Article

Sociological Methods & Research
41(1) 17–56

DOI: 10.1177/0049124112440797
http://smr.sagepub.com

Models of College Entry in the United States and the Challenges of Estimating Primary and Secondary Effects

Stephen L. Morgan[1]

Abstract

In recent work in the sociology of education, the primary effects of stratification are defined as the effects of social class origins on preparation for future educational attainment. The secondary effects are then the net direct effects of social class on the transition to the next higher level of education, which are interpreted as average effects of class conditions on students' choices. In this article, the standard model of primary and secondary effects is laid out in explicit form, and primary and secondary effects are then estimated for college entry among recent high school graduates in the United States. The challenges of estimating these effects for educational transitions in the United States are then explained, focusing on the weak warrant for causal inference that is used to justify the calculation of counterfactual net differences across classes. Alternative estimates of associational primary and secondary effects are then offered, after which the prospects for the identification of causal secondary effects by conditioning on additional confounders are assessed. In conclusion, an appeal is made for attention to the policy-relevant patterns of heterogeneity that research on primary and secondary effects may be able to reveal, and a case is made for elaboration and direct

[1]Cornell University, Ithaca, NY, USA

Corresponding author:
Stephen L. Morgan, Department of Sociology, Cornell University, 358 Uris Hall, Ithaca, NY 14853, USA
Email: morgan@cornell.edu

identification of the presupposed choice mechanism that has been assumed to generate secondary effects in past research.

Keywords

causal effect; college entry; heterogeneity

Since the 1940s, social stratification researchers have devoted considerable effort to modeling the effects of social background on educational attainment in the United States. Erikson and Jonsson (1996), Erikson et al. (2005), and Jackson et al. (2007) have developed new methods for estimating background effects on educational attainment that have shown their value in empirical models of educational transitions in European societies. Should these new methods also be adopted by researchers who investigate similar educational transitions in the United States?

The methods proposed by Erikson and his colleagues adopt a conceptually appealing distinction between the primary and secondary effects of stratification, usually attributed to a model of educational opportunity proposed by Boudon (1974).[1] The primary effects of stratification on educational attainment are the effects of social class origins on the current levels of performance that constitute preparation for higher levels of educational attainment. The secondary effects are then the net effects of social class origins on educational attainment, which are interpreted as choice-based effects that emerge after preparation-determined effects have been purged from the total effects of class. These secondary effects, it is maintained, reveal the prevalence of decisions among lower-class students to forego the option of pursuing educational degrees that they are sufficiently prepared to pursue.

Erikson et al. (2005:9730) note that their empirical methods are based on an implied model where ''a student achieves a level of academic performance and then makes their choice about whether to continue'' to the next level of education. Combining the implicit assumptions of this model with empirical data, strong conclusions have been offered for educational transitions in Great Britain and Sweden. Jackson et al. (2007:224, italics in original), for example, write,

> [I]t is a serious error to concentrate attention entirely on class differences in academic performance. . . . Over and above differences of this kind, class differences further occur in the choices that are made by students, . . . with students from less advantaged class backgrounds being less likely to take educationally more

> ambitious options than students from more advantaged backgrounds, *and even when their academic performance would make such options entirely feasible for them*. This conclusion in turn raises major theoretical and policy issues.

Jackson et al. (2007:224) then argue for policies that target ‘‘resource and informational constraints’’ to promote alternative choices among the high performers in the working class.[2]

All preexisting work on primary and secondary effects has analyzed educational transitions in European societies. This article presents estimates of primary and secondary effects on college entry in the United States and then demonstrates the methodological challenges that confront the methods of this new research. These challenges include (a) a weak causal warrant for interpreting net class coefficients from simple logit models as evidence of class-based choices, (b) reliance on average log-odds measures of relative importance, and (c) inattention to consequential forms of unit-level heterogeneity. The first challenge may be particularly acute for educational transitions in the United States, but the second and third are quite general.

