Maryland Population Research Center

WORKING PAPER

Estimation, Simulation, and Validation of a Two-sex Model of Intergenerational Reproduction of Education

PWP-MPRC-2021-004
April 2021

Authors:

Michael S Rendall
Heide Jackson
Polina Zvavitch
University of Maryland

Shuchi Goyal
Mark S Handcock
University of California Los Angeles

Margaret M Weden
RAND Corporation

Estimation, Simulation, and Validation of a Two-sex Model of Intergenerational Reproduction of Education

Rendall, Michael S.*, Heide M. Jackson*, Shuchi Goyal**, Mark S. Handcock**, Margaret M. Weden***, and Polina Zvavitch*

* University of Maryland, College Park, ** University of California, Los Angeles, ***RAND Corporation

Acknowledgements: We are grateful for support from the National Science Foundation BIGDATA: Applications program, grant NSF IIS-1546259. This project was supported in part also by and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), population research infrastructure grants P2C-HD041041 to the Maryland Population Research Center at UMD (MPRC) and P2C-HD041022 to the California Center for Population Research at UCLA (CCPR). The authors also thank the CCPR for its training support from the NICHD T32-HD007545.

Keywords: microsimulation, revealed preferences matching, data combining, cross-survey multiple imputation

Corresponding author: Michael Rendall, email: mrendall@umd.edu.

Abstract

The intergenerational reproduction of educational inequality across generations is a fundamental process of social stratification. A model of intergenerational reproduction should account for the roles of partnering differences by education, of educational matching in partnering, of partnered and unpartnered fertility differences by education, and parent-child educational correlations. Most models until now have been "one-sex": they consider the characteristics only of one of the parents. Additionally, the model's simulated outcomes have not been evaluated for bias nor presented with confidence-interval estimates around the outcome measures. We address these limitations by specifying, estimating, simulating, and validating a two-sex model of intergenerational reproduction of education inside and outside marital unions. Our estimation builds in bias reduction and sampling-error minimization by using both medium- and large-scale surveys and combined-survey estimation of two of the component processes. The microsimulation model allows for annual birth, marriage, and divorce events and a one-time intergenerational educational-transmission outcome. We find that women's educational attainment, their number of births, and the distribution of births by parental marital status and maternal education, all match reasonably well with estimates from external data sources on the U.S. population, but only when using combined-survey estimation of the education-transmission process.

Estimation, Simulation, and Validation of a Two-sex Model of Intergenerational Reproduction of Education

Introduction

Models of intergenerational social reproduction, including of educational reproduction, have been estimated to provide a broader structure for describing the sources of intergenerational persistence of disadvantage, with respect to occupational class (Preston 1974), IQ (Preston and Campbell 1993), poverty (Musick and Mare 2004), and education (Mare 1997, Maralani 2013, Song and Mare 2017). Additional differentials by race have also been modeled (Preston 1974, Musick and Mare 2004, Maralani 2013). A crucial contribution of these models is that they go beyond measuring intergenerational *correlations*. In them, partnering and fertility processes are integral to generating the context in which intergenerational correlations "reproduce" inequality across generations. The methodological challenges of specification and estimation of an intergenerational reproduction model, however, are great, involving the chaining together of component-process equations into a valid overall model. The purpose of the present study is to construct and evaluate an intergenerational reproduction model of education in ways that overcomes important limitations of models until now.

Overview of the Problem and Our Solution

A limitation of many intergenerational reproduction models is that, with few exceptions, they are "one-sex" models. This means that they consider the characteristics only of one of the parents, whereas both parents' characteristics may be important (e.g., Beller 2009). In the case of models that examine contributions of both sexes, they may use *ad hoc* rules to form pairings (e.g., Maralani 2013). Song and Mare (2017) specify and estimate a "two-sex" model of educational mobility, but their couple-formation process does not account for competition for partners across educational levels. A particular strength of their study, however, that we do not attempt to emulate in our model, is that it allows for indirect and direct influences of a grandparent generation on grandchildren.

The model of the present study may best be viewed for how it builds on the Song and Mare (2017) study. Theirs makes the strongest claim to be a two-sex intergenerational reproduction model that we are currently aware of. Nevertheless it has four limitations, all of which we address in our study. First, they consider only marital fertility. Second, they do not model marital dissolution. Third, their two-sex model of couple formation accounts for the availability of individuals at each given education level, but not for the existence of individuals with different educational attainments who are potentially "competing" for those partners. Fourth, their simulated outcomes are neither evaluated for bias nor presented with estimates of the sampling error around the outcome measures.

Song and Mare used Panel Study of Income Dynamics (PSID) data in which all three generations are observed. The generalization (or unbiasedness) of the population composition represented by the PSID is limited, however, by its construction to exclude immigrants in its "descendent" sampling structure. This weakens its representation especially of Hispanic Americans, whose, education, fertility, and family processes differ from non-Hispanic Americans' (Landale and Oropesa 2007), who account for one in four births in the U.S. (Martin *et al* 2021), and whose educational attainment at childbearing is the lowest of the main race/ethnic groups (Rendall *et al* 2018). Low levels of education among Mexican-born immigrants is a major factor producing this result (Rendall and Parker 2014). Another problem with using a single panel survey is sample size. The level of sampling error around the simulated outcome measures is potentially large when all component processes are estimated from a single panel survey. This is not estimated by Song and Mare, nor is it estimated in any of the other intergenerational simulated model studies cited above.

In the present study, we specify, estimate, and simulate a two-sex model of intergenerational reproduction of U.S. educational attainment inside and outside marriage. Our model structure includes four processes to reproduce educational attainment across generations: (1) marriage; (2) divorce; (3) marital and non-marital fertility; and (4) intergenerational transmission of educational attainment. The latter three processes, as we model them, depend on the education of both parents for children born within marriage, but on the mother only, together with marital status at birth, for children born non-maritally. For the modeling of marriage, we account for the education of all unmarried women and men. For this model, we take advantage of recently-developed methods (Goyal *et al* 2020; Handcock *et al* 2021) that account for competition across men and women from different educational attainment levels, and that allow for the individual's staying unmarried as a simultaneously-modeled outcome. We use multiple sources of smaller-scale, medium-scale, and large-scale survey data to estimate the parameters for each of the four processes. Specifically, we use the American Community

Survey (ACS), the Survey of Income and Program and Participation (SIPP), and the National Longitudinal Surveys (NLSY, 1979 and 1997 cohorts). Our estimation includes cross-survey multiple imputation methods (Rendall *et al* 2013), applied to the divorce and intergenerational education transmission processes. Sampling error across all data sources is incorporated into estimation of confidence intervals around simulation model outputs.

Literature Review: Education in Two-sex Intergenerational Reproduction Processes

We focus our literature review on methodological issues of model structure and estimation. We begin by reviewing the literature on the role of education in family-demographic processes and on the need for a model to be structured as two-sex. To do this, we review work relevant to the roles of male and female education in the four component processes of the intergenerational reproduction of education. We then proceed to review methodological work relevant to the estimation of the component processes. These are notably a new two-sex marriage model and application of combined-survey methods to enable unbiased two-sex estimation of the divorce process and the intergenerational education transmission process with sufficiently large total sample sizes.

Education and Fertility and Family processes

The need for population models of the reproduction of education and educational inequality arises from the recognition that correlations between parents' education and their children's education is just one part of the reproduction of educational inequality across generations. Education has long played a major role in partnership formation, partnership dissolution, and fertility. These processes also contribute to the reproduction of educational inequalities from one generation to the next. It is noteworthy here that Song and Mare's (2017) educational reproduction model produces the result that, in the long run, college-educated children are no more likely to descend from college-educated parents and grandparents than they are to descend from high school educated parents and grandparents. Differential fertility by education, and educationally heterogamous marriages, offset positive parent-child educational correlations to produce this result in their model. However, the population generalizability of this result is difficult to assess because of their restricting the educational reproduction process to marital childbearing. For the last two decades, non-marital childbearing has accounted for about 40% of all U.S. births annually (Solomon-Fears 2014; Martin *et al* 2021). Less educated women in the U.S. have increasingly faced challenging prospects with respect to finding an educational marriage match (Lichter *et al* 2019), and their marriages are likely to dissolve more quickly (Schwartz and Han 2014).

The importance of taking into account both parents' education in the processes of educational reproduction has increased as women's educational attainment and labor force participation have increased strongly in recent decades (DiPrete and Buchmann 2006). First, education of both the woman and man are powerful factors to account for in pair-formation processes (Schwartz 2010, 2013). Second, higher educational attainment and educational homogamy reduce divorce likelihood (Schwartz and Han 2014; Raley and Sweeney 2020). Both marriage (Goldstein and Kenney 2001; Lundberg and Pollak

2015) and divorce (Martin 2006; Smock and Schwartz 2020) have become more unequal over time, favoring college-educated over less educated Americans (Cherlin 2010). Third, education of both the woman and man are powerful factors to account for also in fertility by the educational attainment of the woman (Yang and Morgan 2003; Nitsche and Bruckner 2020) and her partner (Joffe and Li 1994; Kravdal and Rindfuss 2008). Although the relationship between education and fertility has traditionally seen college-educated women experience the highest likelihood of under-achieving with respect to their fertility targets (Quesnel-Vallee and Morgan 2003), Hazan and Zoabi (2014) described the fertility-by-education relationship for U.S. women as moving towards being increasingly U-shaped. Women who are highly educated have begun to have more children than women with a high school degree or some college levels of education, though still fewer than women with less than 12 years of education (see also Lundberg *et al* 2016). College-educated women in the U.S. have long been less likely to have a child outside of marriage, and less likely to have a child at a young age (McLanahan 2004). However, this educational divergence has become even greater over recent decades (Lundberg and Pollak 2015).

New marriage modeling and combined-survey estimation methods

For the modeling of marriage by education, we develop a two-sided discrete choice model for the revealed preferences (Goyal *et al* 2020). This model fully accounts for competition across men and women from different educational attainment levels, and allows for the individual's staying single as a simultaneously modeled outcome. The conceptual and computational challenges of a fully two-sex modeling of pair formation are very substantial. They involve a shift from a standard demographic modeling approach that focuses on rates attributed to individual propensities to rates that are generated by the interaction of individuals' preferences and the availability of other individuals with matching preferences and characteristics. Previous work on this problem includes Schoen (1981), whose harmonic-mean solution is implemented by Song and Mare (2017), Pollak (1986, 1990), Pollard (1997), Choo and Siow (2006), and Logan *et al* (2008).

Estimation of components of a two-sex intergenerational reproduction model requires data on both the woman and man as measured before a given process outcome. For divorce, this involves observation of both spouses' educational attainment before a divorce event. For parent-child educational transmission in our model, this involves observation of both the mother's and father's educational attainment at the time of the child's birth. For both processes, typically this requires panel data, or otherwise the kind of detailed marital history data usually collected only in specialized cross-sectional surveys. These types of data sources have relatively small sample sizes, reducing statistical efficiency. They are also possibly subject to attrition bias in the case of panel data (Fitzgerald *et al* 1998), or to retrospective recall error for marital histories (Kennedy and Ruggles 2014).

Combined-survey estimation can address both those efficiency and bias problems (Ridder and Moffitt 2007). One form of combined-survey estimation is pooled cross-survey multiple imputation (MI). The statistical theory of pooled cross-survey MI is developed in Gelman *et al* (1998) and Rendall *et al* (2013). When two surveys are combined, the larger survey is expected to have advantages of both sample size and unbiasedness, and the smaller survey to have advantages of a fuller range of predictor variables. Applications in social demography include Baker *et al* (2015), Capps *et al* (2018), and Zvavitch *et al* (2020). Like within-survey MI for item non-response, cross-survey MI accounts for increases in variance of the estimates induced by including imputed values. It has a major advantage over within-survey MI, however, of more easily satisfying the missing at random (MAR) assumption needed for unbiased MI, since the reason the value on the variable is missing is that the respondent was randomly sampled into the survey that does not include the question.

Population model outcomes involving multiple equations are typically derived by microsimulation. This is the only practical solution in which the number of possible individual paths or trajectories is high (Moffitt and Rendall 1995, Thomson *et al* 2012). Microsimulation, because it operates similarly to resampling techniques used in statistics, also allows naturally for estimation of variability in the outcomes that is due to sampling error in the estimation of the model parameters (Wolf 2001). In particular, the microsimulation framework may be combined with bootstrap resampling.

Method

Overview of the Intergenerational Reproduction Model

We refer to the entire process that we model as "the intergenerational reproduction of educational attainment." Distributions of education in the grandparent generation (G1), parent generation (G2), and the child generation (G3) are analyzed as occurring through four component processes of the intergenerational reproduction of education: Marriage, Divorce, Fertility, and Parent-child (“intergenerational”) Transmission of Education. Education e is a four-category variable: less than high school graduate; high school graduate; some college (including two-year, associate-degree graduate); and college graduate (four-year college).

