

And Their Children after Them? The Effect of College on Educational Reproduction¹

Matthew Lawrence
Middlebury College

Richard Breen
of Oxford University

Conventional analyses of social mobility and status reproduction retrospectively compare an outcome of individuals to a characteristic of their parents. By ignoring the mechanisms of family formation and excluding childless individuals, conventional approaches introduce selection bias into estimates of how characteristics in one generation affect an outcome in the next. The prospective approach introduced here integrates the effects of college on marriage and fertility into the reproduction of educational outcomes. Marginal structural models with inverse probability of treatment weighting are used with data from the Wisconsin Longitudinal Study to estimate the causal effect of pathways linking graduating from college with having a child who graduates from college. Results show that college increases male graduates' probability of having a child who completes college; for female graduates there is no effect. The gender distinction is largely explained by the negative effects of college on women's likelihood to marry and have children.

Researchers studying intergenerational mobility almost invariably adopt a "retrospective" approach by obtaining data on a sample or population of adults and comparing their statuses with those of their parents. While this approach

¹ We thank Vida Maralani, Seongsoo Choi, Kristian Karlson, and the *AJS* reviewers for their helpful suggestions and advice. Earlier versions of this article were presented at the universities of Amsterdam, Trento, and Edinburgh; at New York University; at the Social Science Research Center (WZB), Berlin; Nuffield College, Oxford; the Danish Institute for Social Research, Copenhagen; and at the 2013 meeting of Research Committee 28 held in Trento. We

has made formidable contributions to our understanding of social mobility (see Breen and Jonsson [2005] and Hout [2015] for summaries of recent findings) and economic mobility (see Black and Devereux 2011; Torche 2015), it actually provides little insight into the question of whether a previous generation reproduces its advantages and disadvantages. As a result, there has always been a certain tension at the heart of conventional approaches: retrospective studies of intergenerational mobility are not really studies of mobility across generations.

In this article, we introduce a “prospective” approach that focuses on the chances of social reproduction for a preceding generation rather than on the chances of social mobility in the generation that succeeds it. Our method builds upon Duncan’s (1966) insight that the parents of respondents in a retrospective sample are not representative of any earlier generation. As Duncan noted, only the respondents (the adult children) in retrospective studies can claim to represent a generation since not all individuals in the prior generation become parents, while those who do may have children at different ages.

Sociologists and demographers have recently renewed interest in Duncan’s observation and the importance of identifying the mechanisms of family formation as the mediators of intergenerational transmission. For example, Mare and Maralani (2006), Mare (2011), Maralani (2013), and Song and Mare (2015), have all emphasized the importance of bringing the determinants of family background to the foreground of studies of intergenerational inequalities. These articles primarily account for how changes in assortative marriage and differential fertility affect the distributions of children and the distributions of children’s statuses in a later generation. Our prospective approach differs by prioritizing the pathways that determine whether that later generation is even created. After all, as Maralani (2013) noted, another generation must exist before a family status can be transmitted.

We use the example of educational reproduction to ask, What is the effect of graduating from college on having a child who also graduates from college? Using data from the Wisconsin Longitudinal Study (WLS), we consider the joint outcomes of marriage, fertility, and the intergenerational transmission of status as potential returns to higher education. For all the cohort members in our sample, we analyze how acquiring a college degree influences whether someone marries, whether someone marries a college-educated spouse, how many children someone has, and, if they have at least one child, whether one or more of these children graduates from college and so reproduces the cohort member’s educational status.

are grateful to participants on all those occasions for their comments. Direct correspondence to Matthew Lawrence, Department of Sociology, Middlebury College, Munroe Hall, 427 College Street, Middlebury, Vermont 05753. E-mail: lawrence@middlebury.edu

Our study is further distinguished from previous descriptive research by our interest in estimating causal effects. We do this by using marginal structural models with inverse probability of treatment weighting (see Robins [1999], Hernán, Brumback, and Robins [2001], Sharkey and Elwert [2011], and Wodtke, Harding, and Elwert [2011] for recent sociological applications). We also draw on work by VanderWeele (2009) that allows us to estimate both the direct effect of a parent completing college on the probability of educational reproduction, and the indirect effect on that outcome that is mediated by the effect of schooling on whom a parent marries and how many children he or she has.

We present the results in two parts. In the first set of analyses, we consider the causal effects of college on the probabilities of having at least one child and having at least one child who graduates from college. While these are minimal definitions of fertility and transmission, they clarify the potential contributions of our prospective approach: improving upon conventional analyses by including nonparents when estimating intergenerational effects and estimating the chance that a status is reproduced at all in a succeeding generation. Our second set of results expands these definitions to consider the causal effects of college on the number of children a graduate has and the proportion of those children who graduate from college.

The results from both approaches are consistent and reveal striking conclusions about gender inequalities in the intergenerational reproduction of educational status. While college has a positive effect on whether male graduates have a child who also completes college, among all the female WLS respondents in our sample there is no effect of graduating from college on having a child who graduates from college. In other words, from the point of view of a generation reproducing its advantages, women's college completion has no effect. We find that this is explained by the lower probability that female college graduates ever have any children and this, in turn, is due to their lower probability of marrying. Among women who do have children, college graduation increases the probability of reproducing that educational status, in part by increasing the probability of marrying a spouse who also graduated from college.

The article is structured as follows. In the next section, we discuss our prospective model, how it differs from the conventional retrospective model, and how it incorporates the role of intermediary mechanisms in the transmission of intergenerational effects. We then introduce our data and analytical approach. Many of the technical details of what we do can be found in appendix A. We discuss the results from our models in the subsequent section. We report sensitivity analyses that test the robustness of our causal claims to the assumptions we make, and we conclude by considering possible developments of the prospective approach. Additional technical details of the robustness tests are in appendix B.

PROSPECTIVE MODELS OF EDUCATIONAL REPRODUCTION

The question of why the children of parents with more schooling tend to complete more schooling themselves has been central to the research agendas of stratification scholars (such as Blau, Duncan, and Tyree 1967; Jencks et al. 1972; Featherman and Hauser 1978), cultural reproduction theorists (Bourdieu 1977; MacLeod 1989; Lareau 2003), and economists (Bowles and Gintis 1976; Becker 1991; Bowles, Gintis, and Groves 2005). Despite their breadth of theories and methods, these studies share the conventional approach of retrospectively comparing children's statuses to those of their parents. But the relationship between schooling in one generation and educational outcomes in the next may be better understood by asking a prospective question: Does completing more schooling increase the probability of eventually having a child who will also complete more schooling?

Prospective and retrospective approaches differ because the conventionally studied relationship between parents' and child's status is the last step in a long sequence of social and biological processes. As Duncan (1966) showed, the demographic mechanisms that link generations mean that the distribution of origins is not representative of the distribution of positions in any earlier generations. A retrospective study's sample of respondents' parents cannot represent a previous generation for multiple reasons: it ignores individuals who never had children, fails to capture individuals who have children outside the study window, and may oversample parents who have multiple children in the sample. Though sociologists frequently write as though their retrospective data are informative about generational reproduction, Duncan's (1966) insight was that their data are informative about the relationship within one generation between destinations and origins. Therefore, using retrospective data to examine inequalities across generations can easily be misleading.

Care is required—but not always taken—in the interpretation of conventional mobility studies. Suppose, for example, we were to observe a decline in the intergenerational correlation between parental and child education. This could reflect a genuine weakening in the relationship or it may arise because of a decline in the fertility of people whose children would have the highest probability of having the same educational attainment as their parents (most likely people at the extremes of the educational distribution). As Mare and Maralani (2006), Mare (2011), Maralani (2013), and Song and Mare (2015), have shown, relying on associations between observed parent-child pairs to estimate intergenerational effects can lead to incomplete accounts if they ignore the effects that a parent's status has on the demographic mechanisms of assortative marriage and differential fertility. Changing a mother's schooling level may influence whom she marries and how many children she has; thus, using a conventional retrospective study to estimate how a child's education changes as his or her parents' education changes reveals little about

how changes in the educational attainments of a generation affect the educational distribution in the succeeding generation.

Recent studies that seek to examine questions of educational reproduction and that acknowledge the intervening role of family formation still share two critical limitations with conventional analyses. The first limitation is reflected in the work of Attewell and Lavin (2007), who show how open access policies enacted by the City University of New York in the 1970s increased educational opportunities for women and, in doing so, also influenced the types of husbands they could marry. Despite recognizing the advantages that educational homogamy created for the children of these new college graduates, Attewell and Lavin nevertheless interpret mothers' and fathers' schooling as having independent effects. However, omitting accurate measures of the pathways that connect parents to each other may lead to biased estimates of the total effect of a "treatment" in one generation on an "outcome" in the next (Sharkey and Elwert 2011).

Second, most researchers attuned to the significance of demographic pathways have conceptualized the importance of education for intergenerational reproduction only for those who eventually become parents. Estimating the effects of education among this restricted group is a classic example of sample selection bias, or, in Greenland, Pearl, and Robins's (1999) phrase, conditioning on a collider. Song and Mare (2015) discuss ways to "correct" for these issues using retrospective data. Maralani (2013) acknowledges similar concerns by demonstrating how the distribution of education in one generation depends on the relationship in the previous generation between educational attainment and becoming a parent (and with whom) in the first place. The simulations in that article allowing women to forgo marriage and childbirth reveal why transmission across generations is both a question of "where families come from" (Mare 2011) and of whether they ever come to be.

Our prospective model extends this view by focusing not on the chances of intergenerational educational mobility in the succeeding generation, as this is conventionally thought of, but on the chances of educational reproduction for the generation that preceded it. We elucidate the causal mechanisms linking schooling across generations by analyzing how acquiring a college degree influences whether a graduate marries at all, marries a college-educated spouse, and has children. It is to the development of the intervening demographic pathways that we now turn.

MARRIAGE AS AN INTERVENING MECHANISM

Marriage is a central feature of household composition and family formation with clear ties to educational outcomes. For example, children who grow up in families with single mothers complete less schooling than children who grow up in households with two parents (McLanahan and Sandefur 1994).

The growing association over time between educational attainment and marriage patterns has been particularly well documented. In their review of trends in assortative mating in the United States, Schwartz and Mare (2005, pp. 623–24) discuss three reasons why rates of educational homogamy have increased: (1) the proportion of Americans extending their educational careers has grown, strengthening the role of schools (rather than workplaces, for example) as places to meet potential partners; (2) a convergence in gender roles has meant that men and women increasingly search for similar characteristics in partners; and (3) the cultural distances and social segregation associated with growing income inequality have limited opportunities for individuals of different socioeconomic backgrounds to meet. Decreases in marriage rates across schooling levels may thus closely relate to questions of inequalities in the reproduction of educational statuses if the benefits of parental educational attainment for intergenerational transmission are concentrated among increasingly segregated segments of the population.