The article proceeds in five sections. First, the standard model of primary and secondary effects is laid out in explicit form. Second, primary and secondary effects are estimated under this formalization. Third, a critique of the existing model is offered with reference to educational attainment processes in the United States. Fourth, a descriptive set of associational models is developed, which it is argued should represent a first step in future modeling of primary and secondary effects. Fifth, the prospects for causal models of primary and secondary effects are then assessed with an augmented model that conditions on some of the variables presumed to lie along the backdoor paths that confound the causal effects in the standard model. The article concludes with an appeal for direct modeling of the choice-based mechanism that is implicitly assumed to generate secondary effects and a discussion of unexamined forms of policy-relevant heterogeneity.

A Simple Structural Model of Primary and Secondary Effects

Figure 1a presents a causal graph that, as noted by Erikson et al. (2005), is the simplest possible causal model of primary and secondary effects. For the graph in Figure 1a, *Class* is a categorical measure of social origins, *P* is a measure of recent performance in school, and *G* is an indicator variable for whether a student goes on to the next level of education. Although

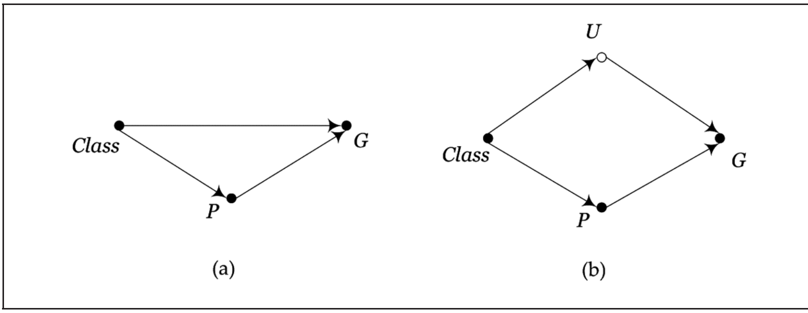

Figure 1. Two variants of the assumed causal model for the putative primary and secondary effects of stratification

models in this tradition of analysis are meant to apply to all educational transitions for which choices are possible, in this article the transition of interest is entry into four-year colleges among recent high school graduates in the United States.

For the graph in Figure 1a, the primary and secondary effects are the causal pathways $Class \rightarrow P \rightarrow G$ and $Class \rightarrow G$, respectively.[3] In path-modeling language, the primary effect is the indirect effect of class origins on college entry that is mediated by academic performance in prior schooling. The secondary effect is the remaining direct effect of class origins on college entry.

In the literature inspired by Boudon, the secondary effect has a choice-theoretic foundation. The causal graph in Figure 1a does not convey in any way what Boudon means when he writes that ‘‘the primary effects of stratification cause the youngsters to be differently distributed as a function of their family status in . . . such dimensions as school achievement’’ and then that ‘‘the secondary effects of stratification have the result that the probabilities of choosing [educational course] *a* rather than *b*, which are associated with each point in this space [defined by the dimension of school achievement], will be greater the higher the social status’’ (Boudon 1974:30). The following model brings this choice-theoretic framing in line with the causal graph in Figure 1a.

Suppose that there are three classes to which students belong based on the occupations of their parents, as in Erikson et al. (2005) and Jackson et al. (2007). Membership in these classes is measured by three indicator variables S, I, and W for the salariat, intermediate, and working classes, respectively.

Suppose that these three classes are mutually exclusive and exhaustive categories defined by *Class* in Figure 1a.