They become five processes when expressed over three generations:

1. Intergenerational education transmission from G1 to G2;
2. Marriage of G2;
3. Divorce of G2;
4. Fertility of G2;
5. Intergenerational education transmission from G2 to G3.

The intergenerational reproduction model can actually be simulated across any number of generations. We limit the number of generations to three to allow us to validate the simulated model outcomes against observed data matching the G2 women, and against observed data matching their G3 children at the time of their birth.

Our focal generation is the G2. We subscript the simulated individuals’ generation by $g = 1,2,3$, respectively for their parents as the “grandparent” generation, for themselves as the “parent” generation, and for their children as the “child” generation. In our microsimulation, we allow for a population of $i = 1,2,\ldots,I$ G2 women and for $j = 1,2,\ldots,J$ G2 men. As we describe in more detail following our description of the estimation of the four processes immediately below, we arbitrarily set I and J each to

6,000. This is a sufficient number to allow for reliable microsimulation of G1, G2, and G3 distributions of education.

Estimation of the Four Component Processes

To summarize, the four component processes in intergenerational reproduction model are Marriage, Divorce, Fertility, which occur to G2 only, and Intergenerational Education Transmission, which occurs to both G2 and G3. We use a single data source, the American Community Survey (ACS), to estimate the Marriage and Fertility equations. We use two data sources for the Divorce equation (ACS and Survey of Income and Program Participation, SIPP) and two data sources for the Intergenerational Education Transmission equations (NLSY79 YA and NLSY97).

1. Marriage

This component process is by far the most complicated of the four, involving simultaneously the preferences and availability, by education and age, of unmarried women and men. Each year, unmarried women and men in the model are either "matched" to each other and become married, or remain unmarried, in a process that accounts for both the preferences and availability of all unmarried individuals. We represent it using a two-sided logit model for the "marriage market" (Menzel 2015; Goyal *et al* 2020). In this model, individuals have *preferences* (represented by utilities) for all population members of the opposite sex and individuals maximize their utility over the set of all people of the opposite sex who are *available*. This means that no man and woman believes he or she can improve their matches by dissolving their current unions and forming a new one with each other, a condition that is called *stability* in the game theory of marriage (Gale and Shapley 1962). Specifically, let $U_{i,j}$ denote the utility female i has for male j, and let $V_{j,i}$ denote the corresponding utility male j has for female i. If these utilities are unique, we can obtain a set of strict rankings of possible partners for each male and each female. Additionally, we include a utility for no union, denoted by $U_{i,0}$ and $V_{j,0}$ for each female and male, respectively.

We represent here a vector of observed characteristics, unobserved characteristics and unobserved utilities. Our analytic sample consists of adults 20-39 who are not currently married and those who, within the past year, have married a person of the opposite sex also between the ages of 20-39.

The observed set of "status characteristics" $W_{i,j}$ for pairing female i and male j are age and education. The models for the utility gained by woman i partnering with man j, and that for the utility gained by man j partnering with woman i are:

$$U_{i,j} = \theta_w^T X_{i,j} + \gamma_{i,j} \qquad\qquad V_{j,i} = \theta_m^T Z_{j,i} + \epsilon_{j,i} \tag{1}$$

where $X_{i,j}$ is the vectorized form of the indicator matrix with $(w_i, w_j)^{th}$ element 1 if the female has status characteristics w_i and the male has status characteristics w_j, and zero otherwise. Similarly, $Z_{j,i}$ is the vectorized form of the indicator matrix with $(w_i, w_j)^{th}$ for men. For example, with four education and four age categories, there are 16 status characteristic types for each sex. Here θ_w is a vector of female preference coefficients for the status characteristics pairings $w_{i,j}$ and θ_m is the male preference coefficients for status characteristics pairings $w_{j,i}$. As discussed in Goyal *et al* (2020), for this model only

$\beta = \theta_w + \theta_m$ is identifiable and our parametrization will reflect this. The random components of the utility model account for unobserved information about individuals which may impact partnership choices. The random terms $\gamma_{i,j}$ and $\epsilon_{j,i}$are assumed to be identically distributed draws from an extreme-value type-I (Gumbel) distribution (this can be relaxed, see Goyal *et al* 2020).

We additionally define the random utility for the choice of remaining single as

$$U_{i,0} = 0 + \max_{1 \le k \le \sqrt{N_m}} \gamma_{i0,k} \qquad\qquad V_{j,0} = 0 + \max_{1 \le k \le \sqrt{N_w}} \epsilon_{j0,k}$$

for females and males, respectively, where N_m is the number of men in the population and N_w is the number of women. This utility specification implies that the deterministic component of the utility for an individual choosing to be unpartnered is 0. If the nondeterministic components of the single utility functions are chosen to be standard Gumbel then the total nondeterministic component is Gumbel with locations $\sqrt{N_m}$ and $\sqrt{N_w}$, respectively. This specification ensures that the share of singles in the market is asymptotically constant with respect to N_m and N_w (Menzel, 2015, Assumption 2.2).

We take advantage of the recent work of Goyal *et al* 2020, who developed the theoretically innovative framework of Menzel (2015) to enable the estimation of the preference parameters, $\beta,$ based on sample survey data on partnerships and population composition. We use data from the 2008 American Community Survey (ACS), obtained through IPUMS (Ruggles *et al* 2020). The ACS has a direct question asked of women in every year since 2008, “did you get married in the last year?” If the answer is “yes,” we use the opposite-sex spouse characteristics to assign the age and education of the husband she married. Additionally, the marital status of every adult member of the ACS household is asked, providing data for the unmarried population of men and women and their ages and educational attainments.

In the microsimulation model, we then use the above two-sided logit model with these estimated parameter values to simulate the matching of women and men to each other. In the model of equation (1), the status characteristics preference parameters drive the partnership process. To understand the matches produced, let $O(i)$ and $O(j)$ denote the opportunity sets for females and males, defined as:

$$O_w(i) = \{j: U_{j,i} > U_{j,\mathrm{fp}(j)}\} \cup 0 \qquad O_m(j) = \{i: V_{i,j} > V_{i,\mathrm{mp}(i)}\} \cup 0$$

where $\mathrm{mp}(i)$ and $\mathrm{fp}(j)$ represent the indices of the male that female i is in a marital union with, and the female that male j is in a marital union with, respectively. In the case the individuals are not in a union, the values are 0. The opportunity set for female i is the set of males who prefer female i to the females with whom they are currently paired. Then stability requires:

$$V_{i,\mathrm{mp}(i)} \geq V_{i,j} \quad \forall j \in O_w(i) \cup 0 \qquad U_{j,\mathrm{fp}(j)} \geq U_{j,i} \quad \forall i \in O_m(j) \cup 0 \qquad (2)$$

These inequalities also allow non-unions to be individually rational. To simulate the marriages in each year, we generate dyad-specific random utilities from equation (1) for each female-male dyad in the population and then find a stable set of unions that satisfy (2) via the Gale-Shapley algorithm (Gale and Shapley 1962). The basic idea is to simulate a population of unions from the preference parameters and count the number of unions between (i, j) pairs of each status characteristics. This general approach to inference, including uncertainty quantification and simulation is implemented in the open-source software package **rpm** (see Goyal *et al* 2020; Handcock *et al* 2021). Estimates and standard errors for the model are provided in Appendix Table A1.

2. Divorce

The ACS has a direct question asked of women in every year since 2008, “did you get divorced in the last year?” We use data from the 2011 ACS for those women who are 20-39 years old and who either are currently married and report no divorce in the last 12 months, or who report a divorce in the last 12 months irrespective of their current marital status. Kennedy and Ruggles (2014) find the ACS to be an excellent source of divorce estimates. While very large sample sizes are available again in the ACS, the ACS data can be used alone only for one-sex estimation of the divorce probabilities. In the case of a divorce event, because the man is no longer in the ACS household after divorce, only the woman’s characteristics (age and education) are known. We rely on Survey of Income and Program Participation (SIPP, United States Census Bureau 2014) panel data to observe a divorce as it occurs between four-month-apart waves, accumulated over one-year intervals. With its panel observation plan, the SIPP provides pre-divorce-exposure educational attainment and other information for both spouses, both in the case that a divorce occurs over the one-year interval and in the case that it does not. (The SIPP’s marital history, conducted in Wave 2 of each Panel, does not have partner information.) Our use of the SIPP panel waves to identify divorces follows Manning, Brown, and Stykes (2016), who used SIPP panel data to code both cohabitation dissolutions and marital dissolutions. The SIPP uses frequently-drawn new panels and thus provides a more accurate picture of the representation of the U.S. population at any given time than do longer-running panels. We use the 2004 and 2008 Panels. The 2004 SIPP Panel respondents were interviewed every four months beginning in the Spring of 2004 and ending in the Fall of 2007. The 2008 SIPP Panel respondents were interviewed every four months beginning in the Fall of 2008 and ending in the Fall of 2013. We assemble the data in a couple-year format, meaning that every observation includes the identification of the man and woman in the couple, their marital status, and education. Details of the SIPP coding of divorce are provided in Appendix 2. To best match the years observed in the SIPP, we use the mid-point year of 2011 in the ACS data. This single year of ACS data provides approximately 10 times the number of person-years of exposure to divorce as in the two SIPP panels. The SIPP panel years accordingly produce fewer divorce events compared to the ACS, and have *a priori* unknown attrition biases. We show in Appendix Table A2a that the overall bias in the SIPP is downward, indicating that attrition is positively correlated with getting divorced. For these reasons, we used combined-survey estimation with the SIPP and ACS. Divorce events by the woman's age and educational attainment are provided in the ACS, which is not subject to attrition bias, and divorce events by both the woman's and the man's ages and educational attainments are provided in the SIPP.

Our divorce prediction equation is based on logistic regression of divorce on the educational characteristics of the women and man:

$$Pr[Divorce] = logit\{ e_f, e_m \} \tag{3}$$

where e_f, e_m are the education of the woman and man, respectively. We combine data from the ACS with the 2004 and 2008 panels of the SIPP using pooled cross-survey multiple imputation (Rendall *et al* 2013). The SIPP has the 'complete data' and the ACS the 'incomplete data'. Because only in the SIPP do we observe the man's education for both intact and divorcing couples ('complete data'), we multiply impute man's education to each ACS record before conducting pooled-survey estimation (with both the ACS and SIPP observations). Using both surveys in a combined-survey approach takes advantage of the sample size and unbiasedness of the ACS and the presence of both partners' education in the SIPP. We estimate a binary logistic regression model with annual divorce versus no divorce as the dependent variable. We use the reported education level at survey for women only in the ACS, even though for couples that remain married, the education levels of both the woman and the man are available. Because the education for the partner of those divorcing in the ACS is not available, using the education of their partner in the case that they remained married in the year would violate the missing at random (MAR) assumption needed for unbiased multiple imputation. Notably, the man's education data are missing on the dependent variable (for all cases of divorce=1).

Because the pooled-survey estimation combines observations from two nationally representative surveys of approximately the same ages and years, we begin by assuming that they sample from a common social process except for a potential difference in levels of the outcome variable (divorce). We test the validity of this assumption by conducting diagnostics under a model-fitting framework, following Rendall *et al* (2013). Model-fit results are presented in Appendix Table A2b. We find model fit improvement only when adding a "SIPP" intercept shift variable for divorce rate differences between the surveys, and not when adding a full set of covariate interactions with "SIPP" survey. Model fit improvement in the latter case would be evidence calling into question the appropriateness of a pooled-survey method. When using the combined-survey parameter estimates to generate predicted divorce probabilities in the microsimulation, we set the SIPP parameter to zero assuming the overall divorce rate will be unbiased in the ACS.

3. Fertility

We estimate the annual birth probability at ages 20 to 39 using data on fertility in marital unions, and in non-marital unions or outside a coresidential union, sampled in the ACS over the years 2001-2011 and 2013-2017. In 2012, data on fertility is suppressed in the ACS public use version for some geographic areas (i.e., 59 PUMAs within the states of Florida, Georgia, Kansas, Montana, North Carolina, Ohio and Texas) due to inconsistencies in data collection. We therefore omit data from 2012 to maintain national representativeness. We identify births using the ACS's question asked of all 15 to 50 year-old women, "in the past 12 months, has this person given birth to any children?" This includes births to women in marital unions as well as non-marital births to women in non-marital unions (cohabiting women) and to women in no coresidential union (unpartnered or single women). Age and educational attainment are available for all married and unmarried women and, if married, for her co-resident husband.