The importance of marriage to intergenerational inequality is often incompletely acknowledged when estimating the effect of family background on educational reproduction. Consider the model shown in figure 1A, which reflects the conventional approach to estimating how parents' college completion affects whether an individual child also completes college. The critical limitation in this design is that it only allows for each parent to have an independent effect on that child's outcome (ϕ_1 and ϕ_2). Because a parent's schooling may also have influenced the likelihood of marrying a spouse with a particular level of schooling, it is inaccurate to think of the total effect of parents' education as the sum of $\phi_1 + \phi_2$ (Sharkey and Elwert 2011; Wodtke et al. 2011). Accounting for the influence of parental education on parental marriage requires measuring a new effect, labeled ψ_1 in figure 1B, which is the effect that one parent's education had on whether that parent would marry a spouse with a certain educational attainment. By estimating separate models for each parent, the total effect of a mother's education, for example, would decompose into a sum of her direct effect (ϕ_1) and the new indirect effect ($\psi_1\phi_2$) that includes how likely she was to marry a college-educated husband. Including this indirect effect in the estimation of the total effect of each parent's education will lead to more accurate conclusions about how children's outcomes are influenced by their parents' schooling. But figure 1B, like the model in figure 1A, only includes individuals who have children and their offspring.

FERTILITY AS AN INTERVENING MECHANISM

The importance of fertility to educational reproduction is straightforward: an educational status can only be transmitted if there is a child to whom it can be passed. The association between schooling and fertility is much broader than that, however. There tends to be an inverse relationship

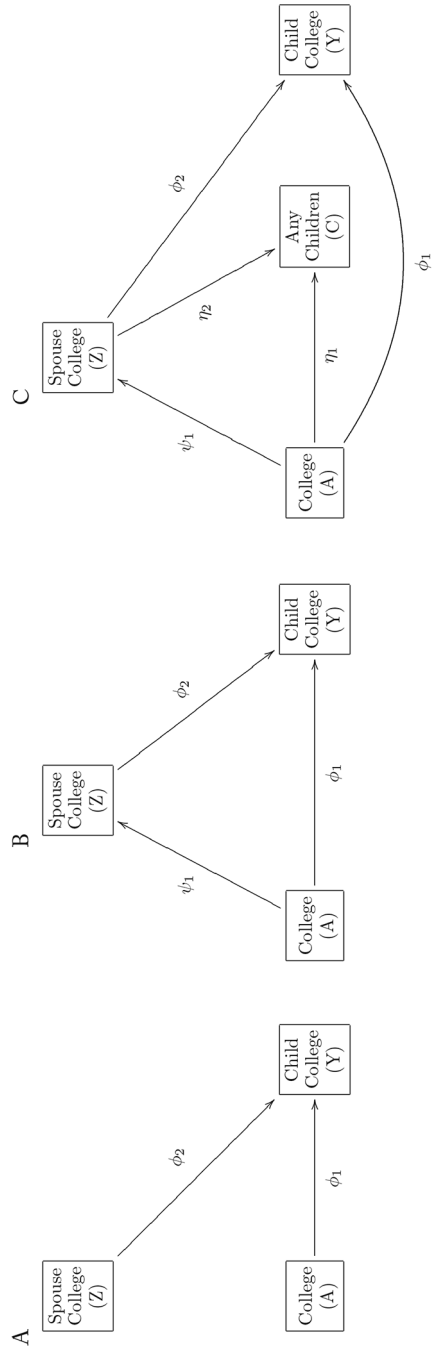


FIG. 1.—A, conventional model of educational reproduction; B, conventional model mediated by parents' marriage; C, prospective model mediated by marriage and fertility.

between education and fertility, with lower-educated women on average having more children (Mare 1996). One recent analysis of the effect of college on fertility (Brand and Davis 2011) also suggests that completing college decreases fertility primarily for those women least likely to attend college, perhaps because they are most attracted to higher education's potential economic (rather than social) rewards. These associations between education and fertility can in turn directly influence the family context in which a child grows up, affecting factors such as sibship size and/or access to financial or social resources that are closely linked to educational outcomes (Becker 1991; Conley 2001; Conley and Glauber 2006; Mare and Maralani 2006).

As the studies above suggest, most research on the effect of schooling on fertility focuses on the different numbers of children that mothers with different levels of education tend to have. Yet as Maralani (2013) discusses, a complete examination of educational reproduction should first consider whether schooling affects who has any children. Figure 1C shows a prospective model of intergenerational transmission that integrates the effect of schooling on fertility by adding the variable C , which captures whether an individual has any children.² Introducing the direct effect of college on becoming a parent (η_1) allows for a more accurate conceptualization of the total effect of college on educational reproduction by including all possible parents of all possible children in the next generation.

In the analyses that follow, we show how prospective models can be used to jointly estimate the direct and indirect effects that link educational outcomes across generations. We estimate separate models for men and women and show how the influence of college completion on the demographic pathways of family formation lead to different conclusions about gender inequalities in educational reproduction.

DATA

Conventional retrospective studies of intergenerational transmission start with a sample of children and add information about their parents. In contrast, the prospective approach we introduce starts with a sample of potential parents and, for those who become parents, adds information about their children. Our analyses also require data about the families of origin of the sampled individuals in order to calculate their probabilities of having children; this, in turn, requires calculating their probabilities of completing college, of ever marrying, and of marrying a spouse who completes college. To our knowledge, the only data to offer these details come from the WLS, a sample of 10,317 men and women who graduated from high schools in Wisconsin in 1957 (WLS Hand-

² Prospective models, in this sense, consider fertility as part of the model, whereas retrospective models condition on the presence of at least one child.

book 2006).³ We refer to the high school graduates in the original WLS cohort as the “respondents.” All other family members are identified by their relationships to members of that initial cohort: the spouses are the husbands and wives of those WLS respondents who ever married, and the children are the biological sons and daughters of those WLS respondents who ever had any children.

At the time of the initial 1957 survey, WLS respondents provided detailed information about the households they lived in during their final year of high school. Data were also collected about the respondents’ academic achievements, postsecondary plans, and whether their parents and teachers encouraged them to attend college. Follow-up surveys in 1964, 1975, 1992, and 2004 provided additional data about marital histories, fertility, and the educational attainments of respondents’ spouses and children. We use data from all of these surveys.

Information about the households in which respondents grew up comes from the original 1957 survey. We restrict our sample to respondents who reported being continuously married or continuously never married in the 1975 and the 1992 surveys. This restriction reduces the potential for our results to be influenced by divorce, remarriage, and widowhood. Information about spouses’ college completion was gathered starting with the 1975 survey, when the question was first asked. We include the last measure reported in any WLS survey in which a respondent reported their spouse’s education.

Data about the respondents’ total number of children (of any age) come from the 1992 survey and are updated with data from the 2004 survey when available. Respondents with five or more children are collapsed into one category. Because the WLS does not assign unique identifiers to children that are consistent across survey years, we are unable to update the educational histories of individual children if and when more recent schooling information was reported. To include the greatest number of observations, the effects we estimate are based on the educational histories collected in 1992. Information on whether children graduated from a four-year college is taken from respondents’ reports of the highest level of schooling completed by each of their children, but we only include those biological children who were at least 25 years of age at the time of the most recent report of the years of schooling they completed.

We have complete information for 2,535 male WLS respondents, of whom 2,252 became fathers and 1,499 had at least one child who graduated from college. We have complete information for 2,745 female WLS respondents, of whom 2,455 became mothers and 1,585 had at least one child who

³ Linking data from the National Longitudinal Survey of Youth (NLSY) and its Young Adult Survey would also provide data on three generations but would only include information about the educational attainments of female NLSY respondents’ children. Furthermore, these children are mostly too young to have completed their education by the time of their 2010 interview.

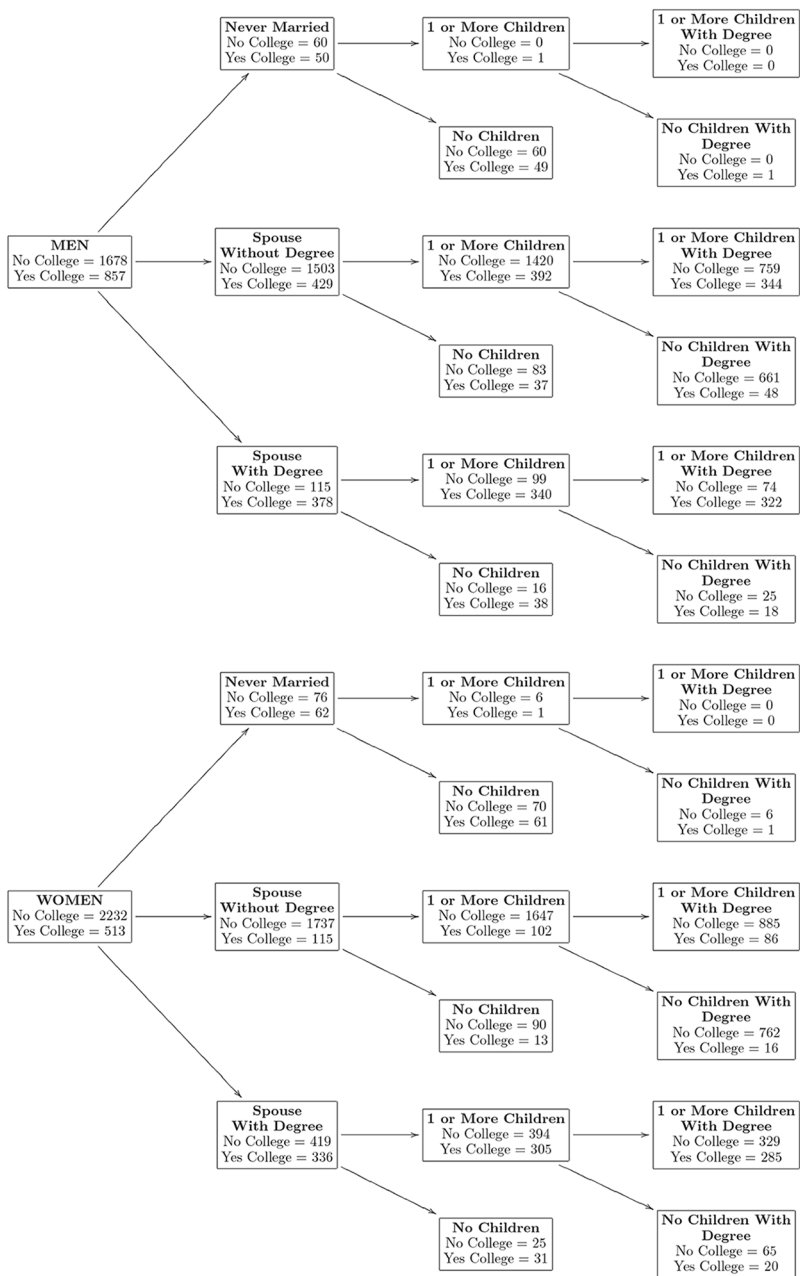


FIG. 2.—Summary of outcomes for WLS respondents by college completion status

graduated from college. Figure 2 shows how the respondents are associated with each of the outcomes we consider.

