B$)	.492	.320	.281	.480	.308	.270
Fraction of $\hat{\beta}$ due to between variation	.699	.537	.500	.669	.490	.452
($\hat{k}_B\hat{\gamma}/\hat{\beta}$)	(.003)	(.003)	(.003)	(.010)	(.008)	(.008)
Calculated within estimate ($\hat{\beta}_W$)	.155	.178	.182	.140	.162	.165
	(.002)	(.002)	(.002)	(.005)	(.004)	(.004)
	Panel C: Social stratification					
		$\hat{\beta} = .280$			$\hat{\beta} = .220$	
Estimated Between parameter ($\hat{\gamma}$)	.414	.481	.498	.319	.366	.372
Estimated Between weight ($\hat{k}_B$)	.504	.325	.297	.489	.305	.280
Fraction of $\hat{\beta}$ due to between variation	.743	.558	.528	.710	.507	.473
($\hat{k}_B\hat{\gamma}/\hat{\beta}$)	(.003)	(.003)	(.003)	(.009)	(.007)	(.007)
Calculated within estimate ($\hat{\beta}_W$)	.145	.184	.188	.125	.156	.161
	(.002)	(.002)	(.001)	(.005)	(.004)	(.004)
	Panel D: LW – scaled to years of schooling					
		$\hat{\beta} = .463$			$\hat{\beta} = .387$	
Estimated Between parameter ($\hat{\gamma}$)	.529	.582	.592	.442	.483	.484
Estimated Between weight ($\hat{k}_B$)	.525	.384	.357	.477	.331	.308
Fraction of $\hat{\beta}$ due to between variation	.600	.482	.456	.545	.413	.386
($\hat{k}_B\hat{\gamma}/\hat{\beta}$)	(.002)	(.002)	(.002)	(.005)	(.004)	(.004)
Calculated within estimate ($\hat{\beta}_W$)	.390	.389	.392	.336	.339	.343
	(.002)	(.002)	(.002)	(.005)	(.004)	(.004)

Note: N = 575,105 observations. Estimates of $\hat{\beta}$ and $\hat{\gamma}$ are taken from Table 3. $\hat{k}_B$ is estimated as the variance of the corresponding dynasty average, which equals the fraction of total variance due to between-dynasty variation since the averages of each type of relative have been standardized to have unit variance. The fraction due to between variation, $\hat{k}_B\hat{\gamma}/\hat{\beta}$, is calculated directly from the preceding estimates, and $\hat{\beta}_W$ is calculated as $(\hat{\beta} - \hat{k}_B\hat{\gamma})/(1 - \hat{k}_B)$. Dynasty level (ii) consists of parents, their siblings, and their spouses; level (iii) adds parents' cousins and their spouses; and level (iv) adds the siblings of the spouses of the aunts and uncles. Bootstrapped standard errors in parentheses, using 1,000 bootstrap samples.

Table 5. Horizontal GPA-schooling regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Years of schooling						
Parents	.359 (.001)	.303 (.001)	.295 (.001)	.286 (.001)	.285 (.001)	.284 (.001)	.296 (.001)
Aunts and uncles		.128 (.001)	.112 (.002)	.103 (.002)	.102 (.002)	.099 (.002)	
Spouses of aunts/uncles			.049 (.001)	.045 (.001)	.045 (.001)	.036 (.002)	
Parents' cousins				.062 (.001)	.053 (.002)	.052 (.002)	
Spouses of parents' cousins					.018 (.001)	.018 (.001)	
Siblings of spouses of aunts/uncles						.025 (.001)	
Dynasty average (excl. parents)							.228 (.002)
Sum of coefficients	.359 (.001)	.431 (.001)	.456 (.002)	.497 (.002)	.504 (.002)	.515 (.002)	.525 (.002)
R^2	.149	.162	.165	.168	.169	.169	.167
	Panel B: LW – scaled to years of schooling						
Parents	.463 (.001)	.399 (.002)	.392 (.002)	.382 (.002)	.382 (.002)	.381 (.002)	.397 (.002)
Aunts and uncles		.132 (.002)	.116 (.002)	.107 (.002)	.107 (.002)	.104 (.002)	
Spouses of aunts/uncles			.041 (.002)	.038 (.002)	.037 (.002)	.031 (.002)	
Parents' cousins				.056 (.001)	.049 (.002)	.048 (.002)	
Spouses of parents' cousins					.014 (0.002)	.014 (.002)	
Siblings of spouses of aunts/uncles						.017 (.001)	
Dynasty average (excl. parents)							.209 (.002)
Sum of coefficients	.463 (.001)	.530 (.002)	.548 (.002)	.584 (.002)	.589 (.002)	.596 (.002)	.605 (.002)
R^2	.184	.195	.197	.200	.200	.200	.197

Note: Each column shows results from a separate regression. N = 575,105 observations. Dependent variable is child's Grade Point Average. Parental generation variable is years of schooling in Panel A, and the LW index of years of schooling, log income, and social stratification in Panel B. Robust standard errors in parentheses.