The decision of whether to go to college after high school is determined by

$$G = 1 \ \text{if} \, f(P, U) > \tau \qquad (1)$$

$$G = 0 \ \text{if} \, f(P, U) \leq \tau$$

where G is an indicator variable for whether a student goes on to college and P is a measure of performance at the end of high school. In addition, U is a composite measure that includes all other substantive determinants of college entry, scaled such that higher values of each component of U always predispose an individual to go to college.[4] Finally, $f(.)$ is an arbitrary function that is weakly increasing in both P and U, and τ is an invariant threshold for the educational transition.[5]

To now link the model in equation (1) to standard procedures for estimating primary and secondary effects, consider the elaborated causal graph in Figure 1b, as well as four assumptions about the relationships between $Class$, P, U, and G that generate its structure:

- Assumption 1 (*Class* is a common cause): P and U are increasing in the levels of *Class*, such that $E[P|S=1] \geq E[P|I=1] \geq E[P|W=1]$ and $E[U|S=1] \geq E[U|I=1] \geq E[U|W=1]$.
- Assumption 2 (P and U are conditionally independent): P and U are independent of each other within *Class*, such that $\Pr[P|Class, U] = \Pr[P|Class]$ and $\Pr[U|Class, P] = \Pr[U|Class]$.
- Assumption 3 (no backdoor paths from *Class*): There are no variables X that lie along backdoor paths that begin at *Class* and end at G, such as $Class \leftarrow X \rightarrow G$, $Class \leftarrow X \rightarrow P \rightarrow G$, or $Class \leftarrow X \rightarrow U \rightarrow G$.
- Assumption 4 (no backdoor paths from P and U): There are no variables V that lie along backdoor paths that begin at either P or U and end at G, such as $P \leftarrow V \rightarrow G$ or $U \leftarrow V \rightarrow G$.

Consider how these four assumptions, when paired with the model in equation (1), generate the causal graph in Figure 1b. Assumption 1 stipulates the existence of the causal effects $Class \rightarrow P$ and $Class \rightarrow U$, and the model in equation (1) stipulates the existence of the causal effects $P \rightarrow G$ and $U \rightarrow G$.[6] Assumption 2 stipulates that the causal pathways $Class \rightarrow P \rightarrow G$

and $Class \rightarrow U \rightarrow G$ are isolated from each other, in the sense that there are no direct causal effects between P and U that would generate the additional pathways $Class \rightarrow P \rightarrow U \rightarrow G$ or $Class \rightarrow U \rightarrow P \rightarrow G$. Assumptions 3 and 4 stipulate that there are no causal variables that generate associations between $Class$ and G, between P and G, or between U and G that do not lie along the direct causal chains from $Class \rightarrow P \rightarrow G$ and $Class \rightarrow U \rightarrow G$.[7] Thus, this assumption states that no common causes of any two variables in the causal graph in Figure 1b have been erroneously suppressed.

Maintaining the simple structural model in equation (1) and assuming the causal structure depicted in the graph in Figure 1b based on Assumptions 1 through 4, the primary and secondary effects of stratification are then the two separable causal pathways $Class \rightarrow P \rightarrow G$ and $Class \rightarrow U \rightarrow G$, respectively. The primary effect is the effect of $Class$ on G that operates indirectly via P, and the secondary effect is the effect of $Class$ on G that operates indirectly via U. The graph in Figure 1a is best regarded as a simplification of the graph in Figure 1b, which is permissible because U is unobserved.[8]

Estimates of Primary and Secondary Effects on College Entry in the United States

In this section, estimates are offered for the primary and secondary effects of stratification on college entry among high school graduates in the United States. The data are drawn from the 2002–2006 Education Longitudinal Study (ELS), which is a national sample of high school sophomores in 2002, followed up in 2004 and 2006. Additional details of the data source and the variables used in the analysis are provided in the appendix (which can be found at http://smr.sagepub.com/supplemental/).