Comparing the paths suggests some important differences between men and women overall and within each group by college degree status. For example, while 3% of all men and women who did not complete college never married, larger proportions of college graduates—6% of male graduates and 12% of female graduates—never married. Further, men who did not graduate from college were very unlikely to “marry up.” Only 7% of these men married a college-educated spouse, while 19% of women in that lower educational category still married husbands with college degrees. Male and female college graduates were both more likely to marry spouses with college degrees, but a wide gap opened between them on this measure: 44% of male college graduates’ spouses were also college graduates, while this pattern of educational assortative mating describes 66% of all female college graduates.

With regard to children, it is striking that nearly identical proportions of men with and without college degrees had at least one biological child (86% and 91%, respectively). A comparable proportion of women without college degrees had children as well (92%). College-educated women are the only group with a different rate on this measure, with one-fifth of them reporting no biological children by the last survey in which they share their fertility histories.

We can calculate simple ratios of educational transmission from these frequencies. In this sample, there are 129 male college graduates for every 100 with a child who also earns a college degree. In contrast, many more men without college degrees—201—are required for 100 men in that category to eventually have a child complete college. For women, there are 138 female college graduates for every 100 who have a child with that same level of education, while there are 184 females with only a high school diploma for every 100 who have a college-educated child.

METHODS

Calculating Causal Effects from Marginal Structural Models

Do the effects of college on the pathways of family formation provide causal explanations for the observed differences in outcomes? Under the counterfactual definition of causality, the causal effect of college on educational reproduction would be easy to calculate if we observed each WLS respondent completing and not completing college. Of course, each respondent can only belong to one of those categories; the “fundamental problem of causal inference” (Holland 1986) arises because we cannot observe respondents in their counterfactual state, and so we only observe one of their two “potential outcomes.”

Our modeling strategy for dealing with this problem comes in two parts. First, we use inverse probability of treatment weighting, IPTW, to estimate

causal effects under the assumption that potential outcomes are independent of treatment, conditional on observed covariates. This is called the “conditional independence assumption” or the “no unmeasured confounders” condition. Second, we carry out sensitivity tests on the robustness of our causal claims to different degrees and types of unmeasured confounding. We explain our sensitivity analyses below; here we focus on the way we estimate causal effects under the conditional independence assumption.

Given selection into treatment (in our case, completion of college) on observable covariates, methods such as regression or propensity score matching might be used to try to estimate unbiased effects of treatment on an outcome (in our case, having a child who graduates from college). In recent years reweighting has become a popular alternative to these approaches; see Robins (1999), Hernán and Robins (2006), and Cole and Hernán (2008) for some illustrations. The idea is to make the distribution of potential outcomes independent of whether an observation is found in the treatment or control group by reweighting the observations by the inverse of their probability of being found in the category (treatment or control) that they were observed in.

Using the treatment and control groups as defined in this article, WLS respondents who graduated college would be weighted by one divided by the probability that they would have graduated college, and those who did not graduate college would be weighted by one divided by the probability they would not have graduated from college.⁴ The probabilities in question are the individuals’ own probabilities, which depend on the individual respondents’ own values on the confounders. In other words, the probability is the propensity score for a respondent in the treatment group, and one minus the propensity score for a respondent in the control group. In the weighted data, the average of the observed outcomes for the treated is an unbiased estimate of the average of the potential outcome, given treatment. Likewise, the average of the observed values for the untreated is an unbiased estimate of the average potential outcomes, given nontreatment (see Angrist and Pischke 2009, pp. 82–83).

As an example, consider the causal effect of graduating from college on the probability of having a child. We write C^A to indicate the potential outcomes of C (whether a WLS respondent had a child [$C = 1$] or not [$C = 0$]) depending on the value of A , where $C^{A=1}$ means the potential outcome of C had the respondent graduated from college and $C^{A=0}$ means the potential outcome had she not graduated. We can then write a marginal structural model (Robins 1999; Hernán et al. 2001; Sharkey and Elwert 2011; Wodtke et al. 2011) as follows:

⁴ A numerator of one and a denominator of the probability of received treatment is the most straightforward definition of an inverse weight. However, in practice it is common to “stabilize” the weights by adjusting the numerator to be the proportion of cases in the sample with the given treatment assignment. As described later, the weights we use are stabilized with these adjusted numerators.

$$E(C^A) = \eta_0 + \eta_1 A. \quad (1)$$

Equation (1) estimates what we will call the “total” causal effect of college on having a child and is estimated using data that have been weighted according to the procedure just outlined. If, as here, we use a linear probability model, our estimate of η_0 is an unbiased estimate of $E(C^{A=0})$ —that is, the average potential outcome under the no treatment (non-college graduation) condition; and $\eta_0 + \eta_1$ is an unbiased estimate of $E(C^{A=1})$, the average potential outcome under treatment (college graduation). Thus, the average causal effect of college graduation on having a child is estimated by η_1 . These estimates are unbiased given not only our maintained assumption of conditional independence but also the assumption that equation (1) has been properly specified⁵ and that the weights have been estimated correctly (i.e., the propensity score should be correctly specified).⁶

The equation for the potential outcomes, Y^A , could be written in the same way,

$$E(Y^A) = \lambda_0 + \lambda_1 A, \quad (2)$$

and estimated using weighted data. When this equation is fitted to the entire sample of WLS respondents, including those who had no children, the resulting estimate of λ_1 is the outcome of two underlying causal processes: how college graduation affects the chances of having at least one child, and, given that someone had one or more children (actually or counterfactually), how college graduation affects the probability of reproducing that educational attainment in the succeeding generation.

Our aim is to estimate the parameters of these underlying causal processes and show how they combine to give rise to what we call the realized causal effect of graduating college on having a child who graduates college. To do that, we first specify and estimate a marginal structural model for whether or not a WLS respondent has at least one child, and then we specify and estimate a model for whether or not at least one child graduates college.⁷ In

⁵ Equation (1) posits a constant treatment effect that does not vary according to measured characteristics. We do, however, allow treatment effects to vary by gender because we fit separate models to male and female WLS respondents. In fact, the design of the WLS requires us to fit these separate models. Since we have far more information about the educational and family histories of the WLS respondents than we have about their spouses, we can only calculate the counterfactual outcomes for the respondents. See Warren and Hauser (1997) for a discussion of this structural feature of the WLS.

⁶ This does not relate to whether there are any variables omitted from the propensity score, because, thus far, we are assuming that there are none. Rather it concerns whether the effects of the predictor variables on the propensity score have been specified correctly.

⁷ Our models estimating the effects of college on having a certain number of children and on the proportion of them who graduate from college follow the same general approach. We describe how the variables differ in a later section.

each case we want to estimate all the pathways linking causes with effects. Fortunately, as shown by VanderWeele (2009), the marginal structural model can be extended to encompass the situation in which we want to estimate direct and indirect causal effects, as well as total effects (for details, see app. A).

We define A to be a binary variable for the WLS respondent's education (coded 1 if he or she received a four-year college degree and 0 otherwise). The variable Z has three categories, indicating whether the WLS respondent did not marry ($Z = 0$), married a non-college-educated spouse ($Z = 1$), or married a college-educated spouse ($Z = 2$). The variable C indicates whether or not the WLS respondent had any biological children (coded 1 if he or she has at least one biological child and 0 otherwise). The variable Y indicates whether or not a WLS respondent has at least one biological child above the age of 25 who graduated from a four-year college (coded 1 if yes and 0 otherwise).⁸

The model for having a child has three marginal structural equations:

$$E(Z0^A) = \alpha_0 + \alpha_1 A, \quad (3)$$

$$E(Z2^A) = \gamma_0 + \gamma_1 A, \quad (4)$$

$$E(C^{A,Z0,Z2}) = \beta_0 + \beta_1 A + \beta_2 Z0 + \beta_3 Z2. \quad (5)$$

Equations (3) and (4) are marginal structural models for the potential outcomes for $Z0$ (not marrying) and $Z2$ (marrying a college-educated spouse). Because $Z1$ acts as the omitted category, the resulting causal effects from equation (3) are interpreted as a respondent's increase in the probability of not marrying, rather than marrying a non-college-educated spouse, because of the respondent's graduating from college (rather than not graduating). Equation (4) should be interpreted similarly, except that now the outcome is marrying a college graduate rather than marrying a nongraduate. Equation (5) is the marginal structural model for a respondent having a child. In this case there are six potential outcomes given by all possible combinations of A and the Z variables. Equations (3), (4), and (5) allow us to estimate the direct causal effects of A , $Z0$, and $Z2$ on C and the indirect effect of A on C .

The marginal structural model for Y , having a child who graduates college, is given by

$$E(Y^{A,Z0,Z2}) = \theta_0 + \theta_1 A + \theta_2 Z0 + \theta_3 Z2. \quad (6)$$

In contrast to equation (2), equation (6) is estimated using only WLS respondents who had a child. Nevertheless, given our assumptions, it provides esti-

⁸ In our second set of analyses, we revise two of the variables. We define C as the number of children each respondent has (from 0 to 5 or more) and we define Y as the proportion of each respondent's biological children above the age of 25 who graduated from a four-year college. We follow Friedman and Mare (2014) in estimating the parental returns to children's schooling using the proportion of children who graduate from college rather than the number of children who do so.

mates of the direct causal effects of A , $Z0$, and $Z2$ on Y that apply to all WLS respondents, not only to those who had children. For those who did not have children they tell us what the effects of A , $Z0$, and $Z2$ on Y would have been in the counterfactual state in which they had children. Appendix A details the construction of the weights used in the estimations and explains how the causal effects are derived from the parameters of the marginal structural models.

ANALYSIS AND FINDINGS

In the first set of results, we use our prospective model to estimate the effect of college on the probability of having a child who also graduates from college. We do so by including all the WLS respondents in our sample—those who have children and those who do not—to estimate equation (2). The results are striking: for men, graduating college increases by 22 percentage points their probability of having a child who also graduates college (the coefficient is estimated as 0.218, $SE = .039$), but among women, college has a small, nonsignificant effect (0.027, $SE = .060$). To understand how this difference arises, we need to turn to the underlying causal processes.

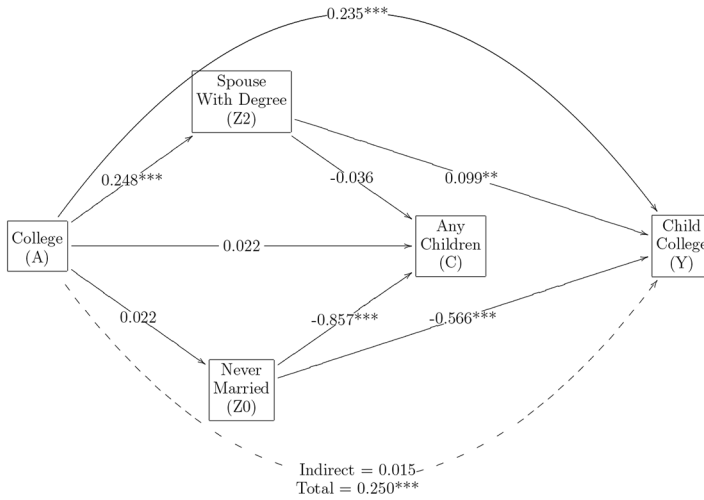
Figure 3 shows these processes in the form of a directed acyclic graph. The figure should be interpreted in the same way as a normal path diagram, except that it displays causal estimates rather than regression coefficients.⁹ This means that the direct path from A to $Z0$, for example, tells us the direct causal effect of graduating college on not marrying. The direct path from A to Y tells us the direct causal effect of college graduation on having a child who graduates from college. This, like all the other causal effects shown in the figure, applies to the whole sample of WLS respondents and not, in this case, only to those who actually had children. This follows because the causal effects are defined as the differences between potential, not observed, outcomes. The estimates of the causal effects of A derived from the marginal structural models, 3–6, are reported in columns 1–4 of table 1 (we discuss the realized causal effects reported in col. 5 later). The table also reports which effects differ between men and women and which do not.