Table 6. Multigenerational GPA regressions

	Years of schooling				LW – scaled to years of schooling			
Ancestor vars.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Ancestors averages							
Parents	.359 (.001)		.324 (.001)	.326 (.002)	.463 (.001)		.426 (.002)	.436 (.002)
Grandparents		.208 (.001)	.077 (.001)	.076 (.002)		.252 (.001)	.071 (.002)	.076 (.002)
Great grandparents				.006 (.002)				.002 (.002)
R^2	.149	.076	.158	.154	.463	.252	.497	.514
10 gen. pred.		.456	.477	.493		.502	.549	.569
	Panel B: Dynasty averages							
Dynasty (parental gen.)	.525 (.002)		.472 (.002)	.478 (.003)	.592 (.002)		.539 (.003)	.550 (.003)
Dynasty (grandparental gen.)		.247 (.002)	.048 (.002)	.061 (.002)		.284 (.002)	.051 (.002)	.076 (.003)
Great grandparents				-.009 (.002)				-.016 (.002)
N	574,683	574,683	574,683	397,282	574,683	574,683	574,683	397,282

Note: Each column shows results from a separate regression. Dependent variable is child's Grade Point Average. Ancestor variable is years of schooling in columns 1-4, and the LW index of years of schooling, log income, and social stratification in columns 5-8. In Panel A, variables are averaged across the direct ancestors in each generation, while in Panel B, variables are averaged across all available types of relatives in each generation. The last row in Panel A shows the AR(1) coefficient that would produce the same intergenerational persistence as the estimated model after 10 generations. Robust standard errors in parentheses.

Table 7. Adoption results GPA on years of schooling

	Years of schooling				LW – scaled to years of schooling			
Ancestor vars.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Dynasty averages							
Parents	.093 (.027)				.126 (.031)			
+Parents' siblings and their spouses		.154 (.035)				.162 (.036)		
+Parents' cousins and their spouses			.148 (.041)				.154 (.042)	
+the siblings of the spouses of the aunts and uncles				.163 (.042)				.173 (.043)
	Panel B: Multivariate							
Parents	.093 (.027)	.068 (.029)	.061 (.030)	.069 (.029)	.126 (.031)	.097 (.033)	.095 (.034)	.111 (.034)
Aunts and uncles		.064 (.032)	.045 (.035)			.097 (.036)	.085 (.040)	
Spouses of aunts/uncles			.050 (.033)				.057 (.035)	
Parents' cousins			-.013 (.027)				-.002 (.030)	
Dynasty average (excluding parents)				.088 (.046)				.094 (.046)
R^2	.136	.142	.150	.152	.140	.150	.160	.158
	Panel C: Multigenerational							
Parents	.093 (.027)		.093 (.028)		.126 (.031)		.121 (.033)	
Grandparents		.043 (.032)	-.000 (.033)			.086 (.044)	.025 (.048)	
R^2	.136	.125	.136		.140	.126	.140	

Note: Each column shows results from a separate regression. N = 975 observations. Dependent variable is child's Grade Point Average. Data is restricted to foreign-born adoptees, with an age at adoption of at most 12 months. Panel A corresponds to Table 3; Panel B to Table 5; and Panel C to Panel A of Table 6. Robust standard errors in parentheses.