Table 1 presents the distribution of the sample, and the central variables of the following analysis, across two alternative operationalizations of ‘‘class.’’ The rows of Table 1 are defined using the standard measure of social class employed in past analyses of primary and secondary effects: the three-tiered Erikson–Goldthorpe (EG) class schema of salariat, intermediate, and working classes. This variable is a recoding of the occupational positions of mothers and fathers into a single ‘‘family class’’ variable. The columns of Table 1 are defined by the most common measure of family background utilized in the modeling of educational attainment in the sociology of education in the United States, socioeconomic status (hereafter, SES). The SES variable that is utilized is a linear composite of mother’s education, father’s education, mother’s occupational prestige, father’s occupational prestige, and

Table 1. College Entry and Performance by Class and SES Group

Class	SES group: Higher tertile	Middle tertile	Lower tertile	Total
Salariat				
Performance	.568	.021	–.460	.331
College entry	65.8%	37.6%	21.4%	54.1%
N	3,424	1,905	399	5,729
% of total	30.0%	16.7%	3.5%	50.3%
Intermediate				
Performance	.465	.051	–.479	–.072
College entry	67.6%	43.4%	25.0%	40.6%
N	296	1,117	837	2,250
% of total	2.6%	9.8%	7.3%	19.7%
Working				
Performance	–.178	.011	–.404	–.296
College entry	29.5%	42.0%	23.9%	23.9%
N	32	813	2,576	3,421
% of total	0.3%	7.1%	22.6%	30.0%
Total				
Performance	.554	.028	–.426	.076
College entry	65.6%	40.2%	23.9%	44.3%
N	3,752	3,835	3,812	11,400
% of total	32.9%	33.6%	33.4%	100%

Source: Education Longitudinal Study of 2002–2006.
Notes: Numbers do not always add perfectly across rows and columns because the data are weighted and then cell sizes are rounded.

family income. For Table 1, and in subsequent analysis, the sample is divided into SES tertiles, thereby obtaining three equal-sized SES groups for comparison with the three (not-equal-sized) EG classes more typically analyzed for models of primary and secondary effects.

Two basic patterns characterize Table 1. First, class and SES groups are strongly but imperfectly related. Because SES includes parental occupation as one of its components, SES and class share a common component. Yet the relationship between the two is not exact because SES is a composite ranking that is based also on variables for parental education and family income.

Second, both performance in high school and subsequent college entry vary with both measures of family background in expected ways. The highest mean level of performance, .568 on a standardized scale, is observed for

Table 2. Logit Models for College Entry Among High School Graduates

	Model 1	Model 2	Model 3
Constant	0.162 (0.043)	–0.236 (0.047)	–0.246 (0.048)
Class			
Intermediate	–0.545 (0.058)	–0.190 (0.069)	–0.187 (0.072)
Working	–1.081 (0.060)	–0.558 (0.067)	–0.543 (0.067)
Performance		1.235 (0.038)	1.266 (0.050)
× Intermediate			0.030 (0.085)
× Working			–0.142 (0.080)
Wald χ^2	336.7	1,170.0	1,206.0
N	11,400	11,400	11,400

Source: See note to Table 1.
Standard errors in parentheses.

members of the salariat who are also in the highest SES group. This group represents 30 percent of the full sample, and 65.8 percent of these students then carried on to a four-year college within six months of graduating from high school. At the opposite extreme, the 22.6 percent of the sample who were in the working class and also the lowest SES group had a mean performance level of –.404 on a standardized scale. Only 23.9 percent of these students entered college after high school graduation. The SES and class marginal distributions in the last row and column of the table, respectively, show that higher levels of SES and class are always associated with higher levels of performance and college entry. In the joint distribution in the cells of the table, slight variation in the pattern is present. Some of the departures from conditional monotonicity may be genuine, but sampling error for some of the small cell sizes (especially for high-SES, working-class students) is a more likely explanation. Classification error is also a candidate explanation.

In the main body of this article, the analysis of primary and secondary effects is offered for EG classes to preserve fidelity with the extant literature on primary and secondary effects that is critiqued. In the appendix, equivalent models are presented using SES groups. As detailed there, the overall patterns are broadly similar, with SES only slightly more predictive.