The direct effect of college on a child's graduating college (the direct effect of A on Y) for those who actually or counterfactually have children is estimated as .235 for men and .228 for women; these are not significantly different (at $P < .05$). Neither the indirect nor the total effects of A on Y differ statistically significantly between men and women either. Were a WLS respondent to have any children, the positive causal effect of college graduation on having a child graduate from college is very similar for the two sexes. For both men

⁹ And we have also added an extra path, indicated by the dashed line, that summarizes the indirect and total effects of A on Y .

Educational Reproduction

Men



Women

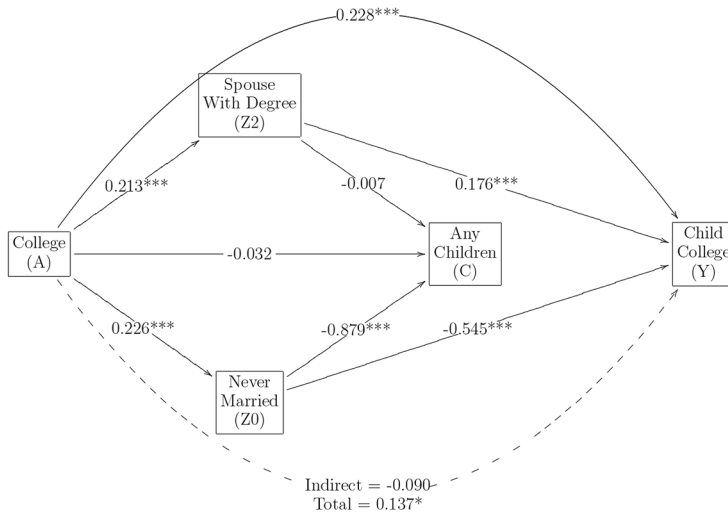


FIG. 3.—Causal effects of graduating from college on marriage, having at least one child, and having at least one child who completes college (for all respondents in sample). * $P < .05$; ** $P < .01$; *** $P < .001$.

and women, marrying a college-educated spouse increases this positive effect (the estimate from the marginal structural model of the causal path from Z2 to Y is 0.176 for women and 0.099 for men).

But alongside these similarities, there is a very large—and consequential—gender difference in the total effect of college on having one or more children. College reduces women's chances of having a child by 23 percentage

TABLE 1
CAUSAL EFFECTS OF COLLEGE (A) ON NEVER MARRYING (Z0), MARRYING A SPOUSE WITH A COLLEGE DEGREE (Z2),
HAVING AT LEAST ONE CHILD (C), AND HAVING AT LEAST ONE CHILD WHO COMPLETES COLLEGE (Y)

| | MEN | | | | | WOMEN | | | | |
|--------------------|--------------|-----------------|---------------|-----------------|-------------------|-----------------|-----------------|------------------|-----------------|-------------------|
| | Z0 (1) | Z2 (2) | C (3) | Y (4) | Realized Y (5) | Z0 (1) | Z2 (2) | C (3) | Y (4) | Realized Y (5) |
| Direct | .02 (.02) | .25*** (.03) | .02 (.02) | .24*** (.04) | .22*** (.04) | .23*** (.06) | .21*** (.05) | -.03 (.03) | .23*** (.04) | .14* (.04) |
| Indirect | | | -.03 (.02) | .02 (.02) | .00 (.03) | | | -.20*** (.06) | -.09 (.05) | -.10*** (.06) |
| Total | | | -.01 (.03) | .25*** (.04) | .22*** (.04) | | | -.23*** (.06) | .14* (.06) | -.05 (.08) |

NOTE.—Nos. in parentheses are SEs. The difference between men and women in the total effect of A on C is significant at $P < .01$. The differences between men and women in the effect of A on Z0, the indirect effect of A on C, the realized indirect effect on A on Y, and the realized total effect of A on Y are significant at $P < .05$.

* $P < .05$.
 ** $P < .01$.
 *** $P < .001$.

points but has no effect on men's chances. This difference stems from one prior behavioral difference: for women, graduating college increases the probability of not marrying by 23 percentage points, whereas there is no effect at all for men.¹⁰ Although being unmarried reduces fertility equally for men and women (reducing their probability of having a child by 86 and 88 percentage points, respectively, as shown in fig. 3), the differential effects of college on ever marrying mean that college also has very different total causal effects on the chances that male and female graduates will have at least one child.

Realized Causal Effects

In column 5 of table 1 we report what we term the direct, indirect, and total "realized" causal effects of A on Y . The realized causal effects are the products of the probability of having a child and the probability that a child graduates college. They are calculated using the parameters of the marginal structural models for C and Y (all of which are reported in app. table A4). We define a new set of potential outcomes as $E([YC]^{A,Z_0,Z_2})$ and, using them, a new total causal effect:

$$E([YC]^{A=1,Z_0(A=1),Z_2(A=1)}) - E([YC]^{A=0,Z_0(A=0),Z_2(A=0)}).$$

This is the potential outcome for Y when $A = 1$ times the potential outcome for C when $A = 1$, minus the potential outcome for Y when $A = 0$ times the potential outcome for C when $A = 0$. This can be expanded to

$$\begin{aligned} &E(Y^{A=1,Z_0(A=1),Z_2(A=1)})E(C^{A=1,Z_0(A=1),Z_2(A=1)}) \\ &- E(Y^{A=0,Z_0(A=0),Z_2(A=0)})E(C^{A=0,Z_0(A=0),Z_2(A=0)}). \end{aligned} \quad (7)$$

As a final step, we can substitute the parameter estimates from the marginal structural models into equation (7), as shown in appendix A, to yield realized total, direct, and indirect effects. The realized total causal effects should be approximately equal to the coefficients estimated from equation (2).

Comparing the realized total causal effects of men and women shown in column 5 of table 1 helps explain why we found a positive effect of college for

¹⁰ The most plausible explanation for this is the norm of hypergamy, which reduced the availability of acceptable marriage partners for highly educated women. It is important to note that the preferences of men and women could have jointly affected the marriage market for college graduates. As suggested by the outcomes in fig. 2, female college graduates may have preferred not to marry rather than marry a husband without a college degree. Male college graduates, in contrast, may have preferred to marry a spouse without a college degree rather than not marry, a preference that would have further reduced the availability of highly educated potential husbands for highly educated women.

men and a negative effect for women when we estimated equation (2). The contrast can now be seen to be a consequence of the indirect effects of college that act through marriage, rather than of differences to the returns to college for WLS respondents who had a child. For men, the estimated total realized effect of A on Y , using equation (7), is 22 percentage points, the same as from equation (2). And because graduating college has no significant impact on fertility, the total realized effect of college on having a child who graduates college is overwhelmingly made up of the direct effect. For women, we estimate a positive direct effect of college of .14. However, this effect is more than offset by the negative indirect effect of $-.19$, which captures the effect of college graduation on reducing a woman's likelihood of marrying and the effects that this has on having any children. For women, then, the total effect from equation (7) is $-.05$, compared with .03 from equation (2); this difference is not statistically significant.

As mentioned previously, having at least one child who graduates from college is a minimal definition of educational reproduction. The same methods used to estimate the effects of college on that outcome can be adapted to estimate the effects of college on the probabilities of having a specific number of children (by revising C) and on the proportion of those children who will graduate from college (by revising Y). To do so, we expand C from a binary variable to a variable with six categories: no children, one child, two children, three children, four children, and five or more children. For each value of C , we use a separate marginal structural model (based on eq. [5]) to estimate the effects of A , Z_0 , and Z_2 on having that number of children. The probability of each outcome for C is estimated from a multinomial logit model, and these probabilities are used to construct a revised set of weights for a new marginal structural model where Y is the proportion of biological children ages 25 years or older who have a college degree:

$$E(Y^{A,Z_0,Z_2}) = \delta_1 C_1 + \delta_2 C_2 + \delta_3 C_3 + \delta_4 C_4 + \delta_5 C_5 + \delta_6 A + \delta_7 Z_0 + \delta_8 Z_2. \quad (8)$$

The parameter estimates from this model yield realized direct, indirect, and total effects of A on our revised Y for every possible value of C following the steps outlined in equation (7).¹¹ Table 2 reports these realized direct, indirect, and total effects by number of children. For every value of C , the total effect of college on the proportion of children who graduate from college is higher for men. For men, these total effects are overwhelmingly made up of the direct effects since there are no significant indirect effects. The direct

¹¹ Appendix table A5 reports the parameters from this model.

TABLE 2
 REALIZED DIRECT, INDIRECT, AND TOTAL CAUSAL EFFECTS OF COLLEGE ON PROPORTION
 OF CHILDREN WHO COMPLETE COLLEGE, BY NUMBER OF CHILDREN

| | MEN | | | | | WOMEN | | | | |
|-------------------------------|---------------|-----------------|-----------------|-----------------|-----------------|--------------|---------------|-----------------|----------------|-----------------|
| | 1 | 2 | 3 | 4 | 5+ | 1 | 2 | 3 | 4 | 5+ |
| Direct | .04* (.02) | .16*** (.03) | .06** (.02) | .00 (.01) | .00 (.01) | .00 (.01) | .06 (.03) | .09 (.06) | -.01 (.02) | .00 (.03) |
| Indirect | .00 (.01) | .01 (.01) | .00 (.01) | .00 (.01) | .00 (.00) | .00 (.00) | -.02 (.02) | -.04** (.02) | -.02* (.01) | -.02** (.01) |
| Total | .03* (.02) | .16*** (.03) | .05** (.03) | .00 (.01) | .00 (.01) | .00 (.01) | .03 (.03) | .04 (.05) | -.03 (.02) | -.02 (.02) |
| Cumulative total | .03* (.02) | .20*** (.03) | .25*** (.03) | .25*** (.03) | .25*** (.04) | .00 (.01) | .04 (.04) | .08 (.07) | .05 (.07) | .04 (.07) |

NOTE.—Nos. in parentheses are SEs.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

effect of college graduation when men have two children is most striking: completing college increases the proportion of those children who complete college by .16.

For women, revising the measures of C and Y does not revise our finding that the indirect effects of college offset the direct effects. Even in the one scenario where the direct effect is stronger for women than for men (when individuals have three children), the indirect effect is also more strongly negative for women, resulting in a total effect that is slightly higher for men. In fact, no indirect effect is positive for women. This confirms that after broadening our definition of educational reproduction, the pathways linking A and Y through the effects of college on marriage remain consequential.