Our fertility prediction equations are estimated separately for married (r) and unmarried (u) women. Again using e_f and e_m to represent the woman's and the man's education for married births, and using a_f to represent the woman's age and p_f to represent her parity, the marital and non-marital fertility equations are respectively:

$$\Pr[Birth_r] = \text{logit}\{\, e_f, e_m, a_f, p_f\} \tag{4a}$$

$$\Pr[Birth_u] = \text{logit}\{e_f, a_f, p_f\} \tag{4b}$$

Regression parameter estimates for the two equations are shown in Appendix Table A3.

4. Intergenerational transmission of education

The child "inherits" educational attainment from both the mother and father, or in the case the parents were not married at the child's birth, from the mother only. Marital versus non-marital birth is implicitly a predictor of the child's "inherited" educational attainment: In our two-parent intergenerational transmission ("inheritance") modeling, the four-category educational attainment of the child generation is predicted in separate multinomial logistic (MNL) equations for maritally-born, $e_{,r}$ and non-maritally-born $e_{,u}$ children, where the comma in the subscript is used to indicate that marital status r or u is a parental characteristic. We need to know the mother's or both parents' educational attainment at the time of the child's birth to link to the previous microsimulation-model process of non-marital or marital fertility by the parent or parents' educational attainments. Indexing again f=female, m=male, we denote mother's education by e_f and father's education by e_m. Denoting by s the gender of the child (male/female), the two parent-child educational transmission equations are:

$$\Pr\left[e_{,r} = e\right] = \text{MNL}\{e_{f,}, e_{m,}, s\} \tag{5a}$$

$$\Pr\left[e_{,u} = e\right] = \text{MNL}\{e_{f,}, s\} \tag{5b}$$

Estimation of these equations requires data including observation of mother's marital status at birth, of both the mother's and father's educational attainment at the time of the child's birth for a marital birth, or of the mother's educational attainment only at the time of the child's birth for a non-marital birth. Panel data are generally needed for this. Moreover, by the time the child's educational attainment is observed, around age 24, they are often no longer living with either or both parents. Therefore, the panel data need to track the child individually into early adulthood. The NLSY79 meets all of these requirements, through its panel observation of the mother and co-resident father at the time of the birth, together with the linked NLSY79-Youth (YA) survey observation of the child as a young adult (Bureau of Labor Statistics 2019a). Sample sizes for observation of both the birth and the child's educational attainment in the NLSY79, however, are relatively small. There is also considerable potential for attrition bias given that both the NLSY79 woman and then her child need to be followed until that child becomes a young adult (in the NLSY79 YA sample). Finally, the NLSY79 YA represents the children of NLSY79 women who were living in the U.S. in the late 1970s, thereby missing the coverage of children born to women who immigrated to the U.S. more recently. A second NLSY cohort, the NLSY97 (Bureau of Labor Statistics 2019b), is useful in addressing all these limitations. We are able to use it in combination with the NLSY79 to add approximately 5,000 children to our intergenerational-transmission estimation sample, as we now describe.

The panel character of the NLSY97, like the NLSY79 YA, can be used to observe the outcome variable of the Intergenerational Transmission of Education process, being the child’s own educational attainment as a young adult. The NLSY97 also provides the predictor variables of the child’s gender and the parents’ marital status when the child was born. The NLSY97, however, is missing observation of the parents’ educational attainments at the time of the child’s birth. It does have reports of the parents’ educational attainments at the time of the child’s adolescence (in the initial survey year 1997), including in cases where one of the parents is not living in the household but is the biological father or biological mother of the child. We use these parental educational attainments at the time of the child’s adolescence to impute parental educational attainments at the time of the child’s birth. To do this, we combine the NLSY79 and NLSY97 in a cross-survey multiple imputation (MI) procedure (see Appendix 4). Parental educational attainments at the time of the child’s adolescence serve as “auxiliary” variables (Schafer 2003) in the cross-survey MI, not used in the analysis equations (5a) and (5b), but used in imputing to the NLSY97 child the mother’s educational attainment at the time of the child’s birth and, if a marital birth, also the father’s educational attainment at the time of the child’s birth. For the multiple imputation, we first divide the two surveys’ samples into marital and non-marital births, as we do in the Intergenerational Education Transmission equations (5a) and (5b). Marital status at the child’s birth is coded in the NLSY97 from the mother’s martial history provided in the 1997 year. Pooled-survey (NLSY79+NLSY97) estimation with multiply-imputed parental educational attainment at the time of the child’s birth for the NLSY97 cases is used to estimate equations (5a) and (5b) (see Appendix Table A4c for regression estimates of the NLSY79-only and pooled NLSY79+NLSY97 models).

Microsimulation

As noted above, the microsimulation model's five processes are conducted over three (G1, G2, G3) generations: Intergenerational Education Transmission from G1 to G2; Marriage of G2; Divorce of G2; Fertility of G2; and Intergenerational Education Transmission from G2 to G3. These processes are simulated to occur in five 20-year periods that we call 'eras.' To build a simulated "analysis" population of (G2) 20-39 year olds, in each year of eras 1, 2, and 3, we simulate 300 G2 individuals (150 women; 150 men) who enter the model at age 0 in each simulation year (see Appendix Figure A1). At ages 20 to 39, the 6,000 women born in eras 1 or 2 marry, divorce, and give birth to G3 children. These family-demographic events, including the G3 births, occur in eras 2, 3, and 4. The G3 children's own education at age 20 is assigned in eras 3, 4, and 5.

In the year that the G2 women and men enter the model, they are first assigned to be a marital or non-marital birth, with a probability estimated from the NLSY79 YA sample that includes their NLSY79 parental marital status, as measured in the survey interview after the birth. Next, for a marital birth, the educational attainments of G1 mother and father are assigned and, for a non-marital birth, the educational attainment of the G1 mother is assigned. The NLSY79 YA sample is again used to estimate these distributions of parent education for marital and non-marital births of G2s.

Era 2 (years 21-40 of the model) represents the first era in which family formation processes take place. In the year that a G2 individual who entered the model in era 1 reaches age 20, they are assigned their educational attainment. This value stays constant throughout their reproductive life. The Intergenerational Education Transmission equations (5a) and (5b) are used to assign the G2's educational attainment probabilistically as a function of G1 parent marital status, G1 education, and G2 gender (see Appendix 4). Each year exposes G2 cohort members between the ages of 20 and 39 to marriage, fertility and divorce processes, in that order. Marriage risk occurs to unmarried women and men. The possible outcomes include staying unmarried and, if marrying, any of four values of education for the new spouse. Using the Gale-Shapley algorithm to calculate stable matches (Gale and Shapley 1962; Goyal *et al* 2020), we generate the utility of marriage options as a function of available G2 men and G2 women's education and age (see Appendix Table A1 for parameter estimates) as well as a random utility component assigned to each individual at each period. Births (0,1) are assigned to G2 women differentially by marital status (see Appendix Table A3 for fertility regression parameter estimates). For unmarried G2 women, the probability of experiencing a birth is a function of age, education, and parity. For married G2 women, the probability of experiencing a birth is a function of her education and that of her married partner, single-year age, and parity. Finally, all married G2 couples, including those who married in the current year, are assigned a divorce value (0,1) which is a function of both partners' education (see Appendix 2).

In era 3 (years 41 to 60), in each year 300 G2 individuals who entered the model in era 2 are assigned at age 20 their educational attainment. These G2 individuals are again exposed to marriage, fertility, and divorce when they are between the ages of 20 and 39. Analogously, in era 4 (years 61 to 80), those G2s born in era 3 are assigned at age

20 their educational attainment and are exposed to marriage, fertility, and divorce when they are between the ages of 20 and 39. However, G2s born in era 3 do not contribute to the analysis group of G2 women. They do contribute to the marriage market and, if marrying a G2 analysis-group woman, contribute as a father of a G3 child. No further G2 individuals enter the model in era 4.

In eras 2, 3, and 4, G3 children are born through the fertility processes experienced by the G2 women who are born in eras 1 and 2, and who live through their reproductive years 20-39 in eras 2, 3, and 4 (see again Appendix Figure A1). The G3 children's G2 parental characteristics are assigned deterministically from the G2 fertility outcomes. That is, a marital birth simulated for a G2 mother and father of educational attainments e_f and e_m results in a G3 child being born with mother's education by $e_{f,G2}$ and father's education by $e_{m,G2}$; a non-marital birth simulated for a G2 mother of educational attainment e_f results in a G3 child being born with mother's education by $e_{f,G2}$ and no father's education. The G3 child's gender is assigned randomly assuming an equal sex-ratio. The G3 children's own educational attainment is assigned at age 20 in eras 3, 4, and 5. In era 5 (years 81 to 100), the only event modeled is the assignment of educational attainment at age 20 to G3 children who are born in era 4.

One note about the G2 cohorts born in era 1 versus era 2: only in era 2 are they born at a simulation-model time when there is already a full population of 20 to 39 year olds; for those born in era 1, the population is still accumulating individuals across those ages. This restricts the ages of the spouse pool, though it restricts it less and less with each successive year. This is not expected to be an issue for microsimulation-model outcome bias in the present version of the model, as in none of the component-process equations is the man's age a predictor variable, and the education of the man is fixed across ages 20-39. For our main results, however, we alternately calculated microsimulation outcomes for women born in era 2 only. We found no substantial differences for women born in era 2 only, compared those outcomes simulated for women born in both eras 1 and 2.

Implicit in our model are several assumptions which deserve mention here. First, the value of educational attainment is assigned at age 20 and is assumed to be fixed after age 20. This value represents eventual completed education. It is estimated from data in which education is observed at approximately age 24 (see again Appendix 4), thus allowing for observation of educational attainment up to a completed college degree. In results not shown, we found evidence of some educational attainment increases in the observed (ACS) population after that age, but these increases in education after age 24 will not be represented by our model. Second, our model constrains the processes of family formation (marriage, fertility, divorce) to occur only between ages 20-39. We show below that this age range reduces the sum of cohort fertility by a surprisingly small proportion. Third, individuals can experience only one of each event in a year (that is, in a given year a G2 member cannot get married or divorced multiple times; however, they can get married, have a birth, and subsequently get divorced within the same year). Fourth, our model is two-sex except for non-marital fertility which assumes only mothers' characteristics to be relevant; and we do not model non-marital fertility for men.

Finally, no mortality is included in the model; all G2 who enter the model at age 0 exit the model at age 40.

To represent sampling error in the data used in the component-process estimation, in our microsimulation estimates, we bootstrap the surveys used to estimate each process: the assignment of G1 marital status and education as parental characteristics for G2 (NLSY79); the transmission of education from the G1 to the G2 and from the G2 to the G3 generations (NLSY97 + NLSY79); G2 marriage (ACS); G2 fertility (ACS); and G2 divorce (ACS + SIPP). For parameters using a combined-survey estimate (educational-transmission and divorce), the bootstrapping is stratified by survey and conducted prior to imputation to preserve the contribution of each survey to the overall estimate. Bootstrapping prior to imputation is consistent with past recommendations to avoid bias in estimates (Schomaker and Heumann 2018). We generate 1,000 bootstrapped parameters for each simulation process and proceed to run the microsimulation for 1,000 iterations, each with a different set of bootstrapped parameters. For each model outcome, we present the median estimate across the 1,000 simulation runs and present 95% confidence intervals which are bounded at the 2.5 percentile of estimates and the 97.5 percentile of estimates across simulation runs.

We present results alternately in the form of the G2 (parent generation) women as the unit of analysis, and of the G3 children (both genders combined) born to G2 women and men as the unit of analysis. This follows the insights of Preston (1976) that parental and child perspectives are both complementary and sometimes very different from each other. We present these main results for G2 women because fertility is modeled as occurring to women (even though their marital status and education of husband if married, and of potential husbands among those not married, are generated through two-sex modeling), and because our external data for use in the model validation are primarily for women. We evaluate both efficiency and bias of the microsimulated outcomes. Although we use combined-survey estimation for two component processes, education transmission and divorce, we focus on evaluating gains to combined-survey estimation of the education-transmission process. The main outcomes we evaluate are distributions of educational attainment of the parent (G2) and child (G3) generations, and on the distribution of the births (producing G3 children) by their G2 parental marital status and maternal education. We also compare the G2 generation's truncated (ages 20-39) total fertility rate (TFR) to the regular and truncated TFR in the population for the same period. To perform our model validation analyses, we use the ACS and birth registration (Vital Statistics) estimates from the 2000s and 2010s to best match to the periods of the G2 generation's educational attainment and fertility, as we describe in the Results section below. Both of these external sources have limitations, which we also note where each is used below, and where it is possible to use both sources for our validation (notably the birth distributions by mother's characteristics), we do so. Because both sources provide information primarily on the mother of children born, our validation is focused on G2 women. For the G3 generation, however, we include children of both sexes combined. Their characteristics used in the validation are their parents' marital status at birth and their mother's education.