Models of College Entry and Predicted Entry Rates

Table 2 presents three logit models for high school seniors that predict entry into a four-year college within six months of high school graduation. All three logit models are weighted by the estimated inverse probability of being in the analysis sample, multiplied by the panel weight developed by the data distributors.[9]

Model 1 is a null model presented to demonstrate that baseline associations between class and college entry are substantial. The magnitude of these effects was already presented in the final column of Table 1. As shown there, the college entry rates are 54.1, 40.6, and 28.5 percent for the salariat, intermediate, and working classes, respectively. The coefficients of Model 1 demonstrate that, when parameterized in a logit model, the lower observed rates for the intermediate and working classes are not only substantially different but also highly statistically significant because the coefficients are high multiples of their standard errors. When performance is added for Model 2, net class differences decline. When interactions between performance and class are included for Model 3, the fit improves slightly. The association between performance and college entry may be slightly weaker for the working class, but this difference is just within conventional levels of expected sampling error.

Either Model 2 or Model 3 could be considered a viable representation of primary and secondary effects in this tradition of analysis since each model estimates net class effects that persist after an adjustment for performance. The recent literature on primary and secondary effects has established the more highly parameterized Model 3 as the standard model.

Predicted probabilities from Models 2 and 3 are presented in the two panels of Table 3. The nine cells of each panel give the estimated probability for each of the three classes that would prevail if the class kept its own performance distribution but transitioned through the estimated college entry regime for each of the three classes.

The diagonals of each of the panels are the same, as they report simple unconditional probabilities of college entry, which are .541, .406, and .285 for the salariat, intermediate, and working classes, respectively. These are simply the predicted entry rates for each of the three classes if they moved through their own observed college entry regime.

The values off the diagonals of each panel are what-if college entry rates, labeled ‘‘counterfactual transition rates’’ in the primary and secondary effects literature (see Jackson et al. 2007, Table 5; also see Erikson and Rudolphi 2010, Table 2, for corresponding ‘‘counterfactual log odds’’). Because the results are very similar for Models 2 and 3 from Table 2, consider only the second panel of Table 3, which is based on the standard model of primary and secondary effects. The value of .423 in the upper-right-hand cell of the table is the estimated probability that the salariat

Table 3. College Entry Rates by Class Origins, Estimated From the Models in Table 2 That Adjust for Prior Performance in the Senior Year of High School

	Model 2 (from Table 2): What-if class transition probability		
	Salariat	Intermediate	Working
Salariat	0.541 (0.011)	0.504 (0.014)	0.433 (0.013)
Intermediate	0.441 (0.012)	0.406 (0.014)	0.340 (0.013)
Working	0.382 (0.011)	0.348 (0.013)	0.285 (0.011)
	Model 3 (from Table 2): What-if class transition probability		
	Salariat	Intermediate	Working
Salariat	0.541 (0.011)	0.507 (0.015)	0.423 (0.014)
Intermediate	0.439 (0.012)	0.406 (0.014)	0.336 (0.013)
Working	0.380 (0.011)	0.346 (0.013)	0.285 (0.011)

Source: See note to Table 1.
See note to Table 2.

would enter college, assuming that the salariat retained its performance distribution but moved through the college entry regime of the working class. Rather than moving through a logit model with an index function of $-.246 + (1.266)$ *Performance* as its argument, the salariat moves through a logit model with an index function of $(-.246 - .543) + (1.266 - .142)$ *Performance* as its argument. Thus, the observed college entry rate of 54.1 percent on the diagonal is reduced in this counterfactual scenario to 42.3 percent. Similar comparisons reveal a strict ordering of transition rates, as determined by the distribution of performance and the estimated logit parameters. Higher classes have lower college entry rates when they pass through the college entry regimes of lower classes, and vice versa.