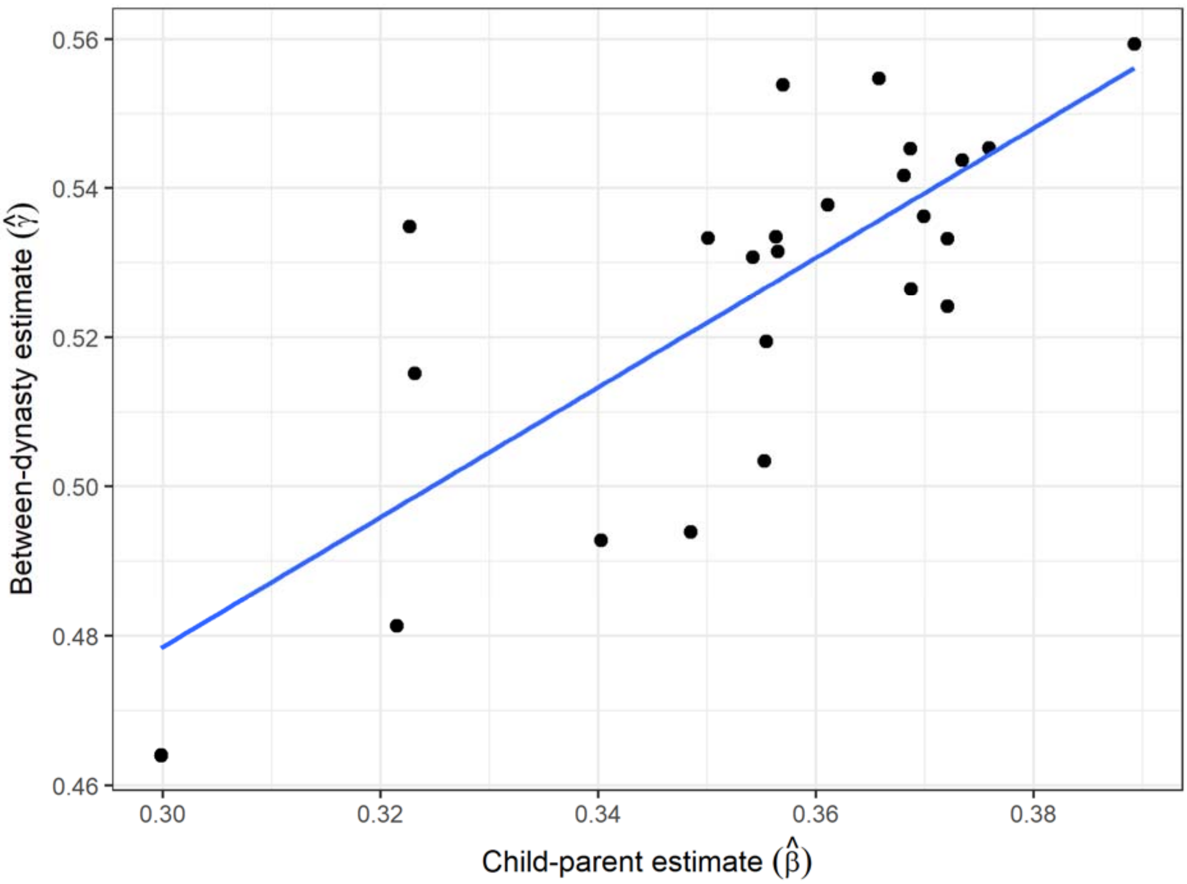

Figure 1: County-level regressions, relation between child-parent estimate and dynasty estimate

Note: The points represent estimates of $\hat{\beta}$ and $\hat{\gamma}$, corresponding to those in Table 3, for each of Sweden's 24 counties. Dependent variable is child's Grade Point Average. Number of observations in each county ranges between 5,011 and 85,181, with mean 23,851 and standard deviation 16,094.