Results

We first present the educational attainment of women in the (G2) Parent Generation and of both genders combined for the (G3) Child Generation (see Table 1). For the G2 generation, we are able to conduct a comparison of the simulated educational distribution to observed data, as the NLSY79-Youth and NLSY97 cohorts used in our estimation were born in the mid-late 1980s. The simulated G2s therefore correspond approximately to real cohorts who were born 24 years before 2010, for whom we have estimated educational attainment distributions from the ACS. We present the 2010 observed distribution of 24 year old women, because this mirrors the age at which educational attainment is observed in the education-transmission NLSY estimation samples.

[TABLE 1 ABOUT HERE]

We expect to see a closer correspondence of the ACS to the NLSY79+NLSY97 version of the microsimulation model estimates than to the NLSY79-only estimates for three reasons. First, the population representativeness of the NLSY97 cohort, sampled from the US population of 14-19 year olds in 1997, is expected to be better than that of the NLSY79-Youth cohort who were born to women sampled in 1979. The 1980s and 1990s saw substantial immigration especially of the Hispanic population, who on average had substantially lower educational attainment than the US-born population (Rendall and Parker 2014). That is, immigration lowered the educational attainment distribution in the

US compared to what it would have been without those 1980s and 1990s immigration inflows, and we want our microsimulation model to capture that. Estimation that includes the NLSY97 cohort should help to do that whereas estimation with the parents and offspring of the NLSY79 cohort will not. Second, more cumulative attrition is expected the NLSY79-Youth cohort, as both the NLSY79 women and the offspring NLSY79 youth may have attrited, whereas only the NLSY97 youth themselves are potential attritors. Their parents' characteristics are instead obtained from current and retrospective reporting in the initial 1997 wave. Third, sampling error for the combined-sample estimation will be lower than for either sample alone. Additionally, sampling error will be greater in the single-survey NLSY79 estimates because it has a smaller sample (of NLSY79 Youth) compared to the NSLY97 cohort sample size. For the G3 child generation's educational attainment distribution, we have no data to use to evaluate bias (the G3 generation will only complete their education in the 2030s and 2040s), but we are able to compare CIs of microsimulation model outcomes between the combined-survey estimation and single-survey estimation versions of the education-transmission process.

Education Distribution and Fertility of the Simulated Parent Generation (G2) Women, and Education Distribution of the Simulated Child Generation (G2)

As expected, the NLSY79+NLSY97 version of the microsimulation model estimates are generally closer to the observed distribution of 24 year old women's education in 2010 than are the NLSY79-only version estimates. We see no statistically-significant differences between observed and simulated G2 education distributions for the

NLSY79+NLSY97 estimates for those with Less than High School Graduate (7.6% versus 8.5% observed), High School Graduates (30.7% versus 29.7% observed), Some College (32.2% versus 30.1% observed), and College Graduate (29.5% versus 31.5% observed) educational attainments. The NLSY79-only estimates, on the other hand, are substantially and statistically-significantly lower than the observed High School Graduates (21.1% versus 29.7% observed) and substantially higher than the observed Some College (37.7% versus 30.1% observed). With 33.8% College Graduates, added to the 36.2% Some College, the NLSY79-only estimates produce a G2 simulated distribution of 70.0% women with any college education. This is substantially higher than both the 61.6% observed, and the 61.7% that is simulated from the NLSY79+NLSY97 estimation.

To evaluate the efficiency gains of combined-survey estimation over single-survey estimation of education-transmission for the microsimulated G2 women's educational attainment, we compare 95% confidence intervals (CIs) between the NLSY79 and NLSY79+NLSY97 education-transmission estimation versions of the model (see again Table 1). We describe first the differences in CI widths in the educational distribution of the generation G2 women. They are generally narrower for the NLSY79+NLSY97 estimation version, but surprisingly not a lot narrower. The differences in CIs are largest at higher education attainment levels. In particular, for College graduate G2 women, the NLSY79+NLSY97 estimation 95% CI indicates that between 26.9% and 32.9% of all the G2 women are in this category (a 5.0 percentage-point CI), whereas the NLSY79-only estimation 95% CI indicates that between 29.4%

and 37.8% of all the G2 women are in this category (a 7.4 percentage-point CI). For High School Graduate G2 women, the NLSY79+NLSY97 estimation 95% CI indicates that between 27.3% and 35.0% of all the G2 women are in this category (a 7.7 percentage-point CI), whereas the NLSY79-only estimation 95% CI indicates that between 16.9% and 25.0% of all the G2 women are in this category (an 8.1 percentage-point CI width). Reasons for the lower-than-expected difference in CIs when also including the NLSY97 data in a pooled-survey estimation of the education-transmission equations may include additional sample restrictions on the NLSY79-Youth sample when used in the cross-survey multiple imputation (see Appendix 4).

The simulated G2 generation women give birth to 10,252 Children (NLSY79+NLSY97 estimates), for a Cohort Total Fertility Rate of 1.71 (1.73 using the NLSY79 only). This is from fertility, however, only between ages 20 and 39. Given the G2 are born in the 1980s, their peak fertility ages 25-29 and 30-34 will have been approximately experienced in the decade of the 2010s. The US TFR (sum of age-specific fertility rates at ages 15 to 44) during the decade of the 2010s fell from 1.93 in 2010 to 1.82 in 2016 and 1.71 in 2019, as presented in the National Center for Health Statistics (NCHS) series from birth registrations and Census Bureau population estimates (Martin *et al* 2018, 2021). Of the 2016 TFR of 1.82, 1.79 was contributed between ages 20-39, indicating that we lose only a small proportion of offspring by restricting our fertility age range to 20-39. It also suggests our cohort TFR of 1.73 (95% CI: 1.68-1.19) between ages 20-39 is only about 5% lower than the 1.79 of 20-39 year old U.S. age-specific fertility rates as summed for 2016.

G2 Parental Characteristics of Newborn Children (G3)

We present point estimates and confidence intervals in G3 microsimulation model outcomes, and compare again the NLSY79 and NLSY79+NLSY97 education-transition estimation versions of the model (see Table 2). The CIs are again generally narrower for the NLSY79+NLSY97 estimation version, but not very much narrower. For the fraction of G3 children born within marriage, the NLSY79+NLSY97 estimation 95% CI is from a lower limit of 60.2% to an upper limit of 64.2%, of all G3 children (a 4.0 percentage-point CI width), whereas the NLSY79-only estimation 95% CI is from 63.7% to 68.7% (a 5.0 percentage-point CI width).

[TABLE 2 ABOUT HERE]

These confidence intervals are needed for evaluation of bias in the microsimulation model outcomes. In this case of the fraction of G3 children born within marriage, they indicate that for the NLSY79+NLSY97 estimation the marital-birth fraction of 62.2%, the CI (95% CI: 60.2-64.2%) includes the fraction from the Vital Statistics (birth registrations) of 63.2%, but is lower than the 68.1% from the ACS (95% CI: 67.8-68.4%). For the NLSY79 estimation, the 66.4% marital-birth fraction the CI (95% CI: 63.7-68.7%) is above the 63.2% from the birth registrations, but overlaps with the 95% CI from the ACS (67.8-68.4% 95% CI). Both the birth registrations and the ACS have deficiencies in coverage. Some states, most notably California and Texas, are missing from the birth registrations by marital status and maternal education (U.S. DHHS no date). For the ACS, we used the question on whether a woman gave birth in the last 12 months. However, previous analyses (e.g., Rendall *et al* 2018) have shown that there are non-trivial numbers of cases in which newborns identified in an ACS household roster do not match up with reported births to ACS women in that household. We expect that non-marital births may be overrepresented among those, and are accordingly cautious about the accuracy of the ACS-based statistics on proportions of births to married versus unmarried mothers that we generate for our validations. The birth registration system has generally been taken as the authoritative source of marital and non-marital birth fractions in the U.S. (e.g., Solomon-Fears 2014). On this basis, we conclude that the NLSY79+NLSY97 estimation of the education transmission process produces an unbiased estimate of the population marital-birth fraction for G3 children, whereas the NLSY79-only estimation of the education transmission process produces an upwardly-biased estimate of the population marital-birth fraction.

The confidence intervals are widest, unsurprisingly, for those births less numerous in the population, in particular non-marital births within each education group. For example, the fraction of G3 children born non-maritally whose mother has Less than High School Graduate women have a 95% CI of 12.8 percentage points (12.2-25.0%) in the NLSY79+NLSY97 estimation and of 14.3 percentage points (12.7-27.0%) in the NLSY79-only estimation. For G3 born maritally, the confidence interval width for the proportion of their mothers who are College Graduates the 95% CI is 7.9 percentage points (40.7-48.5%) in the NLSY79+NLSY97 estimation and is 11.1 percentage points (46.5-56.6%) in the NLSY79-only estimation.

These CIs are again needed for evaluation of bias in the microsimulation model outcomes. Differences are unfortunately again seen between the birth registrations (Vital Statistics) and ACS benchmarks. The fraction of College Graduate mothers among G3 children born within marriage, however, exhibits little difference between the birth registration (44.7%) and ACS (43.3%) benchmarks. Here, the NLSY79+NLSY97 estimation version of the microsimulation model matches the estimates from the birth registrations of 44.7%, and overlaps with the 95% CI from the ACS (95% CI: 43.0-43.6%). For the NLSY79 estimation, however, the microsimulated 95% CI of 46.5-56.6% (point estimate of 51.5%), lies above the estimates from both the birth registrations and ACS, implying upward bias. This is offset by a downward bias in the NLSY79 version's proportion of High School Graduates mothers among marital births, with 13.7% (95% CI: 10.5-16.8%) compared to the birth registration system's 17.8% and the ACS's 24.2% (95% CI: 23.9-24.4%). The NLSY79+NLSY97 version fraction 23.4% (95% CI: 20.0-27.0%) matches the ACS estimate, but is above the birth registration system's 17.8%.

For the education distribution of mothers of non-maritally-born G3 children, even though the CIs are wide for the NLSY79 estimation version, those point estimates match better to the birth registration distribution than to the NLSY79+NLSY97 estimation version. The opposite is the case when comparing against the ACS non-marital births maternal-education distribution, where the NLSY79+NLSY97 estimation version is closer and with all CIs overlapping between the ACS and the NLSY79+NLSY97 estimates. The biggest G3 children's difference between the NLSY79 and NLSY79+NLSY97 estimation distributions, and also between the ACS and birth registration distributions, is with respect to the High School Graduate proportion of mothers for non-marital births. The NLSY79+NLSY97 estimates are more than 10 percentage points greater than the NLSY79 estimates. The latter (36.6%, 95% CI: 30.0-42.8%) matches well to the birth registration 36.6% but is 8 percentage points lower than the ACS 44.4% (95% CI: 44.0-44.9%) educational attainment distributions. The former, at 47.4% High School Graduates (95% CI: 41.7-53.2%), is close to the ACS estimate but 11 percentage points above the birth registration statistic. In summary, the maternal educational distribution of the G3 generation born outside marriage is not clearly better represented by either the NLSY79 version or the NLSY79+NLSY97 version of the microsimulation model, and part of this lack of clarity lies in conflicting statistics from the birth registration versus ACS data sources.

Summary and Conclusions

We find in evaluating our intergenerational reproduction model outcomes that, first, our simulation reproduces reasonably well the observed education distribution of women in the G2 (parent) generation. The observed education distribution of G2 women, however, is only reproduced well when using the combined-survey estimation of the education-transmission process. Single-survey estimation of the education-transmission process results in too few High School Graduates and too many with Some College and too many College Graduate parent-generation women. Confidence Intervals are also somewhat wider about the single-survey estimation version of the microsimulation outcomes than with combined-survey estimation, but not so wide as to explain the discrepancy in educational distributions between the single-survey microsimulation outcomes and those observed in the population. Second, we find that the completed fertility of G2 women matches well to the corresponding period TFRs in the observed U.S. population.

Third, distributions of the child (G3) generation's fractions born non-maritally, both overall and by mother's education, are reproduced reasonably well by our microsimulation model, both when using the combined-survey estimation and single-survey estimation of the education transmission process. The match is somewhat better using the combined-survey estimation for the fractions born inside and outside marriage, and for the maternal education distribution for marital births, and there are some discrepancies in reproducing the educational attainment distribution of mothers in the case of non-marital births. Overall, compared to combined-survey estimation of the educational-transmission process, single-survey estimation produces more G3 outcomes that exhibit bias relative to external sources on the marital status and maternal education distributions. Single-survey estimation also produces confidence Intervals that are somewhat wider about the G3 children's estimated parental marital status and maternal education outcomes. Our using combined-survey methods for the education-transmission equation therefore is shown to have increased statistical efficiency and reduced bias also for the parental characteristics of the child generation.

