Dynastic Human Capital, Ethnic Capital and Intergenerational Mobility of Immigrants*

Adrian Adermon[†] Mikael Lindahl[‡] Mårten Palme[§]
Jonatan Riberth[¶]

January 15, 2026

Abstract

We revisit Borjas' (1992) question of whether slow intergenerational convergence among immigrant groups reflects family transmission or "ethnic capital." We extend Borjas' model by applying recently developed methods designed to better capture the transmission of broad family background—dynastic human capital (Adermon et al., 2021). Using Swedish population registers linking three generations, we can map extended family networks for each third-generation immigrant. We also examine whether initial neighborhood placement and heterogeneity in the transmission of dynastic human capital across immigrant groups contribute to persistence. Our results show that family links accounts for most of the long-run transmission of educational outcomes.

*We are grateful for comment given on seminars at the Department of Economics at Stockholm University, Uppsala University, University of Gothenburg, IFAU in Uppsala, VATT in Helsinki, the 2023 FEW summer meeting in Uppsala, the 2024 Uppsala Immigration Lab/Urban Lab workshop at Rånäs, the 2024 ESPE conference in Rotterdam, the 2024 Workshop on Wealth Inequality, Social Mobility, and Equality of Opportunity in Vienna, the 2025 Workshop on the Intergenerational Mobility of Immigrants in Princeton, University of Luxembourg, University of Duisburg-Essen, TU Dortmund; Workshop on Intergenerational Persistence and Inequality, EUI, Florence, and the BI Norwegian Business School in Oslo. Adermon gratefully acknowledges financial support from the Swedish Research Council for Health, Working Life and Welfare.

[†]IFAU, Uppsala. adrian.adermon@ifau.uu.se.

[‡]Department of Economics, University of Gothenburg. mikael.lindahl@economics.gu.se.

[§]Department of Economics, Stockholm University. marten.palme@su.se.

[¶]IIES, Stockholm University. jonatan.riberth@iies.su.se.

1 Introduction

Immigration has reshaped the demographic and economic landscape of Europe and other advanced economies in recent decades. As of 2020, 281 million individuals—around 3.6 percent of the global population—were international migrants. In the European Union, 9.4 percent of the population was foreign-born in 2010, two-thirds from outside the EU. With ongoing geopolitical conflicts and climate change, migratory pressures are likely to increase further.

Persistent group-level disparities in education and labor-market outcomes among immigrants and their descendants remain a central concern for policymakers and scholars (OECD, various; OECD/European Union, 2018). Long-run integration is critical to the sustainability of European welfare states, yet systematic evidence on outcomes beyond the second generation remains scarce.

Sweden provides a particularly suitable setting to study intergenerational human-capital assimilation—both between immigrants and natives and across immigrant groups of different origins—for at least two reasons. First, because Sweden remained neutral during World War II, large-scale immigration began shortly after the war, making it possible to observe third-generation outcomes today. Second, Sweden's administrative registers, notably the *Multi-Generation Register* linking the entire population across generations, combined with detailed education and income data, permit a comprehensive empirical analysis of intergenerational assimilation.

Figure 1 illustrates the stakes by showing the slow intergenerational convergence in schooling between immigrants and natives, and across immigrant groups. First-generation gaps are large—up to 4.5 years between East Asian and Middle Eastern/North African migrants. Convergence occurs across generations but is incomplete: nontrivial differences remain in the third generation, even among individuals fully educated in Sweden.

Borjas' seminal paper on ethnic capital (Borjas, 1992) was the first to ask why the human capital outcomes of immigrant groups converge so slowly across generations—probing the background to why the melting pot allegory of American society, as he remarked in the introduction, famously never happened. The paper's main contribution was the concept of *ethnic capital*—defined as the social, cultural,

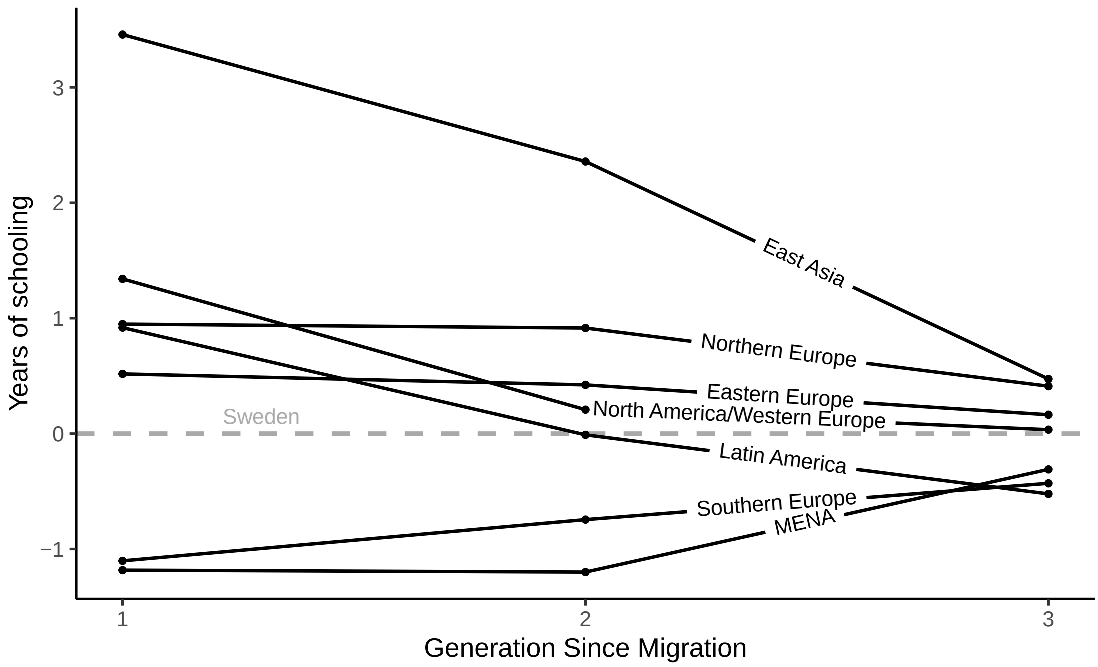

Figure 1: Convergence Between Ethnic Groups

Notes: Average years of schooling relative to Swedes born the same year for different origin groups. 1 = Migrants, 2 = Children, 3 = Grandchildren. Second- and third-generation individuals are weighted by ancestral country group. Sample restricted to dynasties where second-generation immigrants were born in Sweden or arrived before age 9.

and economic factors that are shared among an ethnic group, and measured as the average human capital of the group—which Borjas incorporated sequentially into both his economic and empirical models. The empirical results revealed a remarkably strong role for ethnic capital in explaining intergenerational persistence in human capital: depending on the specification, it accounted for between 28 and 47 percent of the total intergenerational persistence in educational attainment.

In this paper, we revisit the central question posed by Borjas (1992) about the slow intergenerational convergence in human capital between ethnic groups. We extend the *dynastic human capital* model of Adermon et al. (2021) to incorporate *ethnic capital*, thereby integrating two influential strands of research: (i) the ethnic-capital literature following Borjas (1992), and (ii) recent work modeling intergenerational mobility through latent human-capital factors (Braun and Stuhler, 2018; Clark, 2014; Collado et al., 2023; Stuhler, 2014). This synthesis enables a more complete assessment of the relative roles of families and ethnic communities in shaping intergenerational assimilation.

We also explore two potential amplifiers of persistence among immigrants: (i) stronger family transmission if limited institutional access increases reliance on kin networks, and (ii) persistent neighborhood effects from initial settlement patterns. We address these mechanisms empirically by estimating the model separately for immigrant groups and by controlling for neighborhood fixed effects.

Our empirical analysis uses population-wide administrative data covering all individuals born in Sweden between 1968 and 2006. We link each individual to parents, grandparents, and extended family members, defining ethnic origin by grandparents' country of birth. Our main outcome is the grade-point average (GPA) at the end of compulsory schooling—a standard proxy for human capital (e.g., Adermon et al., 2021). We also observe parental education and arrival neighborhoods, enabling consistent comparisons across generations and regions.

Our findings show that dynastic human capital overwhelmingly dominates in shaping third-generation outcomes. In our preferred specification, parental education explains roughly two-thirds of the observed group-level persistence, extended-family factors another third, while ethnic capital plays at most a negligible role. Adjusting for measurement error raises the ethnic-capital share slightly but it remains statistically insignificant. Hence, what has often been attributed to group-level ethnic capital appears largely driven by intergenerational transmission within families. Group heterogeneity and neighborhood influences exist but do not materially alter this conclusion.

These results imply that persistent disparities in immigrant integration cannot be attributed to enduring group norms or neighborhood spillovers. Instead, dynastic channels—operating through the nuclear and extended family—are central. Policies premised on automatic convergence across generations in egalitarian welfare states may therefore be overly optimistic.