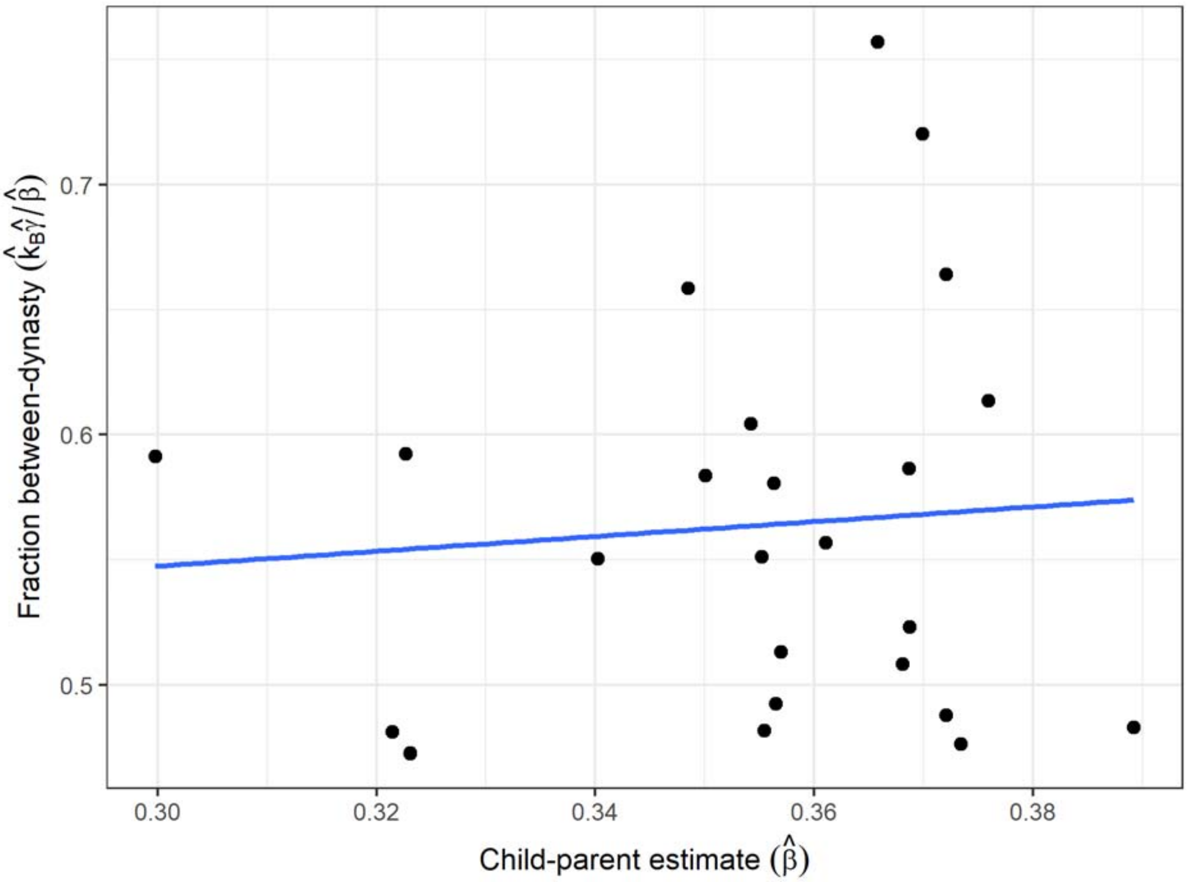

Figure 2: County-level regressions, relation between child-parent estimate and between-dynasty fraction

Note: The points represent estimates of $\hat{\beta}$ and $\hat{k}_B\hat{\gamma}/\hat{\beta}$, corresponding to those in Table 4, for each of Sweden's 24 counties. Dependent variable is child's Grade Point Average. Number of observations in each county ranges between 5,011 and 85,181, with mean 23,851 and standard deviation 16,094.

Appendix Tables: For Online Publication

Appendix Table A1. Summary statistics, schooling sample

	Years of schooling	Log income (residualized)	Social stratification	Observations/ child	Birth year	Observations
	Child generation					
Child	12.32				1979.77	98,052
	(1.96)				(2.82)	
	Parent generation					
Parents	11.00	-.13	46.27	1.99	1955.75	98,052
	(1.59)	(.49)	(9.16)	(.12)	(3.31)	
Parents' siblings	11.19	-.12	46.20	5.22	1955.84	98,052
	(1.34)	(.38)	(7.40)	(2.63)	(3.38)	
Spouses of aunts/uncles	11.39	-.06	47.02	4.37	1957.57	98,052
	(1.40)	(.36)	(8.01)	(2.30)	(5.69)	
Parent's cousins	11.93	-.05	45.77	8.46	1966.43	98,052
	(1.21)	(.37)	(7.34)	(6.60)	(3.98)	
Spouses of parents' cousins	11.97	-.01	45.99	6.58	1965.10	98,052
	(1.26)	(.35)	(8.02)	(5.24)	(4.32)	
Siblings of spouses of aunts/uncles	11.35	-.08	47.05	9.35	1957.45	98,052
	(1.31)	(.32)	(6.85)	(6.50)	(6.50)	
	Grandparent generation					
Grandparents	8.57	-.20	43.85	3.84	1929.80	97,897
	(1.39)	(.39)	(7.14)	(.45)	(4.62)	
Parents' aunts/uncles	9.45	-.17	45.64	4.06	1940.79	97,897
	(1.82)	(.44)	(8.53)	(2.95)	(3.80)	
	Great grandparent generation					
Great grandparents	7.24	-.16	39.91	3.50	1909.65	50,221
	(.65)	(.66)	(8.65)	(1.60)	(3.90)	

Note: Cells show means with standard deviations in parentheses, except for the last column, which shows number of observations with non-missing data on all variables.