The main overall conclusion from our study is that by using a combination of medium- and large-scale survey data sources with a model that allows for annual birth, marriage, and divorce event simulation and a one-time assignment of educational attainment as an intergenerational-transmission process, the U.S. population's educational attainment and births by marital status and maternal education are reproduced in a microsimulation model with a relatively simple structure. Given that no previous intergenerational reproduction model has attempted to evaluate bias against external sources, and given that no previous intergenerational reproduction model has generated confidence intervals or other measures of statistical uncertainty about their estimates, the present study's results represent an encouraging advance in this literature.

The reasonable match of our model outcome distributions to external data sources, however, is achieved only with advances in component-process estimation methodologies. Two-sex estimation of marriage is achieved with a new model that allows appropriately for competition across educational categories for both genders (Menzel 2015; Goyal *et al* 2020). For two of the four component processes, moreover, estimation was performed by the application of pooled cross-survey multiple imputation (Gelman *et al* 1998; Rendall *et al* 2013). The two-sex estimation of divorce by education was only feasible without sizeable downward bias in predicted divorce probabilities by pooling an overall downwardly-biased data source (the SIPP), that nevertheless observes education of both spouses in the year before exposure to divorce, with an overall unbiased data source (the ACS) that observes education of only the wife in the year before exposure to divorce. The intergenerational transmission of educational attainment was estimated by pooling two NLSY cohorts to achieve both larger sample sizes and a sample sufficiently representative of the current U.S. population. This pooled-survey estimation added a more recently-sampled cohort (the NLSY97) to a survey that sampled from a cohort observed in the 1970s, along with the children of this cohort (the NLSY79). We found substantially better matches to current, cross-sectionally observed population outcomes when moving beyond the 1970s-sampled NLSY79 data source alone. This suggests to us that Song and Mare’s (2017) use of a 1960s-sampled data source (the PSID) will have resulted in even more drift from the contemporary U.S. population in its most recent (grandchild) generation’s educational outcomes. This raises an inherent trade-off in models like Song and Mare’s that aim to incorporate direct as well as indirect intergenerational effects across not two but three or more generations (Anderson *et al* 2018). Models like ours instead implicitly or explicitly impose a first-order Markovian assumption that the effects across generations will all occur through contiguous-generation processes. This can be restated as assuming that all effects across a grandparent generation to a grandchild generation will occur through the intermediate, parent generation. We see no practical way to achieve contemporary population representativeness in an intergenerational reproduction model during a time of substantial immigration (or emigration) without making this first-order Markov process assumption.

A related important finding to highlight is that errors in the estimation of one process carry over to generate corresponding errors in outcomes produced by another process in the microsimulation model. In the present case, the upward bias in estimation of the educational-transmission process noted in evaluating the G2 women’s education distributions carried over to upward bias in the maternal education distribution of the G3 children. This upward bias will be produced through both indirect and direct paths. Indirectly, the G2's upwardly biased education distribution is expected to result in more marriage, less divorce, and lower rates of non-marital fertility (Lundberg *et al* 2016; Schwartz and Han 2014). Directly, the G2's upwardly biased education distribution will simply increase the number of more highly educated women at risk of giving birth to G3 children. In particular, using the NLSY79-only version of the microsimulation model resulted in too many G3 children who were estimated to be born to women who went to college. Not only their G2 mother's education, but also their own educational attainment will be correspondingly upwardly biased, given strongly positive intergenerational education correlations. We saw that the educational attainment projected in our simulation for the G3 children is substantially greater for the single-survey NLSY79 version, with 13 percentage points more College Graduates (45.9%) and 13 percentage points fewer High School Graduates (18.2%), compared to the 33.1% College Graduate and 30.7% High School Graduate for the NLSY79+97 version of the simulation. The propagation of error across multiple generations therefore needs to be considered as a major issue for intergenerational reproduction models. This demonstrates again the value to the overall model of both evaluating bias against external data sources and of addressing these biases with strategies such as shown here in our use of combined-survey estimation with at least one recently-sampled data source.

References

Anderson, L.R., P. Sheppard, and C.W.S. Monden (2018) "Grandparent effects on educational outcomes: a systematic review." *Sociological Science* 5:114-142.

Baker, E., Rendall, M.S., and M.M. Weden (2015) “Epidemiological Paradox or Immigrant Vulnerability?” *Demography* 52(4):1295-1320.

Beller, E. (2009) “Bringing intergenerational social mobility into the twenty-first century: why mothers matter” *American Sociological Review* 74:507-528.

Bureau of Labor Statistics (2019a) National Longitudinal Survey of Youth 1979 cohort, 1979-2016 (rounds 1-27). Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Bureau of Labor Statistics (2019b) National Longitudinal Survey of Youth 1997 cohort, 1997-2013 (rounds 1-16). Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Capps, R., Bachmeier, J.D., and J. Van Hook. (2018). Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation. *The ANNALS of the American Academy of Political and Social Science* 677(1):165-179.

Cherlin, A.J. (2010) "Demographic trends in the United States: A review of research in the 2000s." *Journal of Marriage and Family* 72, no. 3: 403-419.

Choo, E., and A. Siow. (2006) “Estimating a Marriage Matching Model with Spillover Effects” *Demography* 43(4):463-490.

DiPrete, T.A., and C. Buchmann (2006) "Gender-Specific Trends in the Value of Education and the Emerging Gender Gap in College Completion" *Demography* 43:1–24.

Fitzgerald, J., P. Gottschalk, and R. Moffitt. (1998) "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33:251-99.

Gale, D and L.S. Shapley. (1962) "College admissions and the stability of marriage" *American Mathematical Monthly* 69(1):9-15.

Gelman, A., King, G., and C. Liu. (1998). Not Asked and Not Answered: Multiple Imputation for Multiple Surveys. *Journal of the American Statistical Association* 94:846-57.

Goldstein, J. R., & Kenney, C. T. (2001). Marriage delayed or marriage forgone? New cohort forecasts of first marriage for US women. *American Sociological Review*, 506-519.

Goyal, S., Handcock, M.S., Yeung, F.C., Jackson, H.M., and M.S. Rendall (2020) "A Practical Revealed Preference Model for Separating Preferences and Availability Effects in Marriage Formation" California Center for Population Research On-Line Working Paper Series PWP-CCPR-2020-010. http://papers.ccpr.ucla.edu/index.php/pwp/article/view/1244/623

Handcock, M. S., Admiraal, R., Yeung, F. C. and Goyal, S. (2021) rpm: Modeling of Revealed Preferences Matchings. Los Angeles, CA. R package version 0.52.

Hazan, M. and H. Zoabi (2014) “Do Highly Educated Women Choose Smaller Families?” *The Economic Journal* 125:1191-1226.

Joffe, M. and Z. Li. (1994) "Male and female factors in fertility." *American Journal of Epidemiology* 140(10): 921-929.

Kennedy, S. and S. Ruggles (2014) “Breaking Up Is Hard to Count: The Rise of Divorce in the United States, 1980-2010” *Demography* 51(2):587-598.

Kravdal, Ø., and R.R. Rindfuss. (2008) "Changing relationships between education and fertility: A study of women and men born 1940 to 1964." *American Sociological review* 73(5): 854-873.

Landale, N.S., and R.S. Oropesa (2007) “Hispanic Families: Stability and Change” *Annual Review of Sociology* 33:381-405.

Lichter, D.T., Price, J.T. and J.M. Swigert (2019) “Mismatches in the Marriage Market” *Journal of Marriage and Family* 82(2): 796-809.

Logan J.A., Hoff, P.D., and Newton M.A. (2008) “Two-Sided Estimation of Mate Preferences for Similarities in Age, Education, and Religion” *Journal of the American Statistical Association* 103(482):559-569.

Lundberg, S., and R.A. Pollak (2015) “The Evolving Role of Marriage: 1950-2010” *The Future of Children* 25(2):29-50.

Lundberg, S., Pollak, R.A., and J. Stearns. (2016) “Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing” *Journal of Economic Perspectives* 30(2):79-102.

Manning, W.D., S.L. Brown, & J.B. Stykes (2016) Same-Sex and Different-Sex

Cohabiting Couple Relationship Stability *Demography* 53:937–953.

Maralani V. (2013) “The Demography of Social Mobility: Black-White Differences in the Process of Educational Reproduction” *American Journal of Sociology* 118(6):1509-1558.

Mare, R.D. (1997) “Differential Fertility, Intergenerational Educational Mobility, and Racial Inequality” *Social Science Research* 26:263–91.

Martin, S.P. (2006) “Trends in Marital Dissolution by Women’s Education in the United States.” *Demographic Research* 15:537-560.

Martin, JA, Hamilton BE, Osterman MJK, Driscoll AK, and P Drake (2018) Births: Final Data for 2016. National Vital Statistics Reports; vol 67 no 1. Hyattsville, MD: National Center for Health Statistics.

Martin JA, Hamilton BE, Osterman MJK, Driscoll AK. (2021) Births: Final Data for 2019. National Vital Statistics Reports; vol 70 no 2. Hyattsville, MD: National Center for Health Statistics.

McLanahan, S. (2004) “Diverging Destinies: How Children Are Faring Under the Second Demographic Transition” *Demography* 41(4): 607-627.

Menzel, K. (2015) “Large matching markets as two-sided demand systems” *Econometrica* 83:897–941.

Moffitt, R.A., and M.S. Rendall (1995) "Cohort Trends in the Lifetime Distribution of Female Family Headship in the United-States, 1968-1985" *Demography* 32(3):407-24.

Musick K and R.D. Mare. (2004) “Family Structure, Intergenerational Mobility, and the

Reproduction of Poverty: Evidence for Increasing Polarization?" *Demography* 41(4):629-48.

Nitsche, N., and H. Brückner. (2020) "Late, But Not Too Late? Postponement of First Birth Among Highly Educated US Women." *European Journal of Population* : 1-33.

Pollak, R. A. (1986) "A reformulation of the two-sex problem" *Demography*. 23: 247–259.

Pollak, R. A. (1990) "Two-Sex Demographic Models" *Journal of Political Economy* 98(2). 399-420.

Pollard, J. H. (1997) "Modelling the interaction between the sexes" *Mathematical and Computer Modelling*. 26:11–24.

Preston S.H. (1974) "Differential Fertility, Unwanted Fertility, and Racial Trends in Occupational Achievement" *American Sociological Review* 39(4):492-506.

Preston, S.H. (1976) "Family Sizes of Children and Family Sizes of Women." *Demography***.** 13(1): 105-14.

Preston S.H. and C. Campbell (1993) "Differential Fertility and the Distribution of Traits: The Case of IQ" *American Journal of Sociology* 98(5):997-1019.

Quesnel-Vallée, A., and S. P. Morgan. (2003) "Missing the target? Correspondence of fertility intentions and behavior in the US." *Population Research and Policy Review* 22(5):497-525.

Raley, R.K. and M.M. Sweeney (2020) "Divorce, Repartnering, and Stepfamilies: A Decade in Review" *Journal of Marriage and Family* 82:81-99.

Rendall, M.S., Ghosh-Dastidar, B., Weden, M.M., Baker, E.H., and Z. Nazarov. (2013). Multiple imputation for combined-survey estimation with incomplete regressors in one but not both surveys. *Sociological Methods and Research* 42(4):483-530.

Rendall, M.S., and S.W. Parker (2014) "Two Decades of Negative Educational Selectivity of Mexican Migrants to the United States" *Population and Development Review* 40(3):421-446.

Rendall M.S., M.M. Weden, J. Brown, and P. Zvavitch (2018) "Why are so many American children born into poverty? Differences in Hispanic and Black new mothers' sources of disadvantage." Maryland Population Research Center Working Paper 2018-002.

Ridder, G., and R.A. Moffitt (2007) "The econometrics of data combination," pp.5469-5547 in J.J. Heckman and E.E. Leamer (Eds.) *Handbook of Econometrics* Vol.6b. Amsterdam: North Holland.

Ruggles, S., K Genadek, R., Goeken, J. Grover, and M. Sobek. (2020) Integrated Public Use Microdata Series: Version 6.0 American Community Survey. Minneapolis: University of Minnesota.

Schafer, J.L. (2003) "Multiple Imputation in Multivariate Problems When the Imputation and Analysis Models Differ" *Statistica Neerlandica* 57(1):19–35.

Schoen, R. (1981) "The harmonic mean as the basis of a realistic two-sex marriage model" *Demography* 18:201–216.

Schomaker, M., & Heumann, C. (2018). "Bootstrap inference when using multiple imputation" *Statistics in Medicine* 37(14):2252-2266.

Schwartz, C.R. (2010) “Pathways to educational homogamy in marital and cohabiting unions” *Demography* 47(3):735–753.