Our paper contributes to three literatures. First, it revisits the ethnic-capital hypothesis (Borjas, 1992), complementing mixed evidence from Canada (Aydemir et al., 2009; Sweetman and Dicks, 1999), Switzerland (Bauer and Riphahn, 2007), and Denmark (Nielsen et al., 2003). Second, we connect to research on the evolution of native–immigrant gaps in education and income (Abramitzky et al., 2021;

Chetty et al., 2020). The findings from comparing outcomes between immigrants and natives, and between immigrant groups, from the rather small literature using data on three generations are mixed (Gielen and Webbink, 2025; Hammarstedt and Palme, 2012; Prokic-Breuer et al., 2024; Ward, 2020; Zhao and Drouhot, 2024; Zorlu and van Gent, 2024). Third, we build on the latent-human-capital literature (Adermon et al., 2021; Braun and Stuhler, 2018; Clark, 2014; Stuhler, 2014), extending it to immigrant dynasties where extended-family effects may be especially strong and measurement challenges particularly severe.

The remainder of the paper is organized as follows. Section 2 presents the empirical framework. Section 3 describes the data and institutional context. Section 4 reports the main findings, and Section 5 concludes.

2 Empirical Framework

Our aim is to measure the rate of human capital convergence between immigrants and natives and decompose this parameter into its underlying components. We focus on descendants of immigrants, and define immigrant groups by country of origin. The rate of intergenerational convergence between ethnic groups is captured by the parameter α_1 in the model

$$\bar{y}_e^c = \alpha_0 + \alpha_1 \bar{y}_e^p + u_e, \tag{1}$$

where $\bar{y}_e^c$ and $\bar{y}_e^p$ are ethnic group averages of outcome y in the child and parent generations, respectively. A value of $0 \leq \alpha_1 < 1$ indicates convergence between groups, while $\alpha_1 > 1$ indicates divergence—i.e., ethnic groups drifting apart over time. Figure 1 shows evidence of convergence among immigrants to Sweden. In the following, we discuss how we estimate α_1, and how we can learn about its underlying components.[1]

[1] Since our framework takes the individual as the unit of observation, and since ethnic groups e vary in size, we work with a weighted version of equation (1)—or equivalently, an individual-level regression where we replace $\bar{y}_e^c$ by y_i^c.

2.1 Extending the Borjas' Ethnic Capital Model

We build on Borjas (1992), who estimated the model

$$y_i^c = \gamma_0 + \gamma_1 \bar{y}^p_{f(i)} + \gamma_3 \bar{y}^p_{e(i)} + \varepsilon_i \tag{2}$$

where y_i^c is (a measure of) human capital for individual i; $\bar{y}^p_{f(i)}$ is average human capital for individual i's parents, indexed as family f; and $\bar{y}^p_{e(i)}$ is average human capital in individual i's ethnic group, $e(i)$.[2] Taking ethnic group averages of equation (2), we see that $\alpha_1 = \gamma_1 + \gamma_3$, allowing for the role of ethnic capital in explaining the slow convergence between immigrant groups.

Since we can map out the extended families of all individuals included in our sample, we are able to use methods similar to that proposed by Adermon et al. (2021) to incorporate channels from the human capital of the extended family and, key for our purpose, to purge the role of the ethnic group from the role of the extended family for intergenerational transmission.[3] This approach can be generalized to include any set of nested averages. We extend equation (2) to also incorporate the role of the extended family:

$$y_i^c = \gamma_0 + \gamma_1 \bar{y}^p_{f(i)} + \gamma_2 \bar{y}^p_{d(i)} + \gamma_3 \bar{y}^p_{e(i)} + \varepsilon_i, \tag{3}$$

where $\bar{y}^p_{f(i)}$ is the average education of the parents, $\bar{y}^p_{d(i)}$ is the average education of the members of the extended family (denoted as d for "dynasty") in the parents' generation, and $\bar{y}^p_{e(i)}$ is the average education of the members of the ethnic group in the parental generation.[4]

Assuming that $i \subset f \subset d \subset e$, we can again average equation (2) by ethnic

[2]The functions $f(i)$ and $e(i)$ assign to each individual i a family f and an ethnic group e, respectively.

[3]A key difference is that Adermon et al. (2021) are interested in estimating the transmission of latent human capital from parents to child, while we want to understand the components of a group-level convergence parameter. Because of this difference in purpose, we also handle the standardization of extended family variables differently: While Adermon et al. (2021) standardize each group average to have unity standard deviation, we allow the different group averages to have different standard deviations.

[4]The argument that ignoring other family members is likely to bias the coefficient for ethnic capital upwards is not new—see Card et al. (2000) and Dustmann and Glitz (2011).

group to obtain

$$\bar{y}_e^c = \gamma_0 + (\gamma_1 + \gamma_2 + \gamma_3)\bar{y}_e^p + \bar{\varepsilon}_e. \tag{4}$$

This approach thus allows us to decompose group level convergence into three channels: the *nuclear family*, the *extended family*, and the *ethnic group*.

Due to intermarriage, members of the same extended family may belong to different ethnic groups, making the assumption of perfectly nested groups ($f \subset d \subset e$) unlikely to hold in the data. In this case equation (4), where we have $\alpha_1 = \gamma_1 + \gamma_2 + \gamma_3$, does not hold exactly. Since a large share of parent couples are formed outside the ethnic group (mostly natives, sometimes immigrants from other ethnic groups), our main analyses include measures of human capital for the parents and the extended family that includes these individuals.

To take into account that groups can be non-nested, we also decompose the group level convergence parameter α_1 using the approach outlined in Gelbach (2016). Using the coefficients from the individual-level version of the baseline model from equation (1), and the full model from equation (3), the underlying sources of the intergenerational group persistence parameter can decomposed as:

$$\underbrace{\alpha_1}_{\substack{\text{ethnic group}\\\text{convergence}}} = \underbrace{\gamma_1\delta_f}_{\substack{\text{parents'}\\\text{human capital}}} + \underbrace{\gamma_2\delta_d}_{\substack{\text{dynastic}\\\text{human capital}}} + \underbrace{\gamma_3}_{\substack{\text{ethnic}\\\text{human capital}}}, \tag{5}$$

where δ_f is the coefficient from regressing $\bar{y}_{f(i)}^p$ on $\bar{y}_{e(i)}^p$, and δ_d is the coefficient from regressing $\bar{y}_{d(i)}^p$ on $\bar{y}_{e(i)}^p$. Note that if the observations are nested, then $\delta_f = \delta_d = 1$, and we are back to the simple model we discussed above, where $\alpha_1 = \gamma_1 + \gamma_2 + \gamma_3$.

2.2 A Latent Variables Model

If observed education in the parental generation fails to fully capture the human capital that is transmitted to children, the models discussed so far can yield misleading estimates. We examine this by specifying a latent variables model, and use observable proxies to get alternative estimates of the parameters in this model.[5]

[5]This type of latent variable model has been discussed by, e.g., Adermon et al. (2021), Braun and Stuhler (2018), Clark (2014), Collado et al. (2023), and Stuhler (2014).

Specifically, let the observed child outcome be determined as

$$y_i^c = \beta_0 + \beta_1 y_{f(i)}^{p*} + \beta_2 y_{d(i)}^{p*} + \beta_3 y_{e(i)}^{p*} + \varepsilon, \tag{6}$$

where $y_{f(i)}^{p*}$ is the parents' latent human capital, $y_{d(i)}^{p*}$ latent human capital of the extended family, and $y_{e(i)}^{p*}$ latent human of the ethnic group. While these latent variables are fundamentally unobservable, we observe a set of proxy variables $y_j^p = y_j^{p*} + v_j$, for $j \in \{f, d, e\}$, where v^j is an i.i.d. measurement error term. We can view equation (3) above as an empirical version of this model, where we use average level of education among parents, the extended family, and the ethnic group as proxies for the corresponding latent variables.

If we have access to multiple proxies, y_{jk}^p, for y_j^{p*}, we can improve our estimates by using the approach proposed by Lubotsky and Wittenberg (2006).[6] This approach proceeds in three steps: first, we estimate an extended model by regressing child outcome on the full set of proxies,

$$y_i^c = \sum_j \sum_k b_{jk} y_{ijk}^p + \varepsilon_i. \tag{7}$$

We then calculate weighted averages of the coefficients from equation (7) as

$$\hat{\beta}_j = \sum_k \rho_{jk} \hat{b}_{jk}, \tag{8}$$

where the weights are functions of covariances $\rho_{jk} = \frac{\text{Cov}(y^c, y_{jk}^p)}{\text{Cov}(y^c, y_{j1}^p)}$. Notice that the weights are normalized to the scale of one of the proxies (y_{j1}^p)—we normalize to years of schooling, so that the Lutbotsky-Wittenberg (LW henceforth) coefficients can be interpreted on the same scale as our main years of schooling regressions.[7]

The additional proxies we use are measures of mid-life labor income and an occupation-based index of social status (see Section 3.4). This means that our resulting estimates of the relationship between child's GPA and parents', extended

[6] For each latent variable j, the measurement errors for each proxy k are assumed to be uncorrelated with the latent variable, but are allowed to be correlated with each other.

[7] The Lubotsky and Wittenberg estimator has previously been used to improve estimates of intergenerational mobility by, e.g., Adermon et al. (2021), Vosters (2018), and Vosters and Nybom (2017).

families', and ethnic groups' "human capital" (or social status) should be interpreted as capturing broader channels not captured by years of schooling, including employers' valuation of human capital (including discrimination), individuals' occupational choices, and job networks.