Appendix Table A2. Horizontal schooling-schooling regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Years of schooling						
Parents	.283 (.003)	.245 (.003)	.240 (.003)	.235 (.003)	.235 (.003)	.235 (.003)	.246 (.003)
Aunts and uncles		.101 (.003)	.092 (.004)	.087 (.004)	.087 (.004)	.085 (.004)	
Spouses of aunts/uncles			.037 (.004)	.035 (.004)	.035 (.004)	.031 (.004)	
Parents' cousins				.046 (.003)	.047 (.004)	.046 (.004)	
Spouses of parents' cousins					-.001 (.004)	-.001 (.004)	
Siblings of spouses of aunts/uncles						.013 (.004)	
Dynasty average (excl. parents)							.174 (.005)
Sum of coefficients	.283 (.003)	.346 (.004)	.369 (.004)	.403 (.005)	.403 (.005)	.408 (.005)	.420 (.005)
R^2	.099	.108	.111	.113	.113	.114	.111
	Panel B: LW – scaled to years of schooling						
Parents	.387 (.004)	.341 (.004)	.336 (.004)	.331 (.004)	.332 (.004)	.331 (.004)	.346 (.004)
Aunts and uncles		.108 (.004)	.099 (.004)	.094 (.004)	.094 (.004)	.093 (.004)	
Spouses of aunts/uncles			.029 (.004)	.027 (.004)	.027 (.004)	.026 (.004)	
Parents' cousins				.043 (.003)	.045 (.004)	.044 (.004)	
Spouses of parents' cousins					-.004 (.004)	-.005 (.004)	
Siblings of spouses of aunts/uncles						.006 (.004)	
Dynasty average (excl. parents)							.160 (.006)
Sum of coefficients	.387 (.004)	.449 (.004)	.464 (.004)	.495 (.005)	.493 (.005)	.495 (.006)	.506 (.005)
R^2	.128	.136	.138	.140	.140	.140	.137

Note: Each column shows results from a separate regression. N = 98,052 observations. Dependent variable is child's years of schooling. Parental generation variable is years of schooling in Panel A, and the LW index of years of schooling, log income, and social stratification in Panel B. Robust standard errors in parentheses.

Appendix Table A3. Horizontal GPA-schooling regressions, schooling sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Years of schooling						
Parents	.286 (.003)	.245 (.003)	.239 (.003)	.234 (.003)	.234 (.003)	.233 (.003)	.243 (.003)
Aunts and uncles		.106 (.003)	.093 (.004)	.088 (.004)	.087 (.004)	.084 (.004)	
Spouses of aunts/uncles			.047 (.004)	.045 (.004)	.045 (.004)	.038 (.004)	
Parents' cousins				.052 (.003)	.048 (.004)	.048 (.004)	
Spouses of parents' cousins					.007 (.004)	.007 (.004)	
Siblings of spouses of aunts/uncles						.021 (.004)	
Dynasty average (excl. parents)							.198 (.005)
Sum of coefficients	.286 (.003)	.351 (.004)	.380 (.004)	.418 (.005)	.421 (.005)	.431 (.005)	.441 (.005)
R^2	.102	.112	.115	.118	.118	.118	.116
	Panel B: LW – scaled to years of schooling						
Parents	.404 (.004)	.354 (.004)	.348 (.004)	.342 (.004)	.342 (.004)	.342 (.004)	.355 (.004)
Aunts and uncles		.116 (.004)	.101 (.004)	.095 (.004)	.095 (.004)	.093 (.004)	
Spouses of aunts/uncles			.042 (.004)	.040 (.004)	.040 (.004)	.035 (.004)	
Parents' cousins				.048 (.003)	.046 (.004)	.045 (.004)	
Spouses of parents' cousins					.004 (.004)	.004 (.004)	
Siblings of spouses of aunts/uncles						.014 (.004)	
Dynasty average (excl. parents)							.186 (.006)
Sum of coefficients	.404 (.004)	.470 (.004)	.491 (.004)	.525 (.005)	.527 (.005)	.404 (.004)	.541 (.005)
R^2	.135	.144	.146	.148	.148	.135	.146

Note: Each column shows results from a separate regression. N = 98,052 observations. Dependent variable is child's Grade Point Average, but using the sample for which we also observe child's years of schooling. Parental generation variable is years of schooling in Panel A, and the LW index of years of schooling, log income, and social stratification in Panel B. Robust standard errors in parentheses.