Schwartz, C.R. (2013) “Trends and Variation in Assortative Mating: Causes and Consequences” *Annual Review of Sociology* 39:451-470.

Schwartz, C.R. and H. Han (2014) “The Reversal of the Gender Gap in Education and Trends in Marital Dissolution” *American Sociological Review* 79(4):605-629.

Smock, P., and C.R. Schwartz (2020) “The Demography of Families: A Review of Patterns and Change” *Journal of Marriage and Family* 82:9-34.

Solomon-Fears, C. (2014). “Nonmarital births: An overview.” Congressional Research Service. 7-5700. R43667. https://fas.org/sgp/crs/misc/R43667.pdf

Song, X. and R.D. Mare (2017) “Short-Term and Long-Term Educational Mobility of Families: A Two-Sex Approach” *Demography* 54(1):145-173.

Thomson, E., M. Winkler-Dworak, M. Spielauer, and A. Prskawetz (2012) "Union instability as an engine of fertility? A microsimulation model for France." *Demography* 49(1):175-95.

United States Census Bureau (2014) “SIPP Introduction and History” http://www.census.gov/programs-surveys/sipp/about/sipp-introduction-history.html# Accessed 10/13/2014.

United States Department of Health and Human Services (US DHHS no date), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Division of Vital Statistics. Natality public-use data 1980-2018.

Wolf, D.A. (2001) “The role of microsimulation in longitudinal data analysis” *Special*

Issue on Longitudinal Methodology, Canadian Studies in Population 28(2):313-39.

Yang, Y., and S. P. Morgan (2003) "How big are educational and racial fertility differentials in the US?" *Social Biology* 50(3-4):167-187.

Zvavitch, P., Rendall, M.S., Hurtado-Acuna, C., and R.A. Shattuck. (2020). "Contraceptive Consistency and Poverty after Birth" *Population Research and Policy Review*. Online Access.

Table 1 Education distributions as an adult simulated for parent generation (G2) women and their (G3) children, and observed for G2 women

	Parent Generation (G2) Women[a]					Child Generation (G3)			
	Simulation				Observed Education at age 24 ACS 2010 women[b]	Simulation			
	NLSY79 only		NLSY79+NLSY97			NLSY79 only		NLSY79+NLSY97	
Educational Attainment	number (women)	percent	number (women)	percent	percent	number (both sexes)	percent	number (both sexes)	percent
< High School Graduate	443	7.4	458	7.6	8.5	673	6.5	739	7.2
[95% CI]	[287-636]	[4.8-10.6]	[292-642]	[4.9-10.7]	[8.1,8.9]	[493-887]	[4.8-8.4]	[615-885]	[6.0-8.5]
High School Graduate	1,268	21.1	1,866	31.1	29.7	1,900	18.2	3,149	30.7
[95% CI]	[1016-1497]	[16.9-25.0]	[1639-2100]	[27.3-35.0]	[29.1,30.4]	[1616-2170]	[15.6-20.7]	[2928-3375]	[28.8-32.7]
Some College	2,263	37.7	1,886	31.4	30.1	3,043	29.2	2,960	28.9
[95% CI]	[1965-2596]	[32.8-43.3]	[1606-2132]	[26.8-35.5]	[29.5,30.9]	[2735-3387]	[26.2-32.6]	[2769-3156]	[27.0-30.8]
College Graduate	2,012	33.5	1,787	29.8	31.5	4,788	45.9	3,395	33.1
[95% CI]	[1764-2267]	[29.4-37.8]	[1616-1971]	[26.9-32.9]	[30.9.5,32.3]	[4287-5320]	[41.7-50.5]	[3120-3690]	[30.5-35.8]
All[c]	5,986	99.8	5,997	99.9	99.8	10,404	99.9	10,243	99.9
Cohort Total Fertility Rate						1.73		1.71	
[95% CI]						[1.68-1.79]		[1.66-1.76]	

Notes:

a. The Parent Generation (G2) women include those who had no children.
b. American Community Survey are from IPUMS-ACS (Ruggles et al 2020). Estimates are weighted.
c. Totals are summed from bootstrapped median values of each education category, and therefore do not necessarily add exactly to 6,000 G2 women. Percentage distributions may not add to 100 due to rounding.

Table 2 (G2) Parental Marital Status and Maternal Education distributions of (G3) children

	American Community Survey 2010-2019 Education of Mothers Aged 20-39 Born Between 1980-1990 [a]	Vital Statistics 2010-2018 Education of Mothers Aged 20-39	Education of G2 Mothers, NLSY 79 Simulation [b]	Education of G2 Mothers, NLSY 79 + NLSY 97 Simulation [b, c]
All Births				
Less than High School Graduate	10.3	13.9	9.8	10.3
[95% CI]	[10.1-10.5]		[6.3-13.9]	[6.6-14.1]
High School Graduate	30.6	24.9	21.5	32.4
[95% CI]	[30.4-30.9]		[17.2-25.5]	[28.4-36.7]
Some College	26.5	30.1	31.1	26.2
[95% CI]	[26.2-26.7]		[26.6-36.5]	[22.1-30.3]
College Graduate	32.6	31.1	37.4	31.0
[95% CI]	[32.2-32.9]		[33.1-42.1]	[28.0-34.2]
Marital Births	68.1	63.2	66.1	62.2
[95% CI]	[67.8-68.4]		[63.7-68.7]	[60.2-64.2]
Less than High School Graduate	7.7	9.2	4.6	5.2
[95% CI]	[7.5-7.9]		[2.8-6.9]	[3.1-7.6]
High School Graduate	24.2	17.8	13.7	23.4
[95% CI]	[23.9-24.4]		[10.5-16.8]	[20.0-27.0]
Some College	24.8	28.4	30.1	26.9
[95% CI]	[24.5-25.1]		[25.6-35.4]	[22.6-31.0]
College Graduate	43.3	44.7	51.5	44.4
[95% CI]	[43.0-43.6]		[46.5-56.6]	[40.7-48.6]
Non-Marital Births	31.9	36.8	33.9	37.8
[95% CI]	[31.6-32.2]		[31.3-36.3]	[35.8-39.8]
Less than High School Graduate	16.0	22.1	19.9	18.6
[95% CI]	[15.6-16.4]		[12.7-27.0]	[12.2-25.0]
High School Graduate	44.4	37.1	36.6	47.4
[95% CI]	[44.0-44.9]		[30.0-42.8]	[41.7-53.2]
Some College	29.9	33.0	33.3	25.1
[95% CI]	[29.5-30.3]		[27.5-40.2]	[20.7-30.0]
College Graduate	9.7	7.7	10.0	8.9
[95% CI]	[9.4-9.9]		[8.0-12.5]	[7.4-10.5]

Notes:

a. American Community Survey (ACS) data are from IPUMS-ACS (Ruggles et al 2020). Estimates are weighted.
b. NLSY79: National Longitudinal Survey of Youth 1979 cohort (with Young Adult sample).
c. NLSY97: National Longitudinal Survey of Youth 1997 Cohort.

Appendix Table A1 Two-sided Logit Estimates of Annual Marriage

Woman Education	Woman Age	Man Education	Man Age	Estimate	Std. Error
Less than High School	Under 25	Less than High School	Under 25	-0.80	0.17
Less than High School	25-29	Less than High School	Under 25	-1.90	0.34
Less than High School	30-34	Less than High School	Under 25	-2.65	0.81
Less than High School	35+	Less than High School	Under 25	-2.99	0.82
High School	Under 25	Less than High School	Under 25	-1.95	0.15
High School	25-29	Less than High School	Under 25	-3.28	0.39
High School	30-34	Less than High School	Under 25	-5.81	0.45
High School	35+	Less than High School	Under 25	-5.24	0.55
Some College	Under 25	Less than High School	Under 25	-3.23	0.22
Some College	25-29	Less than High School	Under 25	-3.42	0.54
Some College	30-34	Less than High School	Under 25	-5.22	0.56
Some College	35+	Less than High School	Under 25	-4.51	0.88
College	Under 25	Less than High School	Under 25	-4.57	0.67
College	25-29	Less than High School	Under 25	-4.69	0.62
College	30-34	Less than High School	Under 25	-6.00	0.00
College	35+	Less than High School	Under 25	-6.00	0.00
Less than High School	Under 25	Less than High School	25-29	-0.90	0.21
Less than High School	25-29	Less than High School	25-29	-0.79	0.21
Less than High School	30-34	Less than High School	25-29	-2.06	0.43
Less than High School	35+	Less than High School	25-29	-3.07	0.75
High School	Under 25	Less than High School	25-29	-2.09	0.19
High School	25-29	Less than High School	25-29	-1.89	0.19
High School	30-34	Less than High School	25-29	-2.59	0.41
High School	35+	Less than High School	25-29	-4.07	0.60
Some College	Under 25	Less than High School	25-29	-3.25	0.26
Some College	25-29	Less than High School	25-29	-2.52	0.35
Some College	30-34	Less than High School	25-29	-3.47	0.56
Some College	35+	Less than High School	25-29	-4.90	0.62
College	Under 25	Less than High School	25-29	-3.75	0.56
College	25-29	Less than High School	25-29	-4.79	0.47
College	30-34	Less than High School	25-29	-5.49	0.53
College	35+	Less than High School	25-29	-6.00	0.00
Less than High School	Under 25	Less than High School	30-34	-1.64	0.41
Less than High School	25-29	Less than High School	30-34	-0.51	0.22
Less than High School	30-34	Less than High School	30-34	-0.42	0.26
Less than High School	35+	Less than High School	30-34	-1.06	0.33
High School	Under 25	Less than High School	30-34	-3.47	0.33
High School	25-29	Less than High School	30-34	-1.80	0.24
High School	30-34	Less than High School	30-34	-1.85	0.23
High School	35+	Less than High School	30-34	-3.24	0.61
Some College	Under 25	Less than High School	30-34	-5.51	0.42
Some College	25-29	Less than High School	30-34	-2.03	0.28
Some College	30-34	Less than High School	30-34	-2.12	0.36
Some College	35+	Less than High School	30-34	-2.80	0.58
College	Under 25	Less than High School	30-34	-3.84	0.77

College	25-29	Less than High School	30-34	-5.96	0.41
College	30-34	Less than High School	30-34	-3.23	0.71
College	35+	Less than High School	30-34	-3.50	0.98
Less than High School	Under 25	Less than High School	35+	-2.50	0.52
Less than High School	25-29	Less than High School	35+	-1.67	0.41
Less than High School	30-34	Less than High School	35+	-0.85	0.27
Less than High School	35+	Less than High School	35+	-0.54	0.31
High School	Under 25	Less than High School	35+	-5.87	0.42
High School	25-29	Less than High School	35+	-2.28	0.34
High School	30-34	Less than High School	35+	-1.80	0.29
High School	35+	Less than High School	35+	-1.89	0.28
Some College	Under 25	Less than High School	35+	-6.00	0.00
Some College	25-29	Less than High School	35+	-2.84	0.38
Some College	30-34	Less than High School	35+	-2.38	0.45
Some College	35+	Less than High School	35+	-1.92	0.34
College	Under 25	Less than High School	35+	-5.52	0.49
College	25-29	Less than High School	35+	-5.92	0.39
College	30-34	Less than High School	35+	-4.39	0.80
College	35+	Less than High School	35+	-3.22	0.99
Less than High School	Under 25	High School	Under 25	-1.93	0.17
Less than High School	25-29	High School	Under 25	-3.96	0.46
Less than High School	30-34	High School	Under 25	-3.38	0.59
Less than High School	35+	High School	Under 25	-4.58	0.78
High School	Under 25	High School	Under 25	-1.51	0.06
High School	25-29	High School	Under 25	-2.87	0.16
High School	30-34	High School	Under 25	-3.60	0.30
High School	35+	High School	Under 25	-5.25	0.43
Some College	Under 25	High School	Under 25	-2.31	0.08
Some College	25-29	High School	Under 25	-2.91	0.17
Some College	30-34	High School	Under 25	-4.16	0.35
Some College	35+	High School	Under 25	-6.00	0.04
College	Under 25	High School	Under 25	-2.92	0.16
College	25-29	High School	Under 25	-3.67	0.20
College	30-34	High School	Under 25	-5.66	0.32
College	35+	High School	Under 25	-5.33	0.56
Less than High School	Under 25	High School	25-29	-2.21	0.25
Less than High School	25-29	High School	25-29	-2.00	0.22
Less than High School	30-34	High School	25-29	-3.37	0.44
Less than High School	35+	High School	25-29	-4.15	0.66
High School	Under 25	High School	25-29	-1.33	0.07
High School	25-29	High School	25-29	-1.16	0.08
High School	30-34	High School	25-29	-2.16	0.16
High School	35+	High School	25-29	-4.12	0.32
Some College	Under 25	High School	25-29	-2.23	0.09
Some College	25-29	High School	25-29	-1.37	0.08
Some College	30-34	High School	25-29	-2.21	0.17
Some College	35+	High School	25-29	-3.82	0.39
College	Under 25	High School	25-29	-2.66	0.17
College	25-29	High School	25-29	-1.77	0.10