2.3 Other Specification Issues

In the models above, we include additional control variables so as to make immigrants from different groups comparable on other dimensions. Our thought experiment is that immigrant families should be similar when they arrive in Sweden, except when it comes to their human and cultural capital. In our main specifications, we focus on parents who were born in Sweden. To account for variation in time of arrival, we control for individual birth year fixed effects as well as anchoring immigrants' human capital levels to those of natives by subtracting the mean human capital of natives born in the same year.[8]

As several studies have pointed out (e.g., Borjas, 1995; Bratu and Bolotnyy, 2023; Ward, 2020), endogenous geographical sorting might play an important role in explaining patterns of intergenerational mobility among immigrant groups. Using detailed data on the *first* neighborhood at arrival for immigrant ancestors, we are able to control for endogenous sorting of immigrant groups to regions with more or less potential for integration at the time of arrival. We do this by including fixed effects for first location.

2.4 Robust and efficient inference

Before the rise of robust inference, correlated error terms were typically handled by explicitly modeling the correlation structure and adjusting the estimator using, e.g., generalized least squares (Angrist and Pischke, 2010). The problem with this approach is that if the model for the error term is misspecified, standard errors will be biased. Robust inference solves this problem by allowing for arbitrary forms of heteroskedasticity (White, 1980) or clustering (Arellano, 1987; Liang and Zeger,

[8]The latter is important since average years of schooling of natives are trending upwards during this period. This also makes our estimates directly interpretable as measuring convergence relative to the native group.

1986). However, this robustness often comes a the cost of reduced statistical power, which has been shown to be a pervasive issue in empirical economics (Ioannidis et al., 2017).[9]

As pointed out by Wooldridge (2003, 2010), robust inference can be combined with explicit models of the error.[10] This leads to gains in statistical power if the error model captures something about the underlying structure, while protecting from size distortion if the error model is misspecified. The approach has been shown to lead to potentially large gains in power in the case of heteroskedasticity (Romano and Wolf, 2017) and difference-in-differences (Brewer et al., 2018), without compromising robustness.[11]

Cameron and Miller (2015) discuss this approach for clustered data, recommending the use of a random effects model estimated using FGLS paired with cluster-robust inference. They write "It is remarkable that current econometric practice with clustered errors ignores the potential efficiency gains of FGLS" (p. 326).

To increase the precision of our estimates, we estimate an FGLS model with random effects at the country group level, using the method of Amemiya (1971) to estimate the covariance matrix. Standard errors are cluster-robust at the country level. FGLS is consistent and $\sqrt{N}$-consistent under weak assumptions, even when the error structure is misspecified (Wooldridge, 2010, p. 182).[12]

3 Data and Brief Historical Background

Our dataset is constructed by linking individual-level data from several Swedish administrative registers. We use three generations of linked individuals, which we

[9]The move from GLS to robust inference has also been criticized by Leamer (2010).

[10]In Wooldridge (2010), see specifically pp. 182–183, 297–298, 866–867. See also Hansen (2022, pp. 617–619).

[11]Ferman and Pinto (2019) derive a robust variance estimator for difference-in-differences estimation with few treated group, and show how it can be combined with FGLS for improved efficiency.

[12]Some researchers are uncomfortable with GLS and related approaches because they can yield point estimates that differ from OLS estimates in finite samples. Furthermore, Angrist and Pischke (2008, p. 93) prefer OLS to WLS because even with a misspecified regression model, OLS has the property of being the best linear approximation to the population conditional expectation function. For these reasons, we show OLS estimates in Table A.5.

refer to as the *child*, *parent*, and *grandparent* generations. We observe educational outcomes for the child and parent generations, while the grandparent generation is used to assign ethnic origin, arrival neighborhood, and to identify extended family relationships.[13]

Our main sample consists of children born in Sweden between 1968 and 2006, the cohorts for which grade point average (GPA) at the end of compulsory school is observed. Using the Swedish Multi-generation Register, we identify each child's grandparents, which allows us to reconstruct extended family networks—including aunts, uncles, and their spouses.

We include children who have at least one parent who: (i) was born in Sweden; (ii) is of foreign background, defined as having both parents born outside the Nordic countries; and (iii) has at least one sibling in the data, which implies that the child has at least one aunt or uncle of foreign origin.

The ethnic composition of our sample reflects historical immigration flows. Most arrivals for the grandparent (first) generation occurred between 1945 and 1980.

3.1 Migration to Sweden 1945–1980

Sweden provides a valuable setting for studying long-term migration effects, having emerged as a stable and attractive destination country for migrants after World War II. The period from 1945 to 1980 encompasses three distinct phases of migration, shaped by geopolitical developments and domestic labor market needs (Lundh, 2005; OECD, various; Swedish Migration Agency, 2025).

1945–1950s: Post-war humanitarian and Nordic migration. Sweden initially received refugees and displaced persons from the Baltic states (Estonia, Latvia, Lithuania), Finland, Germany, and Poland. Following the creation of the Nordic Passport Union in 1954, migration from neighboring Nordic countries increased substantially (Nordic Council, 1954). The Cold War and increased political repression in Eastern Europe also triggered new refugee inflows from these

[13] Siblings are identified via shared parentage. For immigrants, sibling links are typically observed only if individuals migrated to Sweden together with a parent before age 18.

countries (Byström, 2006).

1950s–1960s: Labor migration and industrial expansion. Sweden's economic boom created a demand for labor, prompting government-led recruitment—especially from Italy, Yugoslavia, Greece, and Turkey. Migrants often worked in construction and manufacturing, under regulated bilateral labor agreements (Lundh, 2005).

Late 1960s–1970s: Policy restriction and humanitarian turn. With rising unemployment in the late 1960s, Sweden imposed restrictions on labor immigration. Work permit requirements were introduced in 1967, and labor recruitment ended in 1972. The focus shifted toward family reunification and refugee admissions. Notable groups included refugees from Chile (after the 1973 coup), Uganda (expelled Asians), and Southeast Asia (post-Indochina wars). In 1975, Sweden adopted an official integration policy promoting multiculturalism and equal rights (Borevi, 2002).

Figure A.1 plots the year of immigration for ancestors from different country groups in our sample. The different immigration waves described above are all represented.

3.2 Ethnic Composition and Country of Origin Assignment

Our analysis excludes individuals of Nordic ancestry (i.e., those with grandparents from Denmark, Finland, Iceland, or Norway), as these populations are relatively similar to Sweden in terms of culture, language, and appearance.[14] Table A.1 presents the distribution of countries of origin for the grandparent generation in our dataset. The largest groups originate from other European countries, with Germany and former Yugoslavia being particularly prominent. Turkey also constitutes a major group, while the US and Chile represent the largest non-European origins.

[14] In Section 4.2, we show that our results are robust to including them.

We assign ethnic origin to children based on their grandparents' countries of birth. If all observed grandparents are born in the same country, the child is assigned that country. If grandparents are born in different non-Nordic countries, the child is assigned proportional weights for each country. For example, a child with one set of grandparents from Germany and another from Poland is included in the analysis twice—once under each origin—with a weight of 0.5 in each case.

3.3 Extended Family Structure and Ancestry Composition

Our focus on third-generation immigrants allows us to study long-term outcomes without conditioning on parental marital patterns. In many cases, children have one native-born parent and one foreign-born parent. Figure A.2 displays the distribution of foreign and native ancestry across children in the sample.

For nearly all children in the dataset, we observe both parents. Approximately 86% of children are assigned a single country of origin, while 13% are assigned two countries, and fewer than 1% are assigned three or more. This weighting scheme enables us to include children from mixed-background families without losing statistical representativeness or comparability across groups.

3.4 Variable Definitions

For the child generation, we measure human capital using GPA. For the parental generation, we use years of schooling, income, and an index of social stratification. These are defined as follows.

Grade Point Average (*GPA*) for all compulsory subjects is measured in ninth grade, at the end of compulsary schooling (typically at age 16). We percentile rank the variable within year.

Years of schooling is computed based on educational attainment in administrative data between 1985–2020 and census data from 1960, 1970 and 1990. Individuals who have an education from another country than Sweden report their education to the migration authority upon arrival to Sweden.

Lifetime *income* includes unemployment insurance and sickness benefits. We first regress yearly log incomes on a full set of gender, birth year, and income year controls using the full income panel during ages 35–55. Lifetime income is then calculated by averaging the residuals from this regression for each individual.

Finally, we measure *social stratification* using the Swedish version of the occupation-based CAMSIS index (Lambert and Bihagen, 2012).[15] We link the index using the individual's occupation at or as close as possible to age 50.

We calculate three sets of group averages of the human capital variables in the parental generation. First, *parental* outcomes are simply averaged over both parents. Second, we define the *extended family* as the child's aunts and uncles, their spouses, and those spouses' siblings. Siblings are defined as individuals sharing both biological parents. Third, the *ethnic mean* captures the average outcome for individuals in the parental generation with a given county-of-origin ancestry. For each parent, the ethnic mean is defined as average human capital among individuals who (i) have the same ancestry as the parent,[16] (ii) are born within five years of the parent, (iii) are not members of the child's extended family, (iv) are born in Sweden or moved to Sweden before age 9.[17]

The ethnic mean assigned to the child is then the mean of the two parents' ethnic means.[18] By focusing on individuals that are born in Sweden or came before age nine, we mitigate issues related to differences in age of arrival to Sweden between different country groups.