Appendix Table A4. Multigenerational regressions, schooling

	Years of schooling				LW – scaled to years of schooling			
Ancestor vars.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Ancestors averages							
Parents	.283 (.003)		.274 (.003)	.267 (.005)	.386 (.004)		.377 (.004)	.380 (.006)
Grandparents		.116 (.003)	.026 (.003)	.027 (.005)		.155 (.004)	.022 (.004)	.030 (.006)
Great grandparents				-.011 (.007)				-.009 (.008)
R^2	.100	.031	.103	.098	.128	.042	.131	.127
10 gen. pred.		.341	.344	.404[a]		.394	.426	.402
	Panel B: Dynasty averages							
Dynasty (parental gen.)	.419 (.005)		.413 (.006)	.408 (.009)	.484 (.006)		.481 (.007)	.480 (.010)
Dynasty (grandparental gen.)		.146 (.004)	.000 (.005)	.011 (.007)		.183 (.005)	-.002 (.005)	.018 (.008)
Great grandparents				-.022 (.007)				-.027 (.008)
N	97,897	97,897	97,897	50,221	97,897	97,897	97,897	50,221

Note: Each column shows results from a separate regression. Dependent variable is child's years of schooling. Ancestor variable is years of schooling in columns 1-4, and the LW index of years of schooling, log income, and social stratification in columns 5-8. In Panel A, variables are averaged across the direct ancestors in each generation, while in Panel B, variables are averaged across all available types of relatives in each generation. The last row in Panel A shows the AR(1) coefficient that would produce the same intergenerational persistence as the estimated model after 10 generations. Robust standard errors in parentheses.

[a] No solution is possible with the estimated parameters. The given prediction was calculated by predicting 7 generations ahead instead of 10.

Appendix Table A5. Multigenerational regressions, GPA, schooling sample

	Years of schooling				LW – scaled to years of schooling			
Ancestor vars.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Ancestors averages							
Parents	.286 (.003)		.270 (.003)	.264 (.005)	.404 (.004)		.386 (.004)	.393 (.006)
Grandparents		.135 (.003)	.046 (.003)	.047 (.005)		.178 (.004)	.045 (.004)	.054 (.006)
Great grandparents				-.008 (.007)				-.007 (.008)
R^2	.102	.037	.107	.101	.135	.049	.139	.135
10 gen. pred.		.367	.382	.335		.422	.476	.474
	Panel B: Dynasty averages							
Dynasty (parental gen.)	.441 (.005)		.423 (.006)	.419 (.009)	.518 (.006)		.503 (.007)	.511 (.010)
Dynasty (grandparental gen.)		.167 (.004)	.018 (.004)	.031 (.007)		.208 (.005)	.017 (.005)	.037 (.008)
Great grandparents				-.020 (.007)				-.030 (.009)
N	97,897	97,897	97,897	50,221	9,7897	97,897	97,897	50,221

Note: Each column shows results from a separate regression. Dependent variable is child's Grade Point Average, but using the sample for which we also observe child's years of schooling. Ancestor variable is years of schooling in columns 1-4, and the LW index of years of schooling, log income, and social stratification in columns 5-8. In Panel A, variables are averaged across the direct ancestors in each generation, while in Panel B, variables are averaged across all available types of relatives in each generation. The last row in Panel A shows the AR(1) coefficient that would produce the same intergenerational persistence as the estimated model after 10 generations. Robust standard errors in parentheses.