The sum of the realized total effects across the values of C can be interpreted as the overall realized effect of college completion on the proportion of children who graduate from college. For men, the cumulative effect is .25, meaning that completing college increases the proportion of male graduates' children who complete college by one-quarter. The cumulative total effect reaches this point at three children and remains stable as the number of children increases. For women, the cumulative total effect also peaks at three children (though even at that parity the total effect for women is roughly one-third that of men) and then drops to .04. These results suggest that the effects we found when using a minimal definition of educational reproduction were not obscuring a story about the proportion of children who graduate. In fact, the cumulative total effect of college on the proportion of all children who graduate for women is similar to the total effect when men only have one child.

THREATS TO VALIDITY

Representativeness of Our Sample

One potential limitation of our study is that our findings may not generalize beyond the populations represented by the WLS: midwestern Americans who completed high school in the middle of the twentieth century and their children. We cannot replicate our full study with more recent data (see n. 2); however, we did use data from the National Longitudinal Survey of Youth (NLSY 79) to examine whether the gender differences in the proportion of college graduates who have children is anomalous to our sample.

Following Brand and Davis (2011), we restricted our NLSY sample to those respondents who were 14–17 at the time of the initial interview and completed at least twelfth grade in the year they turned 19. As shown in table 3, smaller proportions of NLSY college graduates reported not having any children, though the difference is only significant for women. Among men, the difference between graduates and nongraduates is .022 in the NLSY sample compared to .050 in our WLS sample (from fig. 2); for women, the difference is .078 in the NLSY sample compared to .120 in our WLS sample. At the same time, higher proportions of men and women in every educational category report not having any children in the NLSY sample. While female college graduates in the WLS stood out for their relatively high rate of not having children, female nongraduates in the NLSY stand out for their relatively low rate of not having children. Thus, our WLS sample may be unique for the high proportion of its respondents who had any children, but the gender difference in the relationship between college completion and fertility that we found is also apparent in the later, and nationally representative, NLSY data.

TABLE 3
PROPORTIONS OF NLSY 79 RESPONDENTS HAVING NO CHILDREN
OR AT LEAST ONE CHILD BY COLLEGE DEGREE STATUS

| | No Children | 1+ Children | Total |
|-------------------------------|--|-------------|-------|
| Men: | | | |
| Without college degree . . . | .28 | .72 | 1,504 |
| With college degree | .31 | .69 | 408 |
| Total | .29 | .71 | 1,912 |
| Test of differences | $\chi^2 = .79$ $P > \chi^2 = .38$ | | |
| Women: | | | |
| Without college degree . . . | .21 | .79 | 1,609 |
| With college degree | .29 | .72 | 4,35 |
| Total | .22 | .78 | 2,044 |
| Test of differences | $\chi^2 = 12.03$ $P > \chi^2 = .00$ | | |

While one concern may be that the WLS sample is too old to reflect contemporary trends, another concern may be that the sample is too young for the children of respondents to have completed their schooling. Measuring children's schooling among those children age 25 or older by the last survey date could raise the possibility that our sample ignored the effect of delayed fertility among college graduates. However, our analysis of the data does not suggest that this is problematic. Of all the WLS respondents in our sample, only 18 men (including 10 college graduates) and three women (only one of whom completed college) had their first biological child outside of the window we constructed. We reestimated our models relaxing the age restriction to include these later parents, and the results were unchanged (results available from authors).

Robustness to Violation of Assumptions

The interpretation of our results as causal parameters rests on the assumptions we made earlier about the conditional independence of potential outcomes and treatments. The validity of this assumption depends on whether there is still some confounding of the treatment-outcome relationship even given the control variables we employed, or, in other words, whether there is some unmeasured bias in selection into treatment. This is not directly testable because the potential outcomes cannot be observed. What we can do is assess how robust our causal estimates are to different degrees of selection bias into treatment. For this purpose we use a method developed by Robins (1999; Brumback et al. 2004) and recently applied by Sharkey and Elwert (2011). We present the method and our application in appendix B and include the results related to the total realized effect of college on child's college (eq. [7]) in appendix table B1, and the total effect of college on having any children (A on C in col. 3 of table 1) in appendix table B2.

We consider several different kinds of selection bias, including positive selection into college graduation (as assumed by economists studying the returns to education, e.g., Heckman 1979; see also Winship and Mare 1992) and negative selection into college, as recently proposed by Brand and Xie (2010). For the total realized effect of A on Y , we find that, for men, if unmeasured selection were as large as the selection we have accounted for through our control variables, the causal effect of college would lie between .096 and .104, compared with our estimate of .220. If unmeasured selection were half as great as measured selection, the effect would lie between .158 and .162. So, for our finding of a positive realized causal effect of college graduation among men to be overturned, there would have to be more selection on unobservables than on observables, something that seems unlikely to be the case. For women, there is not a scenario where we find a significant positive effect of college graduation on having a child who also graduates college.

CONCLUSION

Researchers studying social mobility and intergenerational inequality have almost exclusively focused on retrospective designs, yet for many purposes prospective approaches will be more relevant. The causal models we used confirm the importance of the “new view of intergenerational effects” (Mare and Maralani 2006), emphasizing the need to incorporate the demographic mechanisms of family formation into analyses of social reproduction. This is particularly relevant for understanding how educational statuses are passed down across generations, since the effects of schooling influence marriage and fertility outcomes, especially among female college graduates. For the women in our sample, completing college significantly affected family formation processes. Women graduates were less likely to marry (but more likely to marry a spouse with a college degree) and less likely to have children. Hypergamy, together with norms that dissuaded unmarried people (both male and female) from having children, generated the demographic pathways of family formation that combined to yield our most striking findings: that graduating college had no effect on a woman’s probability of having a child who graduated college and had no effect on the proportion of a woman’s children who graduated college.

For men, however, the effects of college on educational reproduction are unambiguously positive. Graduating from college had little effect on their chances of ever marrying or having any children, thus avoiding the negative effects that college had on both outcomes for women. The consistently strong and positive effect of college on educational reproduction for men also highlights how men with college degrees and men whose highest level of education was high school differ in their chances of having a college-educated child. Further, among those men and women who did not graduate from college, men were much more disadvantaged in terms of the statuses they could pass to their children. In this respect, the norm of hypergamy may have benefited non-college-educated women while disadvantaging their college-educated sisters.

Women college graduates who did not have children may, nevertheless, pass on their advantages to the next generation by their influence on the educational attainment of nieces and nephews. Jaeger (2012, p. 915), using WLS data, reports that aunt’s education has a significant positive effect on the educational attainment of the children of WLS respondents, even when controlling for parents’ and grandparents’ education and for the income and socioeconomic status of fathers, uncles, and grandfathers.¹²

¹² Jaeger finds no corresponding significant effect of uncle’s education. Our results suggest that one reason for this might be the greater likelihood for college-educated uncles, compared with college-educated aunts, to have children of their own in whom to invest.

The historic expansion of women's enrollment in higher education in recent years (see Buchmann and DiPrete 2006) suggests the need to pay attention not only to the educational returns for students but also to how demographic processes shape the ways in which changes in the distribution of education influence family formation. Understanding these changes within the context of contemporary American society is especially important. The effects we estimated in our overwhelmingly white sample of high school graduates may be different in samples with broader educational distributions and more racial diversity; indeed, Maralani (2013) showed that the effects of schooling on fertility are stronger for black women than they are for white women. Extending our methods to samples that better reflect the diversity of family backgrounds and types—in terms of race and immigration as well as changing patterns of divorce, remarriage, and step-parenthood—will be important for the development of prospective approaches.

Many applications of the methods used in this article suggest themselves. For example, the consequences of class differentials in fertility for social mobility are ignored in conventional, retrospective analyses but are easily incorporated into the prospective approach. The same is true of income differentials in parenthood and fertility for studies of intergenerational income mobility. Given the appropriate data, the prospective design could be fruitfully reformulated to analyze the association between the same characteristic or trait among parents and their children in numerous other areas of interest to social scientists (such as political preferences, health behaviors, or religious beliefs to name just a few).

By acknowledging that inequalities in social reproduction are influenced by how—and if—families are formed, the prospective approaches described in this article can continue to improve our knowledge of how statuses are linked across generations.

APPENDIX A

Technical Details

Marginal Structural Models

We use a set of marginal structural models to calculate causal effects. These are equations (3), (4), (5), and (6).

The coefficients of the marginal structural models can be estimated by replacing the potential outcomes in each of the above equations with their observed counterparts, after weighting each respondent by the inverse probability of receiving the treatment he or she actually received. We estimate the propensities to receive each treatment from logistic regression models that include pretreatment covariates commonly used to predict educational at-

tainment and the returns to college (see Brand and Xie [2010] and Brand and Davis [2011] for two recent examples). We describe the covariates and present the coefficients from the estimation models for men in appendix table A1 and for women in appendix table A2. We use the same pretreatment covariates for men and women since research suggests that family background, indicators of academic achievement, and encouragement from parents and teachers are important predictors of college completion for both sexes. However, some researchers have suggested that separate sets of covariates should be used when examining gender differences in the effect of college completion on certain outcomes (especially those related to occupational status). See Marini (1980) for an overview of this debate.

The propensity scores are used to construct the weights. In equations (3) and (4), where $E(Z0^A)$ and $E(Z2^A)$ are the dependent variables, we construct a weight (ω_A) that will make the potential outcomes of the Z s independent of the treatment, A . The weight we use here is computed as

$$\omega_A = \frac{\text{pr}(A = a)}{\text{pr}(A = a|X)},$$

where X indicates the relevant confounders. In our application, the confounders, X , are used in the calculation of the weights but do not appear in the marginal structural models as covariates. The numerator of ω_A is the marginal probability of the observation taking the value of A that it actually had: the purpose of using this as the numerator (the resulting weight is called the “stabilized weight”), rather than 1 is that the stabilized weights are less variable and, when the marginal structural model is not saturated—as in our case—the estimated coefficients also show less variability (Robins, Hernán, and Brumback 2000, p. 554). For male respondents without a college degree, the numerator is simply the proportion of all the male respondents in that model’s sample who do not have a college degree. For male respondents with a college degree, the numerator is the proportion of male respondents with a college degree. We repeat the same steps to compute the numerators for female respondents.

For those respondents who graduated from college, the denominator of ω_A is the estimated probability of having graduated college, which is the usual propensity score. For those respondents who did not graduate from college, the denominator is their estimated probability of not having graduated college, equal to 1 minus the propensity score for graduating from college.

Separate weights are computed for each equation. In equation (5) we weight by

$$\omega_A \times \omega_Z \text{ where } \omega_Z = \frac{\text{pr}(Z = z|A = a)}{\text{pr}(Z = z|A = a, X, \omega_A)},$$

and in equation (6) we weight by

$$\omega_A \times \omega_Z \times \omega_C \text{ where } \omega_C = \frac{\text{pr}(C = c|A = a, Z = z)}{\text{pr}(C = c|A = a, Z = z, X, \omega_A, \omega_Z)}.$$

In practice, well-behaved weights should have a mean close to one (Cole and Hernán 2008). Our weights meet this criterion. Further details about the distributions of the constructed weights are included in appendix table A3.