We also define a child's *neighborhood* as the first parish that a non-Swedish ancestor lived in when they first came to Sweden.

Table A.2 shows summary statistics for our main variables.

4 Results

This Section presents our results. The results from the core specification (see Section 2), measuring the excess intergenerational persistence from Ethnic and

[15]CAMSIS measures social distance based on the occupations of married couples.

[16]Ancestry is defined as grandparents' country of birth or own country of birth if the individual is born outside Sweden.

[17]Böhlmark (2008) shows that arriving after age nine has a strong negative impact on school performance.

[18]If a parent has two ethnic origins (e.g., if the maternal grandparents were born in two different countries), we assign the parent one of these origins at random, and then proceed to calculate their ethnic mean.

Dynastic human capital, are presented in Section 4.1. Section 4.2 assess the robustness of our key findings, while Section 4.3 presents separate estimates by country-of-origin groups.

4.1 Parents, Extended Family and Ethnic Capital

Figure 2 illustrates the paper's starting point: the slow convergence in human capital between ethnic groups. It plots average years of schooling in the second generation (parents) against the third generation (children), by country-of-origin group (corresponding to equation (1)). Origins are defined by the first-generation immigrants (grandparents). The upward-sloping line shows that group differences persist into the third generation.

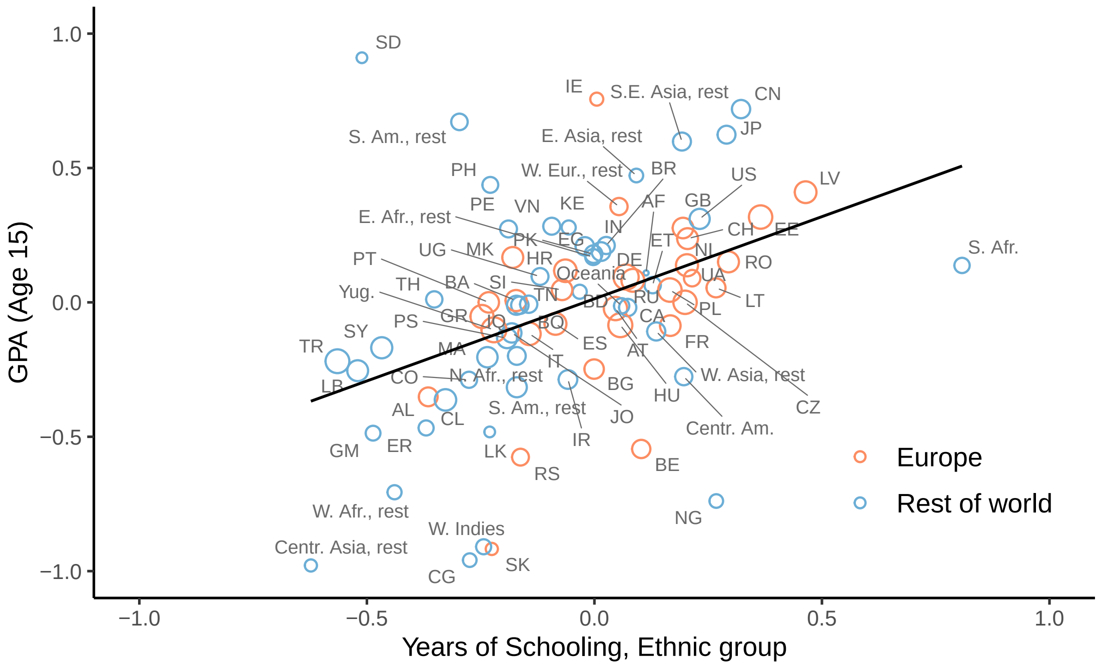

Figure 2: Country-Level Scatter Plot

Notes: This figure plots average GPA against average ethnic group means by country of origin. Circle sizes are proportional to the logarithm of the number of children in our sample from each country group. Nordic countries are excluded.

This group-level intergenerational association is substantially stronger than what is typically found in individual-level data for Sweden. The results suggest that over 60 percent of the average educational advantage observed in the second generation persists to the third generation. To decompose this persistence in educational outcomes into contributions from parents, the extended family, and the ethnic group, Table 1 presents results from a series of regression models. Each column reports estimates from a separate specification in which a child's educational outcome is regressed on different combinations of human capital variables measured at the parental, extended family, and ethnic-group levels. All models include fixed effects for birth year (to account for cohort effects), grandparents' initial neighborhood (to address potential geographic sorting), and grandparents' year of migration to Sweden.

For interpretability, the coefficient on parental years of schooling is standardized to reflect changes in standard deviations (SD). For example, in column 2, a one SD increase in parental schooling is associated with an 0.4 SD increase in child GPA. The coefficients for the extended family and ethnic group schooling variables are also rescaled by the standard deviation of parental years of schooling to ensure direct comparability between the coefficients in the estimated model. Standard errors are clustered at the country of origin level.

Column 1 mirrors the approach in Figure 2, but applies it to individual-level data rather than group-level data. As expected, the results confirm strong intergenerational persistence, with the GPA of the child closely related to the average GPA of their ethnic group in the parental generation. Column 2 estimates a Markovian AR(1) model, widely used in research on intergenerational mobility since Becker and Tomes (1986). The point estimate of 0.44 is in line with previous Swedish studies (e.g., Adermon et al., 2021).

Column 3 presents results from the canonical model of Borjas (1992) (see equation (2)), which includes the (leave-one-out) average educational attainment among individuals in the parental generation from the same country of origin. The coefficient on parental education remains similar to that in column 2, while the estimate for ethnic capital is moderate, substantially below the levels reported in some earlier studies using data on three generations (e.g., Borjas, 1992; Ward, 2020) that found ethnic capital to be comparable in magnitude to the coefficient for parental human capital. Although the coefficient is imprecisely estimated, the data rule out similarly large ethnic effects in this context.

Table 1: GPA schooling regressions

	(1)	(2)	(3)	(4)	(5)
Parents		0.433	0.430	0.362	0.361
		(0.012)	(0.012)	(0.007)	(0.007)
Extended family				0.200	0.199
				(0.016)	(0.016)
Ethnic mean	0.540		0.090		0.029
	(0.058)		(0.042)		(0.045)
Sum	0.540	0.433	0.520	0.561	0.589
	(0.058)	(0.012)	(0.041)	(0.018)	(0.043)
R^2	0.085	0.230	0.230	0.241	0.241
Num. Ind	36 305	36 305	36 305	36 305	36 305

Notes: Each column shows results from separate specifications where standardised Grade Point Average (GPA) is regressed on years of schooling. The years of schooling variables are divided by the standard deviation of years of schooling for parents. All regressions include fixed effects for birth year, age at migrating to Sweden and first parish of arrival of the ancestors. Each parental generation outcome is the average across all members of the given category of relatives. Models are estimated using FGLS to allow for country group random effects, with standard errors clustered by country group in parentheses.

Column 4 extends the baseline model by incorporating the mean schooling of extended family members, following the dynastic human capital framework of Adermon et al. (2021). This allows us to capture broader familial influences beyond the nuclear household. The results confirm that extended family education is a strong predictor of child outcomes, even after accounting for parental education, underscoring the importance of extended family networks in shaping educational attainment.

Finally, column 5 includes all three components: parental, extended family, and ethnic human capital, as in equation (3). In this full specification, the coefficient for ethnic capital becomes very close to zero, suggesting that once family-level human capital is accounted for, ethnic-group effects are no longer important for explaining differences in third-generation outcomes.

A possible alternative interpretation of these results is that ethnic capital is transmitted through the extended family, such that the extended family serves as a better proxy for ethnic capital than the ethnic mean.[19] We assess this interpretation in Table A.3 by estimating specifications that separately distinguish extended-family members of foreign and native origin. If the extended family primarily captures ethnic capital, we would expect, first, a larger coefficient for relatives of foreign origin and, second, a stronger attenuation of the ethnic-mean coefficient when controlling for foreign-origin relatives than when controlling for native-origin relatives. We find neither pattern. Instead, the coefficient on native-origin relatives is at least as large as that on foreign-origin relatives, and the ethnic-mean coefficient is nearly identical across specifications. We therefore conclude that the extended family captures family-specific factors distinct from ethnic capital.

As discussed in Section 2.1, we decompose the overall group convergence parameter into contributions from parents, the extended family (dynastic capital), and ethnic capital. We implement this in two complementary ways: the first simply sums the relevant regression coefficients to recover the share of convergence explained by each channel (sums are reported at the bottom of Table 1), while the second applies a Gelbach decomposition, which accounts for the overlapping na-

[19] One reason is that a single country of origin may comprise multiple ethnic groups, which may be more accurately captured by extended-family links.

ture of family and group characteristics by reweighting the coefficients accordingly (see equation (5)).[20]

Using the estimates in column 5, the direct approach attributes 64 percent of group convergence to parents, 34 percent to the extended family, and a negligible share to ethnic capital. The Gelbach decomposition, based on the overall ethnic mean in column 1, yields similar results: parents account for 73 percent and the extended family for 25.

Taken together, the results shown in Table 1 imply that the observed group-level differences in human capital across ethnic groups in the third generation are not driven by ethnic group specific effects, such as culture, norms or ethnic peer effects. Rather, the persistence seen in Figure 1 appears to reflect initial disparities in human capital and the strength of intergenerational transmission within families broadly defined, which have been ignored in previous studies.