College	30-34	High School	25-29	-2.88	0.20
College	35+	High School	25-29	-4.13	0.43
Less than High School	Under 25	High School	30-34	-2.44	0.36
Less than High School	25-29	High School	30-34	-1.49	0.27
Less than High School	30-34	High School	30-34	-1.81	0.27
Less than High School	35+	High School	30-34	-1.99	0.40
High School	Under 25	High School	30-34	-2.70	0.17
High School	25-29	High School	30-34	-1.51	0.11
High School	30-34	High School	30-34	-0.93	0.11
High School	35+	High School	30-34	-2.37	0.20
Some College	Under 25	High School	30-34	-3.53	0.20
Some College	25-29	High School	30-34	-1.33	0.11
Some College	30-34	High School	30-34	-1.21	0.14
Some College	35+	High School	30-34	-2.16	0.21
College	Under 25	High School	30-34	-4.00	0.47
College	25-29	High School	30-34	-1.83	0.13
College	30-34	High School	30-34	-1.82	0.15
College	35+	High School	30-34	-2.58	0.28
Less than High School	Under 25	High School	35+	-3.12	0.33
Less than High School	25-29	High School	35+	-3.04	0.47
Less than High School	30-34	High School	35+	-2.15	0.42
Less than High School	35+	High School	35+	-2.52	0.46
High School	Under 25	High School	35+	-3.35	0.22
High School	25-29	High School	35+	-2.19	0.16
High School	30-34	High School	35+	-1.39	0.13
High School	35+	High School	35+	-1.34	0.13
Some College	Under 25	High School	35+	-4.41	0.35
Some College	25-29	High School	35+	-2.52	0.20
Some College	30-34	High School	35+	-1.70	0.17
Some College	35+	High School	35+	-1.77	0.18
College	Under 25	High School	35+	-5.93	0.40
College	25-29	High School	35+	-3.09	0.30
College	30-34	High School	35+	-1.84	0.17
College	35+	High School	35+	-1.97	0.20
Less than High School	Under 25	Some College	Under 25	-4.25	0.56
Less than High School	25-29	Some College	Under 25	-5.16	0.62
Less than High School	30-34	Some College	Under 25	-6.00	0.03
Less than High School	35+	Some College	Under 25	-6.00	0.00
High School	Under 25	Some College	Under 25	-2.50	0.11
High School	25-29	Some College	Under 25	-4.01	0.29
High School	30-34	Some College	Under 25	-5.35	0.41
High School	35+	Some College	Under 25	-6.00	0.02
Some College	Under 25	Some College	Under 25	-2.04	0.07
Some College	25-29	Some College	Under 25	-3.17	0.20
Some College	30-34	Some College	Under 25	-4.70	0.49
Some College	35+	Some College	Under 25	-6.00	0.31
College	Under 25	Some College	Under 25	-2.26	0.12
College	25-29	Some College	Under 25	-3.79	0.19
College	30-34	Some College	Under 25	-5.13	0.49

College	35+	Some College	Under 25	-5.84	0.32
Less than High School	Under 25	Some College	25-29	-3.54	0.68
Less than High School	25-29	Some College	25-29	-2.52	0.31
Less than High School	30-34	Some College	25-29	-2.97	0.92
Less than High School	35+	Some College	25-29	-6.00	0.00
High School	Under 25	Some College	25-29	-2.30	0.13
High School	25-29	Some College	25-29	-1.51	0.11
High School	30-34	Some College	25-29	-2.71	0.25
High School	35+	Some College	25-29	-3.23	0.41
Some College	Under 25	Some College	25-29	-1.93	0.11
Some College	25-29	Some College	25-29	-0.60	0.08
Some College	30-34	Some College	25-29	-2.10	0.18
Some College	35+	Some College	25-29	-3.06	0.33
College	Under 25	Some College	25-29	-1.48	0.12
College	25-29	Some College	25-29	-0.87	0.08
College	30-34	Some College	25-29	-2.18	0.19
College	35+	Some College	25-29	-3.52	0.38
Less than High School	Under 25	Some College	30-34	-3.07	0.54
Less than High School	25-29	Some College	30-34	-2.55	0.73
Less than High School	30-34	Some College	30-34	-2.22	0.60
Less than High School	35+	Some College	30-34	-2.89	0.88
High School	Under 25	Some College	30-34	-2.96	0.21
High School	25-29	Some College	30-34	-1.44	0.16
High School	30-34	Some College	30-34	-1.16	0.16
High School	35+	Some College	30-34	-2.90	0.35
Some College	Under 25	Some College	30-34	-2.78	0.18
Some College	25-29	Some College	30-34	-0.59	0.10
Some College	30-34	Some College	30-34	-0.53	0.12
Some College	35+	Some College	30-34	-1.49	0.20
College	Under 25	Some College	30-34	-3.24	0.40
College	25-29	Some College	30-34	-1.01	0.10
College	30-34	Some College	30-34	-0.56	0.11
College	35+	Some College	30-34	-2.04	0.24
Less than High School	Under 25	Some College	35+	-3.45	0.77
Less than High School	25-29	Some College	35+	-3.83	0.88
Less than High School	30-34	Some College	35+	-2.47	0.72
Less than High School	35+	Some College	35+	-3.61	1.17
High School	Under 25	Some College	35+	-3.76	0.46
High School	25-29	Some College	35+	-2.39	0.25
High School	30-34	Some College	35+	-1.71	0.21
High School	35+	Some College	35+	-1.77	0.20
Some College	Under 25	Some College	35+	-3.47	0.35
Some College	25-29	Some College	35+	-1.50	0.17
Some College	30-34	Some College	35+	-1.09	0.16
Some College	35+	Some College	35+	-1.22	0.18
College	Under 25	Some College	35+	-4.49	0.59
College	25-29	Some College	35+	-1.89	0.20
College	30-34	Some College	35+	-0.90	0.14
College	35+	Some College	35+	-1.23	0.17

Less than High School	Under 25	College	Under 25	-4.40	0.98
Less than High School	25-29	College	Under 25	-6.00	0.00
Less than High School	30-34	College	Under 25	-6.00	0.00
Less than High School	35+	College	Under 25	-6.00	0.00
High School	Under 25	College	Under 25	-3.37	0.27
High School	25-29	College	Under 25	-4.61	0.62
High School	30-34	College	Under 25	-6.00	0.00
High School	35+	College	Under 25	-6.00	0.00
Some College	Under 25	College	Under 25	-2.39	0.14
Some College	25-29	College	Under 25	-4.20	0.46
Some College	30-34	College	Under 25	-5.97	0.40
Some College	35+	College	Under 25	-6.00	0.00
College	Under 25	College	Under 25	-0.15	0.07
College	25-29	College	Under 25	-1.96	0.14
College	30-34	College	Under 25	-6.00	0.39
College	35+	College	Under 25	-5.17	0.65
Less than High School	Under 25	College	25-29	-4.99	0.60
Less than High School	25-29	College	25-29	-4.44	0.63
Less than High School	30-34	College	25-29	-4.87	0.77
Less than High School	35+	College	25-29	-6.00	0.00
High School	Under 25	College	25-29	-3.02	0.20
High School	25-29	College	25-29	-2.58	0.21
High School	30-34	College	25-29	-3.28	0.58
High School	35+	College	25-29	-5.60	0.43
Some College	Under 25	College	25-29	-2.49	0.13
Some College	25-29	College	25-29	-1.41	0.11
Some College	30-34	College	25-29	-2.79	0.26
Some College	35+	College	25-29	-4.32	0.75
College	Under 25	College	25-29	-0.60	0.07
College	25-29	College	25-29	0.43	0.04
College	30-34	College	25-29	-1.15	0.11
College	35+	College	25-29	-2.92	0.36
Less than High School	Under 25	College	30-34	-4.08	0.88
Less than High School	25-29	College	30-34	-2.98	0.51
Less than High School	30-34	College	30-34	-2.79	0.97
Less than High School	35+	College	30-34	-3.62	1.32
High School	Under 25	College	30-34	-3.80	0.43
High School	25-29	College	30-34	-2.56	0.20
High School	30-34	College	30-34	-2.09	0.25
High School	35+	College	30-34	-3.12	0.37
Some College	Under 25	College	30-34	-3.41	0.25
Some College	25-29	College	30-34	-1.51	0.14
Some College	30-34	College	30-34	-1.05	0.15
Some College	35+	College	30-34	-2.42	0.33
College	Under 25	College	30-34	-1.66	0.17
College	25-29	College	30-34	0.00	0.06
College	30-34	College	30-34	0.57	0.06
College	35+	College	30-34	-0.76	0.13
Less than High School	Under 25	College	35+	-3.71	1.25

Less than High School	25-29	College	35+	-4.29	0.78
Less than High School	30-34	College	35+	-2.86	0.86
Less than High School	35+	College	35+	-4.49	0.93
High School	Under 25	College	35+	-4.28	0.46
High School	25-29	College	35+	-2.97	0.30
High School	30-34	College	35+	-1.99	0.24
High School	35+	College	35+	-2.55	0.33
Some College	Under 25	College	35+	-3.70	0.31
Some College	25-29	College	35+	-2.10	0.22
Some College	30-34	College	35+	-1.54	0.22
Some College	35+	College	35+	-1.59	0.20
College	Under 25	College	35+	-3.68	0.51
College	25-29	College	35+	-1.20	0.12
College	30-34	College	35+	0.10	0.09
College	35+	College	35+	-0.03	0.09
N					382,079

Source: American Community Survey

Appendix 2 Divorce Estimation

The Survey of Income and Program Participation (SIPP)

Each SIPP panel covers approximately four years and respondents are interviewed every four months. To match the annual event structure of our microsimulation mode, we code an annual outcome variable for divorce versus no divorce in the SIPP. We use the 2004 and 2008 SIPP Panels. The 2004 SIPP Panel respondents were interviewed every four months beginning in the Spring of 2004 and ending in the Fall of 2007. The 2008 SIPP Panel respondents were interviewed every four months beginning in the Fall of 2008 and ending in the Fall of 2013. Since respondents are interviewed every four months, three waves of a panel are equal to 12 months. In 2004 we use waves 1-10 (there are 12 possible waves) and in 2008 we use waves 1-13 (there are 16 possible waves) which correspond to the years 2003-2007 and 2008-2012 to keep exact 12-month cycles to best match the annual structure of the ACS.

2004 SIPP Panel Wave	1 2 3 4 2004-2005	5 6 7 2005-2006	8 9 10 2006-2007	
2008 SIPP Panel Wave	1 2 3 4 2008-2009	5 6 7 2009-2010	8 9 10 2010-2011	11 12 13 2011-2012

We observe the changes in relationship status between waves 1 and 4, 4 and 7, 7 and 10, and 10 and 13. At wave 1, our sample is restricted to only those who are married and a partner in the household. Therefore, we do not use any newcomers into the sample (i.e. people who move into an already sampled household). To identify a divorce, we start with the all married, spouse present sample of wave 1. The year (2004-2005 of the 2004 Panel and 2008-2009 of the 2008 Panel) between wave 1 and wave 4 will be used as an example to explain how marital status and divorce events were determined annually. If a respondent A is married at wave 1 to their partner, person B, and married[1] at wave 4 to the same person B, they were at risk of divorce between those waves but remained married. A couple would be identified as getting a divorce between wave 1 and 4 if in wave 4 their status is identified as “divorced”.

[1] If the couple is separated or married, spouse absent or separated we still consider them to be married in waves 4, 7, 10, and 13. The couple has to explicitly identify themselves as divorced in order to be considered divorced that calendar year. This is to best match to the ACS which only asks about the strict status of “divorced”.

Appendix Table A2a: Descriptives of those at Risk for Divorce Annually during 2004-2011 in the United States, ages 20-39

	All		Divorcing		Remaining Married	
	SIPP	ACS	SIPP	ACS	SIPP	ACS
Women's Education		**		**		**
Less than High School Graduate	0.109	0.101	0.062	0.082	0.110	0.101
High School Graduate	0.197	0.306	0.267	0.382	0.196	0.303
Some College	0.346	0.276	0.425	0.325	0.345	0.275
College Graduate	0.347	0.318	0.240	0.212	0.349	0.321
Men's Education						
Less than High School Graduate	0.130		0.108		0.130	
High School Graduate	0.226		0.277		0.225	
Some College	0.330		0.412		0.329	
College Graduate	0.314		0.203		0.316	
Divorce	0.013	0.028 **				
Sample N	26,098	212,542	372	5,227	25,726	207,315

Notes: All proportions are weighted. Group differences from chi-squared (SIPP vs. ACS), ** p<0.01.