Computing Causal Effects from Marginal Structural Models

The causal effects are calculated from the coefficients of the weighted regression equations. The average causal effects are defined as the difference between the average potential outcomes. So, from equation (3), for example, we have

$$E(Z0^{A=1}) - E(Z0^{A=0}) = \alpha_0 + \alpha_1 - \alpha_0 = \alpha_1,$$

and our estimate of α_1 is our estimate of the causal effect of A on $Z0$. The same argument holds for equation (4). In equation (5) the causal effect of $Z0$ on C is estimated as

$$\begin{aligned} E(C^{A=a, Z0=1, Z2=0}) - E(C^{A=a, Z0=0, Z2=0}) \\ = \beta_0 + \beta_1(A = a) + \beta_2 - (\beta_0 + \beta_1(A = a)) \\ = \beta_2. \end{aligned}$$

In this case, $Z2$ must be set to zero, but A can be fixed at any value because it appears on both sides of the equation and so cancels.

The average total causal effect of A on C is estimated as

$$\begin{aligned} E(C^{A=1, Z0(A=1), Z2(A=1)}) - E(C^{A=0, Z0(A=0), Z2(A=0)}) \\ = \beta_0 + \beta_1 + \beta_2(\alpha_0 + \alpha_1) + \beta_3(\gamma_0 + \gamma_1) - (\beta_0 + \beta_2\alpha_0 + \beta_3\gamma_0) \\ = \beta_1 + \beta_2\alpha_1 + \beta_3\gamma_1. \end{aligned}$$

Here $Z0$ and $Z2$ are set to their expected values given the particular value of A . This total effect decomposes into a direct effect (obtained in a way analogous to the direct effect of $Z0$ on C explained above and equal to β_1) and an indirect effect estimated as

$$\begin{aligned} E(C^{A=0, Z0(A=1), Z2(A=1)}) - E(C^{A=0, Z0(A=0), Z2(A=0)}) \\ = \beta_0 + \beta_2(\alpha_0 + \alpha_1) + \beta_3(\gamma_0 + \gamma_1) - (\beta_0 + \beta_2\alpha_0 + \beta_3\gamma_0) \\ = \beta_2\alpha_1 + \beta_3\gamma_1. \end{aligned}$$

TABLE A1
SUMMARY STATISTICS AND COEFFICIENTS FOR COVARIATES INCLUDED IN PROPENSITY SCORE ESTIMATION MODELS, WLS MEN

| | Degree = No (<i>n</i> = 1,678) Mean (SD) | Degree = Yes (<i>n</i> = 857) Mean (SD) | Effect on Pr(A) Log Odds (SE) | Effect on Pr(Z0 vs. Z1) Log Odds (SE) | Effect on Pr(Z2 vs. Z1) Log Odds (SE) | Effect on Pr(C) Log Odds (SE) |
|--|---|--|-------------------------------------|---|---|----------------------------------|
| Mother's education in years | 10.18 (2.69) | 11.81 (3.05) | .06** (.02) | .01 (.04) | .03 (.02) | |
| Father's education in years | 9.69 (2.82) | 11.87 (3.41) | .09*** (.02) | .03 (.04) | .09*** (.02) | |
| Log of parents' 1957 income + 500 | 8.56 (.56) | 8.88 (.61) | .23* (.11) | | | |
| Number of siblings | 3.40 (2.66) | 2.58 (2.05) | -.02 (.03) | .03 (.04) | -.05 (.03) | |
| High school class rank percentile | 36.60 (24.11) | 66.16 (24.47) | -.25 (.15) | | | |
| IQ score (from eleventh-grade Hennon-Nelson test) | 44.99 (26.93) | 72.08 (22.97) | .03*** (.00) | .00 (.00) | .01* (.00) | |
| Rural background (0 = no, 1 = yes) | .34 (.47) | .17 (.38) | -.01*** (.00) | | | |
| Catholic (0 = no, 1 = yes) | .40 (.49) | .39 (.49) | .07 (.12) | -.11 (.21) | | -.16 (.12) |
| College prep academic program (0 = no, 1 = yes) | .55 (.50) | .92 (.27) | .70*** (.16) | | | |

| | | | | | |
|--|--------------|--------------|------------------|-------------------|--------------------|
| Parents encouraged college (0 = no; 1 = yes) | .48 (.50) | .91 (.28) | 1.00*** (.16) | | |
| Teachers encouraged college (0 = no, 1 = yes) . . . | .37 (.48) | .76 (.43) | .34** (.13) | | |
| Friends' college plans (0 = no, 1 = yes) | .22 (.42) | .69 (.46) | .98*** (.12) | | |
| College degree (A) (0 = no, 1 = yes) | | | | 1.11*** (.23) | 2.06*** (.14) |
| Never married (Z0) (0 = no, 1 = yes) | | | | | -.23 (.19) |
| Spouse with degree (Z2) (0 = no, 1 = yes) | | | | | -7.38*** (1.01) |
| $w(A)$ | | | | -.04 (.06) | -.47* (.20) |
| $w(Z)$ | | | | | .10 (.08) |
| Intercept | | | | | -.52 (.39) |
| | | | | -3.48*** (.49) | -3.80*** (.30) |
| | | | | | 3.19*** (.00) |

NOTE.— $N = 2,535$. WLS = Wisconsin Longitudinal Study.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE A2
SUMMARY STATISTICS AND COEFFICIENTS FOR COVARIATES INCLUDED IN PROPENSITY SCORE ESTIMATION MODELS, WLS WOMEN

| | Degree = No (n = 2,232) Mean (SD) | Degree = Yes (n = 513) Mean (SD) | Effect on Pr(A) Log Odds (SE) | Effect on Pr(Z0 vs. Z1) Log Odds (SE) | Effect on Pr(Z2 vs. Z1) Log Odds (SE) | Effect on Pr(C) Log Odds (SE) |
|--|---|--|-------------------------------------|---|---|-------------------------------------|
| Mother's education in years | 10.12 (2.78) | 12.32 (3.17) | .08** (.02) | -.01 (.04) | .08*** (.02) | |
| Father's education in years | 9.72 (2.90) | 12.08 (3.38) | .05* (.02) | .07* (.03) | .15*** (.02) | |
| Log of parents' 1957 income + 500 | 8.58 (.54) | 8.93 (.66) | .42*** (.11) | | | |
| Number of siblings | 3.43 (2.64) | 2.38 (1.72) | -.05 (.03) | -.05 (.04) | -.07*** (.02) | |
| High school class rank percentile | 55.68 (26.29) | 80.62 (18.70) | .44** (.16) | | | |
| IQ score (from eleventh-grade Hemmon-Nelson test) | 47.10 (26.95) | 72.50 (21.94) | .03*** (.00) | .00 (.00) | .01*** (.00) | |
| Rural background (0 = no, 1 = yes) | .29 (.45) | .20 (.40) | .01** (.00) | | | |
| Catholic (0 = no, 1 = yes) | .44 (.50) | .31 (.46) | -.21 (.13) | .13 (.19) | .21* (.10) | |
| College prep academic program (0 = no, 1 = yes) | .47 (.50) | .89 (.31) | .74*** (.18) | | | |

| | | | | | |
|--|--------------|--------------|-----------------|-------------------|-------------------|
| Teachers encouraged college (0 = no, 1 = yes) | .38 (.49) | .77 (.42) | .50*** (.14) | | |
| Friends' college plans (0 = no, 1 = yes) | .31 (.46) | .77 (.42) | .71*** (.14) | | |
| College degree (4) (0 = no, 1 = yes) | | | | 2.31*** (.22) | 1.94*** (.14) |
| Never married (Z0) (0 = no, 1 = yes) | | | | | -.70** (.21) |
| Spouse with degree (Z2) (0 = no, 1 = yes) | | | | | -5.55*** (.40) |
| $w(4)$ | | | | .24** (.08) | -.01 (.20) |
| $w(Z)$ | | | | | -.08 (.10) |
| Intercept | | | | | .04 (.25) |
| | | | | -4.00*** (.48) | -4.29*** (.28) |
| | | | | | 2.93*** (.28) |

NOTE.— $N = 2,745$. WLS = Wisconsin Longitudinal Study.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE A3
SUMMARY STATISTICS FOR WEIGHTS

| WEIGHT | MEAN | SD | PERCENTILES | | | |
|--------------------------------------|------|------|-------------|------|------|------|
| | | | 1st | 25th | 75th | 99th |
| Men: | | | | | | |
| All models: | | | | | | |
| $w(A)$ | 1.01 | 1.67 | .36 | .67 | .88 | 6.48 |
| $w(Z)$ | 1.00 | .17 | .65 | .96 | 1.03 | 1.64 |
| $w(A) \times w(Z)$ | 1.00 | 1.56 | .26 | .64 | .89 | 6.84 |
| C = any children: | | | | | | |
| $w(C)$ | 1.00 | .08 | .90 | 1.00 | 1.01 | 1.24 |
| $w(A) \times w(Z) \times w(C)$. . . | 1.01 | 1.62 | .26 | .63 | .90 | 6.80 |
| C = number of children: | | | | | | |
| $w(C)$ | 1.00 | .11 | .72 | .98 | 1.02 | 1.33 |
| $w(A) \times w(Z) \times w(C)$. . . | 1.00 | 1.72 | .26 | .62 | .90 | 6.36 |
| Women: | | | | | | |
| All models: | | | | | | |
| $w(A)$ | .97 | .91 | .23 | .82 | .95 | 3.61 |
| $w(Z)$ | 1.00 | .32 | .36 | .87 | 1.07 | 2.25 |
| $w(A) \times w(Z)$ | .95 | 1.11 | .19 | .71 | .99 | 3.49 |
| C = any children: | | | | | | |
| $w(C)$ | 1.00 | .02 | .95 | 1.00 | 1.00 | 1.13 |
| $w(A) \times w(Z) \times w(C)$. . . | .96 | 1.22 | .18 | .71 | .99 | 3.44 |
| C = number of children: | | | | | | |
| $w(C)$ | 1.00 | .08 | .81 | .98 | 1.02 | 1.27 |
| $w(A) \times w(Z) \times w(C)$. . . | .95 | .97 | .18 | .71 | .99 | 3.84 |

Equation (6) can be estimated only for those WLS respondents who had a child. But the marginal structural model refers to potential, rather than observed, outcomes, and those WLS respondents without children nevertheless are considered to have potential outcomes for Y that would have been realized had they had children. Thus we want to estimate the potential outcomes involving $C = 1$, but we have to use data on those whose observed outcome was $C = 1$. But this yields unbiased estimates because the potential outcomes, $Y^{A,Z0,Z2}$, are independent of C once we apply our weights (and given the assumptions that justify their use).

The causal effects of A on Y are calculated in the same way as the effects already discussed.

The direct effect of A on Y is given by

$$E(Y^{A=1,Z0(A=1),Z2(A=1)}) - E(Y^{A=0,Z0(A=1),Z2(A=1)})$$

and, omitting the algebra, is equal to θ_1 .