4.2 Robustness of Key Results

We assess the robustness of our main estimates in five sensitivity checks:

1. *Neighborhood effects.* To examine whether our results are driven by neighborhood composition, we include fixed effects for children's neighborhoods.

2. *Inclusion of Nordic immigrants.* The main analysis excludes immigrants from other Nordic countries (Denmark, Finland, Iceland, and Norway), who could be considered as ethnically relatively similar to Swedes.[21] As a robustness check, we expand the sample to include these groups.

3. *Reweighting of country groups.* To reduce the influence of larger origin groups, we reweight each observation by the inverse of the number of individuals in the group, so that each country of origin is given the same weight in the regressions.

[20]The Gelbach weights are estimated by regressing parental and extended family education on the ethnic mean, controlling for cohort and neighborhood fixed effects.

[21]However, Finnish immigrants and their descendants—who make up the vast majority of this group—had a distinct ethnic identity and were a socially disadvantaged group in Sweden well into the 1990s (see, e.g., Koivunen, 2017; Saarela and Finnäs, 2007; Saarela and Rooth, 2006; Weckström, 2011).

4. *Multiple proxies for parental human capital.* Parental, extended family, and ethnic schooling measures are imperfect proxies for the respective group's latent human capital, as discussed in Section 2.2. Since they are averages of the same underlying variable (years of schooling in the parental generation) over different sample sizes, they also differ in the extent of measurement error.[22] which may bias the estimated relative contributions.[23]

5. *Excluding arrival neighborhood effects.* The main analysis includes fixed effects for the parish of arrival for the first ancestor to immigrate to Sweden. We check the sensitivity of our estimates by dropping these controls.

Figure 3 shows the results of each of these sensitivity checks.[24] The boxes correspond to the columns in Table 1, with rows showing the different human capital variables. Coefficients from the different sensitivity analyses outlined above are delineated using different shapes and colors.

Overall, our findings are robust across specifications. Ethnic capital plays only a modest role in explaining the slow convergence across groups, especially once dynastic human capital is incorporated into the Borjas framework.

In addition to these checks, Table A.4 reports results using alternative proxies for ethnic human capital. We consider education and cultural indices from the World Values Survey (WVS) as well as average years of schooling in the descendant's country of origin from Barro and Lee (2013) as alternative proxies for ethnic human capital. These measures yield a somewhat larger coefficient on the ethnic mean. When we combine all proxies, including the ethnic mean, the estimated contribution of ethnic human capital is very similar to that in the baseline model.

[22] If the error is classical, it will average out more in the ethnic mean than in the extended family and parental means,

[23] Ward (2020) documents this in US data. We follow the proxy-variable framework in Lubotsky and Wittenberg (2006), combining years of schooling, labor income, and occupation-based social stratification to form a broader measure of parental human capital and to mitigate attenuation bias. See Section 2.2 for details.

[24] These estimates are also presented in Table A.6.

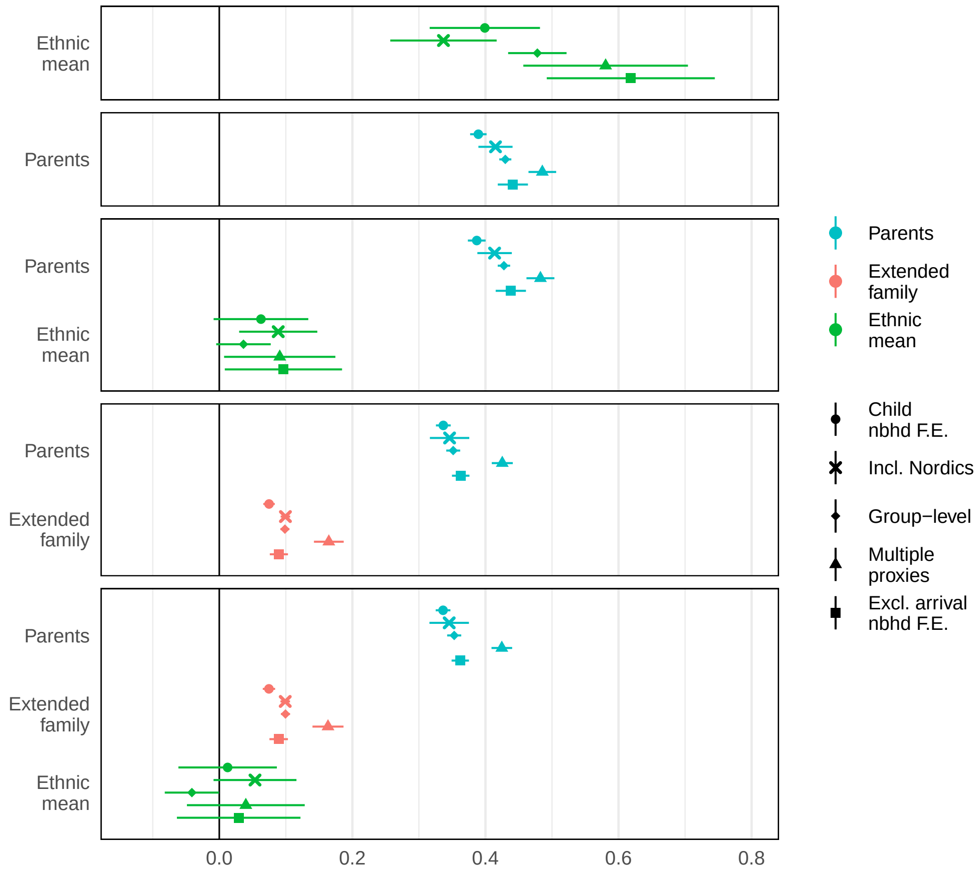

Figure 3: Robustness to alternative specifications

Notes: This figure displays results for four alternative specifications. Each subfigure corresponds to one column in Table 1. **Child nbhd. F.E.** displays coefficients where we include neighborhood fixed effects for the children. **Incl Nordics** displays results where we include children of Nordic descent. **Group-level** displays coefficients where we weight each observation by the inverse of the number of individuals in each country group. **Multiple proxies** displays coefficients based on Lubotsky-Wittenberg index constructed based on years of schooling, income and an occupation stratification index.

4.3 Between-Group Heterogeneity

Cross-country differences in intergenerational mobility are well documented (Corak, 2013; Manduca et al., 2024; Solon, 2004). Explanations point to inequality, welfare-state institutions, and cultural norms. Similar mechanisms may also operate across ethnic groups within a society. Immigrants often have weaker language skills and networks, which can limit access to public institutions for human capital formation such as schools and libraries. As a result, the family plays a larger role in human capital transmission, which may lead to stronger intergenerational persistence within ethnic groups.

Table 2 reports separate estimates of the dynastic human capital model from Adermon et al. (2021). We distinguish between parental education, the extended family, and the full family (our measure of within-group persistence). Estimates are shown for natives, all immigrants, and immigrant subgroups by region of origin. This allows us to compare how the strength of family-based transmission varies across groups.

Two results stand out. First, immigrants as a group show slightly stronger intergenerational persistence than natives ($p = 0.097$), a difference driven by the role of the extended family. Second, persistence varies markedly between origin groups: it is highest for Nordics (mainly Finland, 0.66) and lowest for Northern Europeans (0.47). Most of this variation comes from differences in the importance of the extended family: while the parental coefficient varies from 0.32 to 0.41, the extended family coefficient spans from 0.12 to 0.28. These patterns suggest that cultural factors and proximity to Swedish society may influence the strength of family-based transmission.

5 Conclusions

We study intergenerational persistence in educational outcomes among immigrant-origin families and address the longstanding question of why human capital convergence appears slower than predicted by canonical intergenerational-mobility models (AR(1) and Becker–Tomes). Using Swedish administrative data linking three generations of descendants of postwar immigrants, we map extended fam-

Table 2: GPA schooling regressions—Heterogenous transmission

	Parents	Extended Family	Parents + Extended Family
Swedish Origin	.361 (.020)	.160 (.029)	.521 (.025)
Foreign Origin	.363 (.005)	.195 (.008)	.559 (.006)
Nordics*	.405 (.004)	.251 (.007)	.656 (.006)
Eastern Europe	.369 (.011)	.245 (.015)	.614 (.012)
East Asia	.320 (.053)	.284 (.076)	.604 (.062)
North America/Western Europe	.359 (.010)	.206 (.014)	.565 (.012)
Southern Europe	.381 (.011)	.182 (.015)	.563 (.014)
Latin America	.412 (.046)	.127 (.064)	.540 (.059)
MENA	.343 (.019)	.151 (.030)	.495 (.027)
Northern Europe	.349 (.014)	.120 (.021)	.469 (.018)

Notes: Each column shows results from separate regressions. The dependent variable is standardised GPA. Column one corresponds to column 5 in Table A.5. In column 2 we allow β^p to vary by country. In column 3 we allow β^e to vary by country. In column 4 we allow both β^p and β^e to vary by country. Robust standard errors in parentheses.

ily networks into the third generation. We find that most of the residual third-generation disadvantage is accounted for by slow convergence in the population at large rather than by a distinct effect of the ethnic group's human capital. Robustness analyses indicate only modest roles for neighborhood composition and for stronger within-group persistence among immigrant families.