Appendix Table A6. Summary statistics, adoptees sample

	Years of schooling	Log income (residualized)	Social stratification	Observations/ child	Birth year	Observations
	Child generation					
Child (GPA)	46.34				1988.45	975
	(26.55)				(3.95)	
	Parent generation					
Parents	12.13	0.12	51.76	1.99	1956.41	975
	(2.03)	(0.35)	(9.75)	(0.08)	(3.54)	
Parents' siblings	11.93	0.02	49.32	4.36	1956.43	975
	(1.56)	(0.34)	(8.55)	(2.25)	(3.62)	
Spouses of aunts/uncles	12.06	0.05	49.82	3.79	1957.97	975
	(1.64)	(0.34)	(8.94)	(1.97)	(5.60)	
Parent's cousins	12.31	-0.01	47.04	7.48	1966.33	975
	(1.35)	(0.38)	(7.66)	(5.63)	(4.06)	
Spouses of parents' cousins	12.29	0.03	46.99	5.87	1965.09	975
	(1.33)	(0.36)	(8.37)	(4.49)	(4.28)	
Siblings of spouses of aunts/uncles	11.82	-0.04	48.80	7.65	1957.87	975
	(1.55)	(0.34)	(7.56)	(5.28)	(6.74)	
	Grandparent generation					
Grandparents	9.17	-0.09	46.51	3.88	1929.75	974
	(1.74)	(0.36)	(7.85)	(0.41)	(4.55)	

Note: Cells show means with standard deviations in parentheses, except for the last column, which shows number of observations with non-missing data on all variables. The first row shows Grade Point Average for the child in the "years of schooling" column.

Appendix Table A7. Test of quasi-randomization of adopted children to rearing nuclear and extended families. Children adopted within a year

	No fixed effects			Region-of-birth fixed effects		
	Years of schooling	Log income	Social stratification	Years of schooling	Log income	Social stratification
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Dependent variable is child gender					
Parents	-.023	-.020	-.015	-.014	-.009	-.003
	(.014)	(.023)	(.016)	(.014)	(.023)	(.016)
+Parents' siblings and their spouses	-.017	-.031	-.036	-.004	-.021	-.022
	(.018)	(.029)	(.022)	(.019)	(.029)	(.023)
+Parents' cousins and their spouses	-.014	.001	-.033	.001	.012	-.015
	(.022)	(.031)	(.027)	(.022)	(.031)	(.027)
+the siblings of the spouses of the aunts and uncles	-.011	-.012	-.036	.004	.001	-.019
	(.023)	(.034)	(.029)	(.023)	(.033)	(.029)
	Panel B: Dependent variable is adoption age					
Parents	-.039	.231	.017	-.102	.211	-.054
	(.078)	(.118)	(.084)	(.073)	(.115)	(.080)
+Parents' siblings and their spouses	-.083	-.104	.106	-.209	-.174	-.015
	(.104)	(.165)	(.119)	(.099)	(.151)	(.116)
+Parents' cousins and their spouses	-.063	-.149	.058	-.200	-.180	-.091
	(.122)	(.172)	(.152)	(.116)	(.165)	(.145)
+the siblings of the spouses of the aunts and uncles	-.058	-.096	.055	-.203	-.156	-.091
	(.129)	(.189)	(.157)	(.122)	(.180)	(.151)

Notes: Each cell shows results from a separate regression. N = 957 observations. Data is restricted to foreign-born adoptees, with an age at adoption of at most 12 months. Dependent variable is child's gender in Panel A, and child's age at adoption in Panel B. Parental variables are averages across parents, while dynasty variables are averages over the indicated types of relatives. All regressions include linear and quadratic controls for average years of birth for each included type of relative and birth year indicators for the children, and columns 4-6 include fixed effects for region-of-birth. Robust standard errors in parentheses.

Appendix Table A8. Region F.E. GPA for children, Schooling for parents

	Main	Parish F.E.	Municipality F.E.
	(1)	(2)	(3)
Parents	.359	.355	.360
	(.001)	(.002)	(.002)
+parents' siblings and their spouses	.450	.446	.452
	(.002)	(.003)	(.003)
+parents' cousins and their spouses	.510	.502	.511
	(.002)	(.003)	(.004)
+the siblings of the spouses of the aunts and uncles	.526	.519	.528
	(.002)	(.003)	(.004)

Note: Each cell shows results from a separate regression. N = 572,414 observations. Dependent variable is child's Grade Point Average. Parental variables are averages across parents, while dynasty variables are averages over the indicated types of relatives. Fixed effects for mother's parish of residence are included in Column 2, and for mother's municipality of residence in Column 3, both from the 1985 census. Sweden had 2,579 parishes and 284 municipalities at this time. All regressions include linear and quadratic controls for average years of birth for each included type of relative and birth year indicators for the children. Robust standard errors in parentheses.