The indirect effect is given by

$$E(Y^{A=0,Z0(A=1),Z2(A=1)}) - E(Y^{A=0,Z0(A=0),Z2(A=0)})$$

and is equal to $\theta_2\alpha_1 + \theta_3\gamma_1$.

The total effect is the sum of these. We also have direct effects on Y of $Z0$ (θ_2) and $Z2$ (θ_3). The total, direct, and indirect causal effects of A on Y are the average causal effects of graduating college on having a child who graduates college. They tell us what the returns to college would be for all respondents of each sex, if they had at least one child. However, what we observe in the data is the product of two causal processes: having a child and having a child graduate college. To capture this, we define another set of causal effects as the product of these two: these are what we call, in the text, the realized causal effects. As we explain in the text, we can write the realized total effect in terms of potential outcomes as follows:

$$E(Y^{A=1, Z0(A=1), Z2(A=1)})E(C^{A=1, Z0(A=1), Z2(A=1)}) \\ - E(Y^{A=0, Z0(A=0), Z2(A=0)})E(C^{A=0, Z0(A=0), Z2(A=0)}).$$

This total effect decomposes into direct and indirect effects:

$$\text{Direct} = E(Y^{A=1, Z0(A=1), Z2(A=1)})E(C^{A=1, Z0(A=1), Z2(A=1)}) \\ - E(Y^{A=0, Z0(A=1), Z2(A=1)})E(C^{A=0, Z0(A=1), Z2(A=1)}),$$

and

$$\text{Indirect} = E(Y^{A=0, Z0(A=1), Z2(A=1)})E(C^{A=0, Z0(A=1), Z2(A=1)}) \\ - E(Y^{A=0, Z0(A=0), Z2(A=0)})E(C^{A=0, Z0(A=0), Z2(A=0)}).$$

In terms of parameters, some algebra yields, for the direct effect:

$$\beta_1(\theta_0 + \theta_2(\alpha_0 + \alpha_1) + \theta_3(\gamma_0 + \gamma_1)) \\ + \theta_1(\beta_0 + \beta_1 + \beta_2(\alpha_0 + \alpha_1) + \beta_3(\gamma_0 + \gamma_1)),$$

and for the indirect effect:

$$(\theta_2\alpha_1 + \theta_3\gamma_1)(\beta_0 + \beta_2(\alpha_0 + \alpha_1) + \beta_3(\gamma_0 + \gamma_1)) \\ + (\beta_2\alpha_1 + \beta_3\gamma_1)(\theta_0 + \theta_2\alpha_0 + \theta_3\gamma_0).$$

The sum of these two is the total effect.

The estimated coefficients from the four marginal structural models, equations (3) through (6), which are used to calculate the direct, indirect, and total effects are reported in appendix table A4 (for the models with C defined as having at least one child and Y defined as having at least one child with a college degree) and appendix table A5 (for the models with C defined as the number of children and Y defined as the proportion of children with college degrees).

TABLE A4
PARAMETER ESTIMATES FOR EQUATIONS FOR Z0, Z2, C (Any Children), AND Y (Any Child with College Degree)

| | MEN | | | | WOMEN | | | |
|----------------|-----------------|-----------------|------------------|------------------|-----------------|-----------------|------------------|------------------|
| | Z0 | Z2 | C | Y | Z0 | Z2 | C | Y |
| A..... | .02 (.02) | .25*** (.04) | .02 (.02) | .24*** (.04) | .23*** (.07) | .21*** (.05) | -.03 (.03) | .23*** (.04) |
| Z0..... | | | -.86*** (.08) | -.82*** (.04) | | | -.88*** (.03) | -.58*** (.03) |
| Z2..... | | | -.04 (.02) | -.14*** (.04) | | | -.01 (.02) | .19*** (.03) |
| Intercept..... | .04*** (.01) | .08*** (.01) | .93*** (.01) | .59*** (.02) | .03*** (.00) | .23*** (.01) | .95*** (.01) | .58*** (.01) |
| N..... | 2,535 | 2,535 | 2,535 | 2,252 | 2,745 | 2,745 | 2,745 | 2,455 |

NOTE.—Nos. in parentheses are SEs.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE A5
PARAMETER ESTIMATES FOR EQUATIONS FOR Z0, Z2, C (Number of Children),
AND Y (Proportion of Children with College Degrees)

| | C = 1 | C = 2 | C = 3 | C = 4 | C = 5 + | Y = Proportion |
|---------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Men: | | | | | | |
| A | .03 (.03) | .13** (.05) | -.02 (.04) | -.08*** (.02) | -.04** (.01) | .25*** (.03) |
| Z0 | -.08*** (.01) | -.25** (.08) | -.30*** (.02) | -.14*** (.01) | -.08*** (.01) | -.66*** (.03) |
| Z2 | -.01 (.03) | .04 (.07) | -.05 (.05) | -.01 (.04) | -.01 (.02) | .14*** (.04) |
| C = 1 | | | | | | .49*** (.04) |
| C = 2 | | | | | | .42*** (.02) |
| C = 3 | | | | | | .36*** (.02) |
| C = 4 | | | | | | .29*** (.03) |
| C = 5+ | | | | | | .28*** (.03) |
| Intercept | .07*** (.01) | .28*** (.02) | .31*** (.02) | .17*** (.01) | .10*** (.01) | |
| Women: | | | | | | |
| A | -.02 (.01) | .03 (.05) | .08 (.09) | -.08* (.04) | -.05 (.03) | .20*** (.03) |
| Z0 | .00 (.02) | -.23*** (.03) | -.35*** (.05) | -.17*** (.02) | -.13*** (.02) | -.39*** (.04) |
| Z2 | .02 (.01) | .06* (.03) | -.04 (.04) | .00 (.03) | -.06** (.02) | .21*** (.02) |
| C = 1 | | | | | | .38*** (.05) |
| C = 2 | | | | | | .38*** (.02) |
| C = 3 | | | | | | .36*** (.02) |
| C = 4 | | | | | | .29*** (.02) |
| C = 5+ | | | | | | .24*** (.02) |
| Intercept | .06*** (.01) | .22*** (.01) | .32*** (.02) | .20*** (.01) | .15*** (.01) | |

NOTE.—Nos. in parentheses are SEs.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

APPENDIX B

Sensitivity Analyses

Our assumptions about the absence of unmeasured confounders imply:

$$E(Y^{A=1}|A = 1, X) = E(Y^{A=1}|A = 0, X), \quad (B1)$$

and

$$E(Y^{A=0}|A = 0, X) = E(Y^{A=0}|A = 1, X). \quad (B2)$$

In words, the expected value of each counterfactual outcome, Y^A , is the same among those who graduated ($A = 1$) and those who did not graduate ($A = 0$) from college, given the same values on the confounders, X . Unmeasured selection bias would mean that one or both of these equalities did not in fact hold. Positive selection bias for both outcomes replaces both equalities with $>$, indicating that the potential outcome for treatment (nontreatment) is larger among those who received treatment (did not receive treatment) than among those who did not (did). Negative selection into college graduation, as recently argued by Brand and Xie (2010), means, in our case, that those who actually graduated from college have lower potential outcomes for college graduation than those who did not graduate, whereas potential outcomes for nongraduation do not differ. This means that in (B2) the equality holds but in (B1) the equality is replaced by $<$.

Both these forms of selection bias have been argued for in the literature, and so we test the robustness of our results to both of them. A third form of bias that we also consider arises when the potential outcomes in both states are larger for those who actually graduated from college. In this case the equality in (B1) is replaced with $>$ and that in (B2) with $<$. We refer to this as uniform selection bias.

To model the impact of selection bias we first define:

$$c(a, x) = E[Y^{A=a}|A = a, X] - E[Y^{A=a}|A = (1 - a), X].$$

Here a indicates the value of A : that is, 0 or 1 to denote control and treatment, and x is the individual's values on X . The difference in potential outcomes due to unobservables is thus given by $c(a, x)$. Assuming no selection on unobservables, the average treatment effect can be estimated as

$$E[Y|A = 1, X] - E[Y|A = 0, X].$$

But, given nonzero $c(a, x)$, this estimate will have a bias equal to

$$c(1, x)\text{pr}(A = 1|X) - c(0, x)\text{pr}(A = 1|X) \quad (B3)$$

(see Brumback et al. [2004, pp.753–54] for the derivation of this quantity). The idea, then, is to specify different ways in which unobserved factors can

play a role, allow the size of the hypothesized bias to vary, adjust the estimate of the causal effect for the bias, and see how robust the original estimate is to different types and extents of bias.

Different kinds of bias are specified by picking a functional form for $c(a, x)$. Setting $c(a, x) = \alpha$, where α is a constant, captures the positive selection case: people have a better outcome in the group (graduate or non-graduate) in which they find themselves. In this case, using (B3), the adjusted outcomes, Y^* , would be:

$$Y^* = Y - \alpha(\text{pr}(A = 0|X)), \quad \text{for } A = 1,$$

and

$$Y^* = Y - \alpha(\text{pr}(A = 1|X)), \quad \text{for } A = 0.$$

In both cases the observed value of the outcome is reduced to counter the inflation of these observed values by positive selection.

For negative selection into college graduation, we define $c(a, x) = \alpha\alpha$ and this gives:

$$Y^* = Y - \alpha(\text{pr}(A = 0|X)), \quad \text{for } A = 1,$$

and

$$Y^* = Y, \quad \text{for } A = 0.$$

Finally, the case in which college graduates do better in both potential outcomes, is captured by $c(a, X) = \alpha(2\alpha - 1)$ giving:

$$Y^* = Y - \alpha(\text{pr}(A = 0|X)), \quad \text{for } A = 1,$$

and

$$Y^* = Y + \alpha(\text{pr}(A = 1|X)), \quad \text{for } A = 0.$$

Here the outcomes for those who did not graduate are inflated to compensate for the fact that unmeasured selection bias means that this group contains an overrepresentation of people with low potential outcomes, $Y^{A=0}$.

As Sharkey and Elwert (2011, p. 1956) point out, this method “captures the totality of possible selection bias from any source and any number of omitted variables, in one simple function.” The bias is captured by the size of the α parameter, so, by varying its size and recomputing our causal estimates using the resulting Y^* in place of Y , we can gauge the sensitivity of those estimates to different degrees of bias.

Nevertheless, to interpret these results we need to calibrate α and we do this in the same way as Sharkey and Elwert (2011, pp. 1957–58) so that a unit change in α “corresponds to the amount of observed confounding previously eliminated by adjusting for all observed covariates in the MSM [marginal

structural model].” We apply our method to the causal estimate of the total realized effect of graduating college on having a child who also graduates college (results shown in table B1) and to the estimate of the total effect of graduating college on having a child (results shown in table B2).