Our results differ markedly from those of the pioneering work by Borjas (1992), which found a strong independent effect of ethnic-group human capital on individual outcomes. We attribute these differences to methodological and institutional factors, as well as to differences in measurement. The latent-variable model employed in our study implies substantially stronger intergenerational persistence than the simple Markovian model used by Borjas, thereby absorbing some of the persistence that he attributed to ethnic-group effects. However, when we estimate the Borjas model using our data, we also find a weaker role for ethnic-group influences within that framework compared to Borjas' findings. This pattern suggests that part of the discrepancy may reflect institutional differences between the two countries studied and/or differences in measurement.

Taken together, our findings highlight the central role of the (extended) family—

rather than ethnicity or neighborhoods—in shaping intergenerational persistence. Nevertheless, further research is needed to disentangle and quantify the various channels through which the family contributes to the intergenerational transmission of human capital, such as parental investments in time and consumption, as emphasized by the Becker–Tomes model, or through parental social networks.

References

Abramitzky, R., L. Boustan, E. Jacome, and S. Perez (2021). "Intergenerational Mobility of Immigrants in the United States over Two Centuries". *American Economic Review* 111.2, pp. 580–608.

Adermon, A., M. Lindahl, and M. Palme (2021). "Dynastic Human Capital, Inequality, and Intergenerational Mobility". *American Economic Review* 111.5, pp. 1523–1548.

Amemiya, T. (1971). "The Estimation of the Variances in a Variance-Components Model". *International Economic Review* 12.1, pp. 1–13.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

— (2010). "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics". *The Journal of Economic Perspectives* 24.2, pp. 3–30.

Arellano, M. (1987). "Computing Robust Standard Errors for Within-groups Estimators". *Oxford Bulletin of Economics and Statistics* 49.4, pp. 431–434.

Aydemir, A., W.-H. Chen, and M. Corak (2009). "Intergenerational Earnings Mobility among the Children of Canadian Immigrants". *The Review of Economics and Statistics* 91.2, pp. 377–397.

Barro, R. J. and J. W. Lee (2013). "A New Data Set of Educational Attainment in the World, 1950–2010". *Journal of Development Economics* 104, pp. 184–198.

Bauer, P. and R. T. Riphahn (2007). "Heterogeneity in the Intergenerational Transmission of Educational Attainment: Evidence from Switzerland on Natives and Second-Generation Immigrants". *Journal of Population Economics* 20.1, pp. 121–148.

Becker, G. S. and N. Tomes (1986). "Human Capital and the Rise and Fall of Families". *Journal of Labor Economics* 4.3, S1–S39.

Böhlmark, A. (2008). "Age at Immigration and School Performance: A Siblings Analysis Using Swedish Register Data". *Labour Economics* 15.6, pp. 1366–1387.

Borevi, K. (2002). *Välfärdsstaten i det mångkulturella samhället: Socialdemokratisk integrationspolitik under efterkrigstiden.* Uppsala: Acta Universitatis Upsaliensis.

Borjas, G. J. (1992). "Ethnic Capital and Intergenerational Mobility". *The Quarterly Journal of Economics* 107.1, pp. 123–150.

— (1995). "Ethnicity, Neighborhoods, and Human-Capital Externalities". *The American Economic Review* 85.3, pp. 365–390.

Bratu, C. and V. Bolotnyy (2023). "Immigrant Intergenerational Mobility: A Focus on Childhood Environment". *European Economic Review* 151, p. 104353.

Braun, S. T. and J. Stuhler (2018). "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany". *The Economic Journal* 128.609, pp. 576–611.

Brewer, M., T. F. Crossley, and R. Joyce (2018). "Inference with Difference-in-Differences Revisited". *Journal of Econometric Methods* 7.1.

Byström, M. (2006). *En flykting korsar sitt spår: Folkbokföring, flyktingar och den svenska välfärdsstaten 1944–1951.* Stockholm: Almqvist & Wiksell.

Cameron, A. C. and D. L. Miller (2015). "A Practitioner's Guide to Cluster-Robust Inference". *Journal of Human Resources* 50.2, pp. 317–372.

Card, D., J. DiNardo, and E. Estes (2000). "The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s". In: *Issues in the Economics of Immigration.* University of Chicago Press, pp. 227–270.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). "Race and Economic Opportunity in the United States: An Intergenerational Perspective". *The Quarterly Journal of Economics* 135.2, pp. 711–783.

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

Collado, M. D., I. Ortuño-Ortín, and J. Stuhler (2023). "Estimating Intergenerational and Assortative Processes in Extended Family Data". *The Review of Economic Studies* 90.3, pp. 1195–1227.

Corak, M. (2013). "Income inequality, equality of opportunity, and intergenerational mobility". *Journal of Economic Perspectives* 27.3, pp. 79–102.

Dustmann, C. and A. Glitz (2011). "Migration and Education". In: *Handbook of the Economics of Education*. Ed. by E. A. Hanushek, S. Machin, and L. Woessmann. Vol. 4. Handbook of The Economics of Education. Elsevier, pp. 327–439.

Ferman, B. and C. Pinto (2019). "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity". *The Review of Economics and Statistics* 101.3, pp. 452–467.

Gelbach, J. B. (2016). "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34.2, pp. 509–543.

Gielen, A. C. and D. Webbink (2025). "Unexpected colonial returns". *Journal of Human Resources.*

Hammarstedt, M. and M. Palme (2012). "Human Capital Transmission and the Earnings of Second-Generation Immigrants in Sweden". *IZA Journal of Migration* 1.1, pp. 1–23.

Hansen, B. (2022). *Econometrics.* Princeton: Princeton University Press.

Ioannidis, J. P. A., T. D. Stanley, and H. Doucouliagos (2017). "The Power of Bias in Economics Research". *The Economic Journal* 127.605, F236–F265.

Koivunen, A. (2017). "Economies of Pride and Shame: Politics of Affect in New Narratives about Sweden Finns". In: *Citizenships under Construction: Affects, Politics and Practices.* Ed. by K. De Graeve, R. Rossi, and K. Mäkinen. COLLeGIUM: Studies across Disciplines in the Humanities and Social Sciences 23. Helsinki: Helsinki Collegium for Advanced Studies, pp. 50–66.

Lambert, P. S. and E. Bihagen (2012). *CAMSIS Sweden.* `https://www.camsis.stir.ac.uk/Data/Sweden90.html`.

Leamer, E. E. (2010). "Tantalus on the Road to Asymptopia". *The Journal of Economic Perspectives* 24.2, pp. 31–46.

Liang, K.-Y. and S. L. Zeger (1986). "Longitudinal Data Analysis Using Generalized Linear Models". *Biometrika* 73.1, pp. 13–22.

Lubotsky, D. and M. Wittenberg (2006). "Interpretation of Regressions with Multiple Proxies". *Review of Economics and Statistics* 88.3, pp. 549–562.

Lundh, C. (2005). *Invandringens arbetsmarknad: Ett historiskt perspektiv.* Stockholm: SNS Förlag.

Manduca, R., M. Hell, A. Adermon, J. Blanden, E. Bratberg, A. C. Gielen, H. van Kippersluis, K. Lee, S. Machin, M. D. Munk, M. Nybom, Y. Ostrovsky, S. Rahman, and O. Sirniö (2024). "Measuring Absolute Income Mobility: Lessons from North America and Europe". *American Economic Journal: Applied Economics* 16.2, pp. 1–30.

Nielsen, H. S., M. Rosholm, N. Smith, and L. Husted (2003). "The School-to-Work Transition of 2nd Generation Immigrants in Denmark". *Journal of Population Economics* 16.4, pp. 755–786.

Nordic Council (1954). *Nordic Passport Union Agreement (1954).* `https://www.norden.org/en/info-norden/nordic-passport-union`. Accessed: 2025-05-13.

OECD (various). *International Migration Outlook: Sweden Country Notes and Labor Migration Reports.* `https://www.oecd.org/migration`. Accessed: 2025-05-13.

OECD/European Union (2018). *Settling In 2018: Indicators of Immigrant Integration.* Paris: OECD Publishing.

Prokic-Breuer, T., S. Vermeulen, and D. Webbink (2024). "On the economic prospects of non-Western migrants in Europe". Mimeo.

Romano, J. P. and M. Wolf (2017). "Resurrecting Weighted Least Squares". *Journal of Econometrics* 197.1, pp. 1–19.

Saarela, J. and F. Finnäs (2007). "Adjustment failures in an immigrant population: Finns in Sweden". *Social Indicators Research* 82.3, pp. 545–563.

Saarela, J. and D.-O. Rooth (2006). "How Integrated are Finns in the Swedish Labour Market? Outcomes of Free Labour Mobility". *International Migration* 44.2, pp. 119–152.

Solon, G. (2004). "A Model of Intergenerational Mobility Variation Over Time and Place". In: *Generational Income Mobility in North America and Europe.* Ed. by M. Corak. Cambridge: Cambridge University Press, pp. 38–47.

Stuhler, J. (2014). “Mobility Across Multiple Generations: The Iterated Regression Fallacy”. Mimeo.

Swedish Migration Agency (2025). *Official statistics and historical overview.* `https://www.migrationsverket.se`. Accessed: 2025-05-13.