Sources: Survey of Income and Program Participation (AIPP) 2004 and 2008 panels and the American Community Survey (ACS) 2011.

Appendix Table A2b: Model Fit Statistics for Pooled Logistic Regressions of Divorce on Education

Pooled SIPP and ACS Model Fit statistics	no SIPP intercept or regressor interaction	SIPP intercept, no regressor interaction		SIPP intercept and regressor interaction
AIC	57,482.6	57,273.3	*	57,277.0
BIC	57,524.2	57,325.2	*	57,360.0

* best fitting model (lower = better fit)

Sources: Survey of Income and Program Participation (SIPP) 2004, 2008. American Community Survey (ACS) 2009-2011.

Appendix Table A2c: Logistic Regression of Divorce on Own and Spouse's Education in the United States, ages 20-39 2004-2011

	SIPP			ACS			ACS + SIPP		
	Coefficient		S.E.	Coefficient		S.E.	Coefficient		S.E.
Women's Education									
Less than High School Graduate	-0.103		0.250	0.201	**	0.074	0.107		0.146
High School Graduate	0.683	**	0.170	0.646	**	0.045	0.435	**	0.104
Some College	0.582	**	0.151	0.583	**	0.046	0.378	**	0.086
Men's Education									
Less than High School Graduate							0.140		0.234
High School Graduate							0.446	*	0.224
Some College							0.463	*	0.185
SIPP							-0.728	**	0.056
Constant	-4.685	**	0.119	-3.969		0.034	-4.142	**	0.079
Sample N	26,098			212,542			238,640		

Notes: Estimates are weighted. ** p<0.01 and * p<0.05. ACS+SIPP estimates use cross-survey multiple imputation.

Sources: Survey of Income and Program Participation 2004 and 2008 panels and the American Community Survey 2011.

Appendix Table A3 Logit Regression Estimates of Annual Birth

	Non-Marital Birth		Marital Birth	
Woman's Education (Ref. Less than High School Graduate)				
High School Graduate (HS)	-0.41	(0.01)	-0.26	(0.02)
Some College (SCO)	-1.00	(0.01)	-0.27	(0.03)
College Graduate (CO)	-1.73	(0.02)	-0.12	(0.05)
Parity (Ref. 0 Prior Births)				
1 Prior Birth	0.74	(0.01)	0.35	(0.01)
2 Prior Births	0.51	(0.02)	-0.58	(0.01)
3+ Prior Births	0.38	(0.02)	-0.76	(0.01)
Woman's Age	-0.06	(0.00)	-0.07	(0.00)
Man's Education (Ref. Less than High School)				
High School (HS)			-0.21	(0.02)
Some College (SCO)			-0.24	(0.04)
College (CO)			-0.01	(0.06)
Interactions of Parent Education				
Mother HS Father HS			0.12	(0.03)
Mother HS Father SCO			0.20	(0.05)
Mother HS Father CO			0.12	(0.07)
Mother SCO Father HS			0.12	(0.04)
Mother SCO Father SCO			0.24	(0.05)
Mother SCO Father CO			0.21	(0.07)
Mother CO Father HS			0.17	(0.05)
Mother CO Father SCO			0.26	(0.06)
Mother CO Father CO			0.22	(0.08)
Intercept	-0.62	(0.03)	0.73	(0.02)
N	2,184,117		1,995,207	

Notes:

Standard errors in parentheses

Source:

American Community Survey 2000-2011, 2013-2017

Appendix 4 Educational Transmission

Data

Data for this study comes from U.S. men and women born in the early 1980s in the National Longitudinal Survey of Youth 1979 Young Adult Sample and 1997 Sample (NLSY79, NLSY97). These two cohort panel surveys have the major strength of collecting nationally representative longitudinal data for those cohorts, and having education information on both a parent and child generation.

Starting in 1979, the NLSY79 (Bureau of Labor Statistics 2019a) interviewed a nationally representative sample of 12,686 young men and women born in 1957 through 1964. To best match the 1997 sample, we use the NLSY79 Young Adults survey, alternately described as the NLSY79 Youth. The youth sample includes children born to the female respondents of the NLSY79 and has been collected biennial since 1986.

Starting in 1997, the NLSY97 (Bureau of Labor Statistics 2019b) interviewed a nationally representative sample of 8,984 individuals who were between ages 12 and 16. 92.1% of eligible respondents completed the first round, 1997, interview. Black and/or Hispanic and Latino populations were oversampled. Respondents were interviewed annually until 2011 and biennially since. Approximately 83% of the 1997 sample was interviewed in 2011.

Measures

Education of child generation

Education is measured similarly in both the NLSY79 Young Adult sample and the NLSY97 sample. In the NLSY79 Young Adult, the respondents are interviewed biennially, we take the maximum education completed at age 24-28. In the NLSY97, respondents' own education represents their education at age 24.

Parent or Parents' Education at the Child's Adolescence and at Child's Birth

The NLSY79 respondents record their education status every two years per the NLSY79 survey protocol. When the NLSY79 Young Adult is 14 years old, we record the age of their parents in the NLSY79 survey as the parents' education during the child's adolescence. In the NLSY97, the parental questionnaire administered in 1997 records the parents' education at the time of interview, when the respondents themselves are 12-16 years old. Mother's and father's education at the time of the child's birth is recorded only in the NLSY79 sample. We use the educational attainment variable in the first survey year after the birth of the child.

Marital Status at Birth of child generation

Using the birth date of the NLSY97 cohort member child and the start and end dates of marriages of their parents, we determine whether the birth of the NLSY97 cohort member was to a married or unmarried mother. The NLSY79 respondents record their marriage entry and dissolution as well as the birth of their child. If the birth of the NLSY79 Young Adult occurred during a current marriage of the NLSY79 cohort member, that birth was considered to be a marital birth. For the NLSY97 cohort members, in the initial, 1997 wave, a parent of the cohort member was asked to complete a parental questionnaire which includes a marital history.

Imputation

Education of the parent at the time of the child's birth is needed for the microsimulation model's intergenerational-transmission equation, but is observed only in the NLSY79. Education of the mother at the time of the child's birth is needed for non-marital births. Education of the mother and the father at the time of the child's birth is needed for marital births. Cross-survey multiple imputation is used to impute to every NLSY97 cohort member the education of the parent at the time of the child's birth based on the education of the mother (non-marital birth) or the mother and father (marital birth). The imputation equations are estimated on the NLSY79 data, and coefficients from this equation used to multiply impute a value of parent education at the time of the child's birth to each NLSY97 cohort member.

References

Bureau of Labor Statistics (2019a) National Longitudinal Survey of Youth 1979 cohort, 1979-2016 (rounds 1-27). Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Bureau of Labor Statistics (2019b) National Longitudinal Survey of Youth 1997 cohort, 1997-2013 (rounds 1-16). Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Appendix Table A4a: Descriptives of Women in the United States at age 24 born in 1980-1984.

	NLSY 79	NLSY 97							
Marital Birth Proportion	0.80	0.78							
Sample N	1,908	5,290							
	Marital Birth			**Nonmarital Birth**			**All**		
	NLSY 79	**NLSY 97**		**NLSY 79**	**NLSY 97**		**NLSY 79**	**NLSY 97**	
Female	0.47	0.49		0.49	0.49		0.47	0.49	
Own Education			**			**			**
< High School Graduate	0.04	0.06		0.17	0.19		0.06	0.09	
High School Graduate	0.19	0.36		0.34	0.47		0.22	0.39	
Some College	0.29	0.28		0.34	0.26		0.30	0.27	
College Graduate	0.48	0.30		0.15	0.08		0.41	0.25	
Dad's Education at Child's Adolescence			**						
< High School Graduate	0.08	0.14		-	-		-	-	
High School Graduate	0.38	0.35		-	-		-	-	
Some College	0.22	0.21		-	-		-	-	
College Graduate	0.31	0.30		-	-		-	-	
Mom's Education at Child's Adolescence			**			**			**
< High School Graduate	0.03	0.11		0.17	0.33		0.05	0.16	
High School Graduate	0.40	0.34		0.53	0.40		0.43	0.35	
Some College	0.27	0.27		0.21	0.21		0.26	0.26	
College Graduate	0.31	0.28		0.09	0.07		0.26	0.23	
Household Composition at child's age 2			**			**			**
Bio mom+ Dad	0.73	0.91		0.25	0.12		0.64	0.74	
Other	0.27	0.09		0.75	0.88		0.36	0.26	
Dad Education at Birth									
< High School Graduate	0.09	-		-	-		-	-	
High School Graduate	0.41	-		-	-		-	-	
Some College	0.21	-		-	-		-	-	
College Graduate	0.29	-		-	-		-	-	
Mom Education at Birth									
< High School Graduate	0.05	-		0.20	-		0.08	-	
High School Graduate	0.45	-		0.56	-		0.47	-	
Some College	0.24	-		0.18	-		0.23	-	
College Graduate	0.26	-		0.05	-		0.22	-	
N	1,313	3,758		595	1,532		1,908	5,290	

Notes: All proportions are weighted

Group differences from chi-squared (NLSY79 vs. NLSY97), ** p<0.01.

Sources: National Longitudinal Survey of Youth 1979, 1997

Appendix Table A4b. Model Fit Statistics for Pooled Logistic Regressions of Own Education on Parental Education at Adolecense of Child

	Pooled NLSY79 and NLSY97 Model Fit statistics	No 79 intercept or regressor interaction	79 intercept, no regressor interaction	79 intercept and regressor interaction
Non-Marital Births	AIC	5,152.3	5,130.3 *	5,132.4
	BIC	5,254.2	5,249.2 *	5,336.3
Martial Births	AIC	11,399.8	11,218.2	11,193.4 *
	BIC	11,576.2	11,414.2 *	11,546.1

Notes: * best fitting model (lower= better fit)

Sources: National Longitudinal Survey of Youth 1979, 1997

Appendix Table A4c. Multinomial Logistic Regression of Own Education at age 24 on Mother's and Father's Education by Marital Status at Birth

Own Education Outcome[a]		Non Marital Births				Marital Births			
		NLSY 79		NLSY 79 + 97		NLSY 79		NLSY 79 + 97	
		Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
High School Graduate	**Intercept**	-0.43	0.55	0.54	0.11	0.61	0.38	0.41	0.15
	Mother's Education at Birth[a]								
	High School Graduate	1.04	0.34	0.58	0.14	0.80	0.37	0.89	0.16
	Some College	1.72	0.52	0.69	0.26	0.69	0.47	1.16	0.24
	College Graduate	2.24	0.93	1.46	0.70	1.23	0.84	2.16	0.48
	Fathers's Education at Birth[a]								
	High School Graduate					0.37	0.32	0.95	0.16
	Some College					1.01	0.56	0.83	0.21
	College Graduate					0.76	0.75	2.11	0.43
	Own Gender Female	0.22	0.32	-0.07	0.12	-0.40	0.30	-0.02	0.13
Some College	**Intercept**	-0.80	0.52	-0.28	0.13	0.68	0.38	0.24	0.16
	Mother's Education at Birth								
	High School Graduate	0.99	0.34	1.03	0.16	0.69	0.37	1.26	0.17
	Some College	2.44	0.48	1.66	0.26	1.13	0.46	1.80	0.25
	College Graduate	2.34	0.90	2.70	0.71	1.45	0.82	2.89	0.48
	Fathers's Education at Birth								
	High School Graduate					1.00	0.33	0.54	0.16
	Some College					2.07	0.56	0.61	0.21
	College Graduate					1.94	0.73	2.05	0.43
	Own Gender Female	0.39	0.32	-0.21	0.13	-1.01	0.29	-0.38	0.14
College Graduate	**Intercept**	-2.15	0.70	-1.91	0.25	-1.33	0.56	-0.62	0.20
	Mother's Education at Birth								
	High School Graduate	1.70	0.49	1.75	0.28	2.09	0.53	2.12	0.22
	Some College	3.19	0.60	2.59	0.33	2.92	0.60	3.16	0.28
	College Graduate	2.17	1.04	3.60	0.77	4.02	0.89	5.04	0.49
	Fathers's Education at Birth								
	High School Graduate					1.30	0.37	-0.02	0.17
	Some College					2.62	0.58	0.33	0.22
	College Graduate					3.50	0.74	2.20	0.43
	Own Gender Female	0.24	0.40	-0.44	0.18	-1.04	0.30	-0.71	0.14
	N	595		2,127		1,313		5,071	

Notes: All results are weighted.

NLSY79+97 estimates are from pooled cross-survey multiple imputation estimation.

a. Reference education category is Less Than High School Graduate.

Sources: National Longitudinal Survey of Youth 1979, 1997.