Each column of the tables shows the application of a different function representing, respectively, uniform selection bias favoring those who graduated, positive selection, and negative selection into college graduation. Given our causal estimate (shown in the first row of the table) we searched for a value of α such that it generated Y^* which, when calculated using coefficients from an unweighted regression equation, yielded an estimate equal to our causal estimate. This is shown in the second row of the table, and

TABLE B1
SENSITIVITY ANALYSES FOR REALIZED TOTAL CAUSAL EFFECT OF A (Respondent's College)
ON Y (Having at Least One Child Graduate from College)

| | Uniform Selection $c(\alpha) = \alpha(2\alpha - 1)$ | Positive Selection $c(\alpha) = \alpha$ | Negative Selection $c(\alpha) = \alpha(\alpha)$ |
|--|--|--|--|
| | $\alpha = 0.13$ | $\alpha = 0.38$ | $\alpha = -0.19$ |
| Men: | | | |
| <i>AY</i> total weighted | .22*** (.04) | .22*** (.04) | .22*** (.04) |
| <i>AY*</i> total unweighted | .22*** (.03) | .22*** (.03) | .22*** (.03) |
| <i>AY*</i> total ($\alpha/2$) weighted . . . | .16*** (.04) | .16*** (.04) | .16*** (.04) |
| <i>AY*</i> total (α) weighted | .10*** (.04) | .10* (.04) | .10** (.04) |
| <i>AY*</i> total (2α) weighted . . . | -.01 (.04) | -.03 (.04) | -.02 (.04) |
| <i>AY*</i> total (3α) weighted . . . | -.13*** (.04) | -.15*** (.04) | -.13*** (.04) |
| Women: | | | |
| | $\alpha = -0.05$ | $\alpha = -0.10$ | $\alpha = 0.06$ |
| <i>AY</i> total weighted | -.05 (.08) | -.05 (.08) | -.05 (.08) |
| <i>AY*</i> total unweighted | -.05 (.08) | -.05 (.08) | -.05 (.08) |
| <i>AY*</i> total ($\alpha/2$) weighted . . . | -.03 (.08) | -.03 (.08) | -.03 (.08) |
| <i>AY*</i> total (α) weighted | -.01 (.08) | -.01 (.09) | -.01 (.08) |
| <i>AY*</i> total (2α) weighted . . . | .02 (.09) | .03 (.09) | .02 (.09) |
| <i>AY*</i> total (3α) weighted . . . | .05 (.09) | .07 (.10) | .06 (.09) |

* $P < .05$.

** $P < .01$.

*** $P < .001$.

TABLE B2
SENSITIVITY ANALYSES FOR REALIZED TOTAL CAUSAL EFFECT OF A
(Respondent's College) ON C (Having at Least One Child)

| | Uniform Selection $c(\alpha) = \alpha(2\alpha - 1)$ | Positive Selection $c(\alpha) = \alpha$ | Negative Selection $c(\alpha) = \alpha(\alpha)$ |
|--|--|--|--|
| | $\alpha = -0.07$ | $\alpha = -0.21$ | $\alpha = 0.10$ |
| Men: | | | |
| AC total weighted | -.01 (.03) | -.01 (.03) | -.01 (.03) |
| AC* total unweighted | -.01 (.02) | -.01 (.02) | -.01 (.02) |
| AC* total ($\alpha/2$) weighted . . | .03 (.03) | .03 (.03) | .03 (.03) |
| AC* total (α) weighted . . . | .06* (.03) | .07** (.03) | .07* (.03) |
| AC* total (2α) weighted . . . | .13*** (.03) | .14*** (.03) | .14*** (.03) |
| AC* total (3α) weighted . . . | .20*** (.03) | .22*** (.03) | .21*** (.03) |
| | $\alpha = -0.02$ | $\alpha = -0.03$ | $\alpha = 0.02$ |
| Women: | | | |
| AC total weighted | -.23*** (.06) | -.23*** (.06) | -.23*** (.06) |
| AC* total unweighted | -.23*** (.06) | -.23*** (.06) | -.23*** (.06) |
| AC* total ($\alpha/2$) weighted . . | -.22*** (.06) | -.22*** (.06) | -.22*** (.06) |
| AC* total (α) weighted . . . | -.22*** (.06) | -.22*** (.06) | -.22*** (.06) |
| AC* total (2α) weighted . . . | -.20** (.06) | -.20** (.06) | -.20** (.06) |
| AC* total (3α) weighted . . . | -.18** (.06) | -.18** (.06) | -.18** (.06) |

* $P < .05$.
 ** $P < .01$.
 *** $P < .001$.

at the head of each column we report the value of α that generated this outcome. For example, for positive selection among men in table B1, this was $\alpha = .38$. The remaining four rows show the causal estimates derived from an MSM using Y^* based on $\alpha/2$, α , 2α , and 3α . This allows us to assess how sensitive our causal estimates are to unmeasured selection ranging in magnitude from one-half to three times as large as measured selection.

In table B1 the results are consistent across all three types of bias. Among men, selection on unobservables would have to be twice as great as selection on observables to remove the positive realized causal effect of college graduation. We would estimate a significant negative effect only when selection

on unobservables is three times as great as selection on observables. Among women, in no case do we find a significant positive effect of college graduation. The results for fertility in table B2 show that, for women, the negative effect of college on the chances of having a child are robust to all forms and degrees of selection on unobservables. Among men, selection on observables equal to selection on unobservables could lead us to find a significant positive effect of college on men's probability of having a child. But, even if one finds this degree of selection plausible, this would magnify the gender difference in the effects of college on having a child.

REFERENCES

- Angrist, Joshua, and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, N.J.: Princeton University Press.
- Attewell, Paul, and David E. Lavin. 2007. *Passing the Torch*. New York: Russell Sage.
- Becker, Gary. 1991. *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.
- Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." Pp. 1487–1541 in *Handbook of Labor Economics*, vol. 4b. Edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier.
- Blau, Peter M., Otis Dudley Duncan, and Andrea Tyree. 1967. *The American Occupational Structure*. New York: John Wiley & Sons.
- Bourdieu, Pierre. 1977. "Cultural Reproduction and Social Reproduction." Pp. 487–510 in *Power and Ideology in Education*, edited by J. Karabel and A. H. Halsey. New York: Oxford University Press.
- Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne Groves. 2005. *Unequal Chances: Family Background and Economic Success*. New York: Russell Sage.
- Brand, Jennie E., and Dwight Davis. 2011. "The Impact of College Education on Fertility: Evidence for Heterogeneous Effects." *Demography* 48:863–87.
- Brand, Jennie E., and Yu Xie. 2010. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75 (2): 273–302.
- Breen, Richard, and Jan O. Jonsson. 2005. "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility." *Annual Review of Sociology* 31:223–44.
- Brumback, Babette, Miguel A. Hernán, Sebastien J. P. A. Haneuse, and James M. Robins. 2004. "Sensitivity Analyses for Unmeasured Confounding Assuming a Marginal Structural Model for Repeated Measures." *Statistics in Medicine* 23:749–67.
- Buchmann, Claudia, and Thomas A. DiPrete. 2006. "The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement." *American Sociological Review* 71:515–41.
- Cole, Stephen R., and Miguel A. Hernán. 2008. "Constructing Inverse Probability Weights for Marginal Structural Models." *American Journal of Epidemiology* 168 (6): 656–64.
- Conley, Dalton. 2001. "Capital for College: Parental Assets and Postsecondary Schooling." *Sociology of Education* 74:59–72.
- Conley, Dalton, and Rebecca Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size." *Journal of Human Resources* 41 (4): 722–37.

- Duncan, Otis Dudley. 1966. "Methodological Issues in the Analysis of Social Mobility." Pp. 51–97 in *Social Structure and Mobility in Economic Development*, edited by N. J. Smelser and S. M. Lipset. Chicago: Aldine.
- Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change*. New York: Academic Press.
- Friedman, Esther M., and Robert D. Mare. 2014. "The Schooling of Offspring and the Survival of Parents." *Demography* 51:1271–93.
- Greenland, Sander, Judea Pearl, and James M. Robins. 1999. "Causal Diagrams for Epidemiologic Research." *Epidemiology* 10 (1): 37–48.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–61.
- Hernán, Miguel A., Babette Brumback, and James M. Robins. 2001. "Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments." *Journal of the American Statistical Association* 96:440–48.
- Hernán, Miguel A., and James M. Robins. 2006. "Estimating Causal Effects from Epidemiological Data." *Journal of Epidemiology and Community Health* 60 (7): 578–86.
- Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.
- Hout, Michael. 2015. "A Summary of What We Know about Social Mobility." *ANNALS of the American Academy of Political and Social Science* 657 (1): 27–36.
- Jaeger, Mads Meier. 2012. "The Extended Family and Children's Educational Success." *American Sociological Review* 77 (6): 903–22.
- Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Lareau, Annette. 2003. *Unequal Childhoods*. Princeton, N.J.: Princeton University Press.
- MacLeod, Jay. 1989. *Ain't No Makin' It: Aspirations and Attainment in a Low-Income Neighborhood*. Boulder, Colo.: Westview Press.
- Maralani, Vida. 2013. "The Demography of Social Mobility: Black-White Differences in the Process of Educational Reproduction." *American Journal of Sociology* 118 (6): 1509–58.
- Mare, Robert D. 1996. "Demography and the Evolution of Educational Inequality." Pp. 117–51 in *Social Differentiation and Social Inequality: Theoretical and Empirical Inquiries; Essays in Honor of John Pock*, edited by J. Baron, D. Grusky, and D. Treiman. Boulder, Colo.: Westview Press.
- . 2011. "A Multigenerational View of Inequality." *Demography* 48 (1): 1–23.
- Mare, Robert D., and Vida Maralani. 2006. "The Intergenerational Effects of Changes in Women's Educational Attainments." *American Sociological Review* 71:542–64.
- Marini, Margaret Mooney. 1980. "Sex Differences in the Process of Occupational Attainment: A Closer Look." *Social Science Research* 9:307–61.
- McLanahan, Sara S., and Gary D. Sandefur. 1994. *Growing Up with a Single Parent*. Cambridge, Mass.: Harvard University Press.
- Robins, James M. 1999. "Association, Causation, and Marginal Structural Models." *Synthese* 121:151–79.
- Robins, James M., Miguel A. Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11:550–60.
- Schwartz, Christine R., and Robert D. Mare. 2005. "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography* 42 (4): 621–646.
- Sharkey, Patrick, and Felix Elwert. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." *American Journal of Sociology* 116 (6): 1934–81.
- Song, Xi, and Robert D. Mare. 2015. "Prospective versus Retrospective Approaches to the Study of Intergenerational Social Mobility." *Sociological Methods and Research* 44 (4): 555–84.

- Torche, Florencia. 2015. "Analyses of Intergenerational Mobility: An Interdisciplinary Review." *ANNALS of the American Academy of Political and Social Science* 657 (1): 37–62.
- VanderWeele, Tyler J. 2009. "Marginal Structural Models for the Estimation of Direct and Indirect Effects." *Epidemiology* 20: 18–26.
- Warren, John R., and Robert M. Hauser. 1997. "Social Stratification across Three Generations: New Evidence from the Wisconsin Longitudinal Study." *American Sociological Review* 62:561–72.
- Winship, Christopher, and Robert D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18:327–50.
- WLS Handbook. 2006. *Wisconsin Longitudinal Study Handbook*. Madison: University of Wisconsin—Madison.
- Wodtke, Geoffrey, David Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective." *American Sociological Review* 76 (5): 713–36.