Sweetman, A. and G. Dicks (1999). “Education and Ethnicity in Canada: An Intergenerational Perspective”. *The Journal of Human Resources* 34.4, pp. 668–696.

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* 35.3, pp. 869–901.

Ward, Z. (2020). “The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration”. *American Economic Journal: Applied Economics* 12.4, pp. 73–102.

Weckström, L. (2011). *Representations of Finnishness in Sweden.* Studia Fennica Linguistica 16. Helsinki: Finnish Literature Society.

White, H. (1980). “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”. *Econometrica* 48.4, pp. 817–838.

Wooldridge, J. M. (2003). “Cluster-Sample Methods in Applied Econometrics”. *The American Economic Review* 93.2, pp. 133–138.

— (2010). *Econometric Analysis of Cross Section and Panel Data.* Second Edition. Cambridge: The MIT Press.

Zhao, L. and L. G. Drouhot (2024). “The Grandchildren of Immigrants in Western Europe: Patterns of Assimilation Among the Emerging Third Generation”. *Demography* 61.2, pp. 463–491.

Zorlu, A. and W. van Gent (2024). “Economic Assimilation of the “Third Generation”: An Intergenerational Mobility Perspective”. *International Migration Review* 58.2, pp. 734–763.

A Additional results

Table A.1: Number of observations by country

countrygroup	N. Ind.	countrygroup	N. Ind.
Germany	8 900	Chile	392
Yugoslavia	6 601	Romania	379
Estonia	3 979	Switzerland	310
Hungary	3 257	Bosnia	304
Poland	2 488	Syria	302
Turkey	2 480	United Kingdom	289
Czechia	1 973	North Macedonia	274
Austria	1 883	Lebanon	254
Russia	1 824	France	243
Greece	1 413	Portugal	213
Italy	1 149	Slovenia	203
Croatia	899	Morocco	178
Latvia	832	South America, rest	168
Netherlands	763	Palestine	129
Spain	606	Bulgaria	101
USA	401	Others (50 regions)	1 455

Note: Number of children in the main sample by countrygroup. Category "Others" include the following countries: Lithuania, Ireland, Belgium, Slovakia, Western Europe, rest, Ukraine, Eastern Europe, rest, Serbia, Albania, Canada, Central Asia, rest, Jordan, Iraq, Western Asia, rest, Iran, Afghanistan, Pakistan, India, Bangladesh, Sri Lanka, Tunisia, Egypt, Sudan, North Africa, rest, Nigeria, Gambia, West Africa, rest, Congo, Eritrea, Ethiopia, Somalia, Uganda, Kenya, East Africa, rest, Southern Africa, Central America, Colombia, Brazil, Peru, Bolivia, China, South Korea, Japan, East Asia, rest, Vietnam, Thailand, Philippines, South east Asia, rest, Oceania, West Indies.

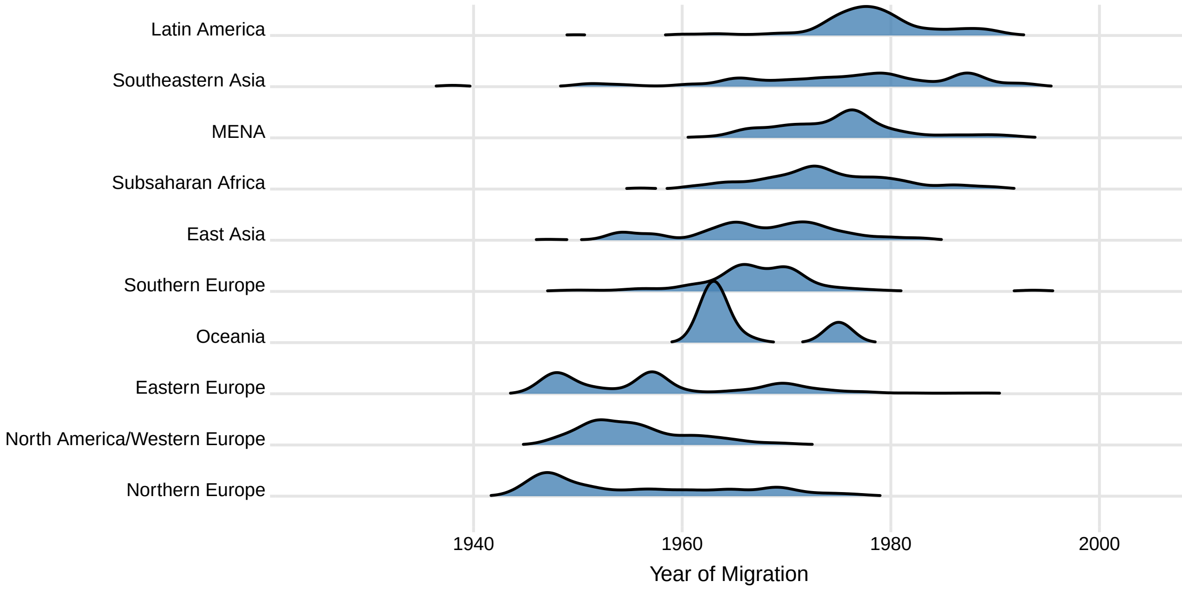

Figure A.1: Year of immigration to Sweden for the grandparent generation.

Notes: This figure shows the distribution of arrival years to Sweden for the first-generation immigrants in our sample.

Table A.2: Summary statistics

	Years of schooling	Log income (residualised)	Social stratification	Obs. /child	Birth year
Child (GPA)	0.31				1987.92
	28.61				8.57
Parents	0.14	-0.00	0.69	1.99	1957.78
	(2.24)	(0.32)	(11.09)	(0.11)	(7.72)
Extended family	-0.02	-0.01	0.11	5.61	1957.55
	(1.57)	(0.20)	(7.23)	(3.31)	(8.07)
Ethnic mean	-0.26	-0.04	-0.17	55 095.24	1962.99
	(0.58)	(0.07)	(2.33)	(10 3717.32)	(6.86)

Notes: Cells show means with standard deviation in parentheses for the main sample.

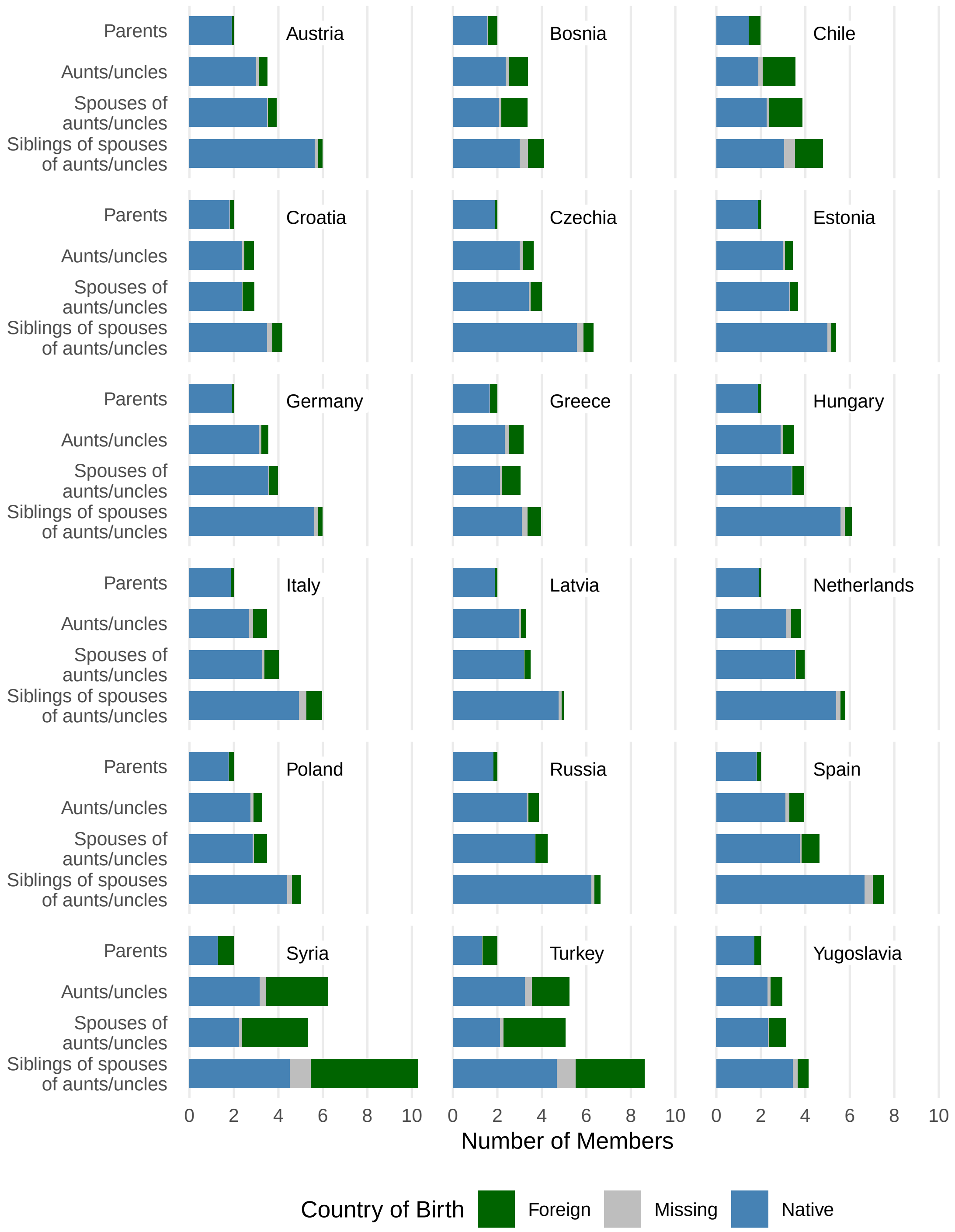

Figure A.2: Composition of native and foreign-born individuals per child.

Notes: Green bars indicate the average number of foreign-born individuals (with observed schooling) per child; blue bars indicate native-born individuals; gray bars represent foreign-born individuals for whom schooling data is missing.

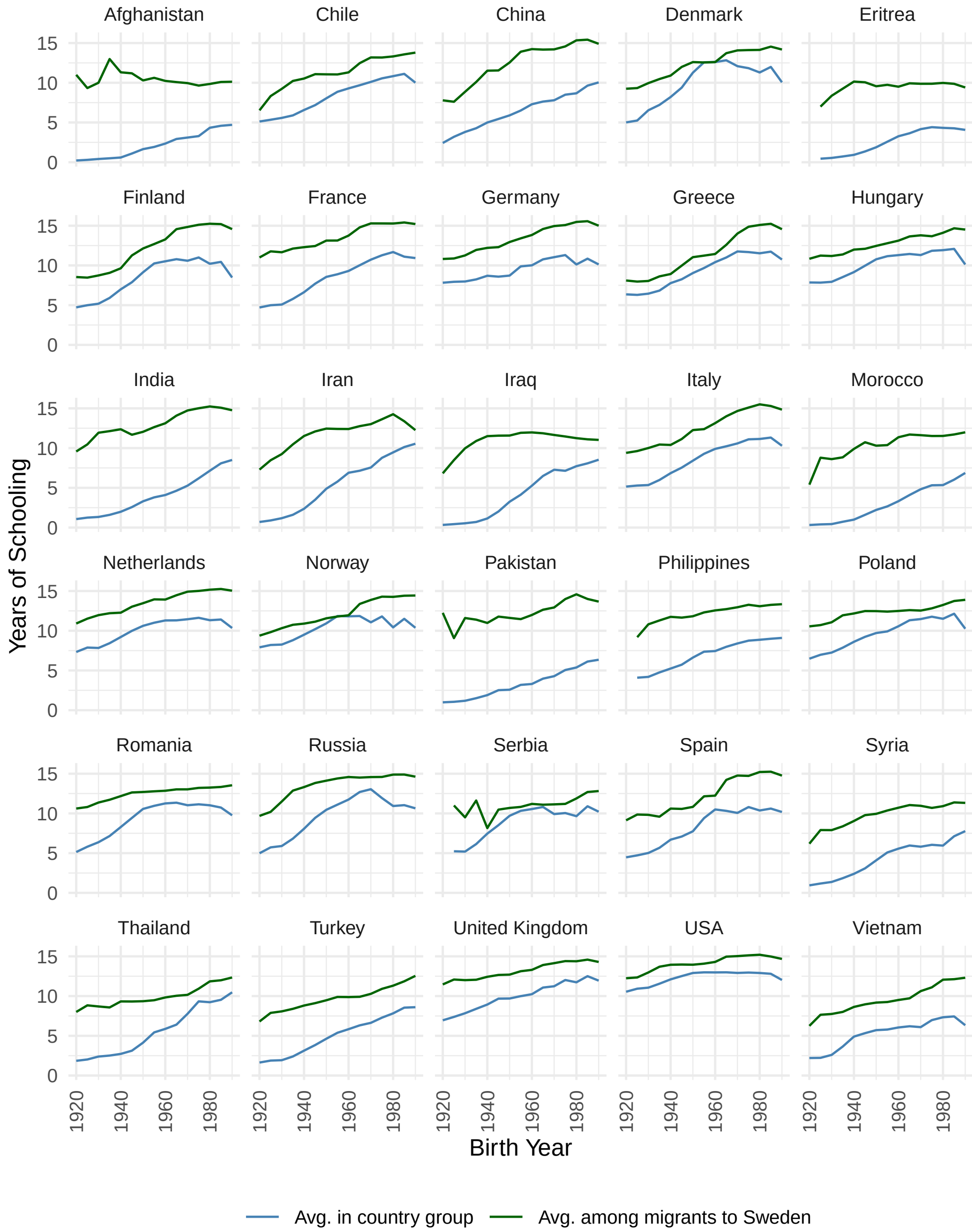

Figure A.3: Selection of Migrants across Birth Years

Notes: This figure shows average years of schooling by birth year and migration status. The green lines reflect the educational attainment of migrants as reported upon their arrival in Sweden (and potentially including education completed after arrival). The blue lines show educational attainment in the origin countries, based on the Barro–Lee dataset.

Table A.3: Separate relatives

	(1)	(2)	(3)	(4)
Parents	0.360	0.387	0.387	0.363
	(0.008)	(0.010)	(0.008)	(0.008)
Extended family	0.189			
	(0.015)			
Extended family (Natives)		0.122		0.097
		(0.011)		(0.011)
Extended family (Foreign Origin)			0.103	0.081
			(0.012)	(0.012)
Ethnic mean	0.041	0.064	0.065	0.040
	(0.044)	(0.044)	(0.042)	(0.045)
R^2	0.235	0.231	0.232	0.235
Num. Ind	39 556	39 556	39 556	39 556

Notes:

Table A.4: Alternative ethnic capital proxies

	(1)	(2)	(3)	(4)	(5)
Parents	0.368	0.368	0.368	0.368	0.368
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Extended family	0.210	0.210	0.210	0.210	0.209
	(0.022)	(0.021)	(0.022)	(0.022)	(0.021)
Ethnic mean	−0.035	0.034	−0.031	0.011	0.035
	(0.072)	(0.050)	(0.070)	(0.066)	(0.052)
WVS education index		0.039			0.058
		(0.010)			(0.031)
WVS culture index			0.008		−0.016
			(0.008)		(0.012)
Barro-Lee Years of schooling				−0.023	0.011
				(0.019)	(0.030)
R2	0.283	0.283	0.283	0.283	0.284
Num.Obs.	16 463	16 463	16 463	16 463	16 463

Notes: Each column shows results from separate regressions. The dependent variable is standardised GPA in both panels. WVS education index, WVS culture index and Barro and Lee (2013) years of schooling are standardised. Robust standard errors in parentheses.

Table A.5: GPA schooling regressions

	(1)	(2)	(3)	(4)	(5)
Parents		0.436	0.432	0.365	0.365
		(0.009)	(0.011)	(0.007)	(0.007)
Extended family				0.196	0.196
				(0.013)	(0.014)
Ethnic mean	0.631		0.089		0.014
	(0.063)		(0.055)		(0.056)
Sum	0.631	0.436	0.521	0.561	0.574
	(0.063)	(0.009)	(0.049)	(0.015)	(0.048)
R^2	0.096	0.236	0.237	0.247	0.247
Num. Ind.	36 305	36 305	36 305	36 305	36 305

Notes: Each column shows results from separate specifications where standardised Grade Point Average (GPA) is regressed on years of schooling. The years of schooling variables are divided by the standard deviation of years of schooling for parents. All regressions include fixed effects for birth year, age at migrating to Sweden and first parish of arrival of the ancestors. Each parental generation outcome is the average across all members of the given category of relatives. Standard errors are clustered by country group in parentheses.

Table A.6: Alternative specifications, FGLS

	(1)	(2)	(3)	(4)	(5)
	Child nbhd F.E.				
Parents		0.389	0.387	0.337	0.336
		(0.006)	(0.007)	(0.006)	(0.006)
Extended family				0.075	0.075
				(0.004)	(0.005)
Ethnic mean	0.399		0.063		0.013
	(0.042)		(0.036)		(0.038)
	Excl. arrival nbhd F.E.				
Parents		0.441	0.438	0.363	0.362
		(0.012)	(0.012)	(0.007)	(0.007)
Extended family				0.090	0.089
				(0.007)	(0.007)
Ethnic mean	0.618		0.096		0.029
	(0.064)		(0.045)		(0.047)
	Group-level				
Parents		0.430	0.428	0.352	0.353
		(0.005)	(0.005)	(0.005)	(0.005)
Extended family				0.099	0.099
				(0.003)	(0.003)
Ethnic mean	0.478		0.036		-0.041
	(0.022)		(0.021)		(0.021)
	Incl. Nordics				
Parents		0.415	0.414	0.346	0.346
		(0.013)	(0.013)	(0.015)	(0.015)
Extended family				0.099	0.099
				(0.004)	(0.004)
Ethnic mean	0.337		0.089		0.054
	(0.041)		(0.030)		(0.032)
	Multiple proxies				
Parents		0.485	0.483	0.425	0.425
		(0.011)	(0.011)	(0.008)	(0.008)
Extended family				0.165	0.163
				(0.011)	(0.012)
Ethnic mean	0.581		0.091		0.040
	(0.063)		(0.043)		(0.045)