From College Entry Models to Estimates of Primary and Secondary Effects

Where are the primary and secondary effects in Table 3? This section presents two ways of converting the predicted probabilities in Table 3 into estimates of causal primary and secondary effects: probability-based estimates and odds-based estimates.

Probability-based estimates. Six secondary effects exist for counterfactual movement in both directions between the three classes. For the comparison of the salariat to the working class, there are two contrasts

$$SE_{S\text{ to }W} \equiv \sum_p \{E[G|W=1,P] - E[G|S=1,P]\} \Pr[P|S=1], \tag{2}$$

$$SE_{W\text{ to }S} \equiv \sum_p \{E[G|S=1,P] - E[G|W=1,P]\} \Pr[P|W=1]. \tag{3}$$

where the summation is over the countable set of values p that performance, P, can take on. For the comparison of the salariat to the intermediate class, there are two analogous contrasts:

$$SE_{S\text{ to }I} \equiv \sum_p \{E[G|I=1,P] - E[G|S=1,P]\} \Pr[P|S=1], \tag{4}$$

$$SE_{I\text{ to }S} \equiv \sum_p \{E[G|S=1,P] - E[G|I=1,P]\} \Pr[P|I=1]. \tag{5}$$

For the comparison of the intermediate class to the working class, there are two final contrasts:

$$SE_{I\text{ to }W} \equiv \sum_p \{E[G|W=1,P] - E[G|I=1,P]\} \Pr[P|I=1], \tag{6}$$

$$SE_{W\text{ to }I} \equiv \sum_p \{E[G|I=1,P] - E[G|W=1,P]\} \Pr[P|W=1]. \tag{7}$$

Accordingly, point estimates of these six effects can be calculated directly from the cells in Table 3. For the standard model in the second panel, the secondary effects are as follows:

$$\begin{aligned} SE_{S\text{ to }W} &= .423 - .541 = -.118, \\ SE_{W\text{ to }S} &= .380 - .285 = .095, \\ SE_{S\text{ to }I} &= .507 - .541 = -.034, \\ SE_{I\text{ to }S} &= .439 - .406 = .033, \\ SE_{I\text{ to }W} &= .336 - .406 = -.070, \\ SE_{W\text{ to }I} &= .346 - .285 = .061. \end{aligned} \tag{8}$$

Notice the patterned directionality of the effects, which is determined by the structure of the definitions and the parameter estimates from the underlying logit model. Counterfactual downward moves generate negative effects on college entry, and counterfactual upward moves generate positive effects on college entry. For example, $SE_{S \text{ to } W} = -.118$, but $SE_{W \text{ to } S} = .095$. These shifts are interpretable in clear substantive terms, if they are regarded as causal effects. For example, they suggest that 11.8 percent of the salariat that entered college would not have chosen to enter college if they instead had the choice-based secondary effect that applied to the working class. Given the scale of typical effects in college entry research, 11.8 percent of students in one class and 9.5 percent of students in another class would represent large purported effects of choice.

Primary effects are then defined as the portion of the total class differential that is not attributed to the secondary effect,

$$\begin{aligned}
PE_{S \text{ to } W} &\equiv [E(G|W=1) - E(G|S=1)] - SE_{S \text{ to } W},\\
PE_{W \text{ to } S} &\equiv [E(G|S=1) - E(G|W=1)] - SE_{W \text{ to } S},\\
PE_{S \text{ to } I} &\equiv [E(G|I=1) - E(G|S=1)] - SE_{S \text{ to } I},\\
PE_{I \text{ to } S} &\equiv [E(G|S=1) - E(G|S=1)] - SE_{I \text{ to } S},\\
PE_{I \text{ to } W} &\equiv [E(G|W=1) - E(G|I=1)] - SE_{I \text{ to } W},\\
PE_{W \text{ to } I} &\equiv [E(G|I=1) - E(G|W=1)] - SE_{W \text{ to } I},
\end{aligned} \tag{9}$$

where the differences in brackets in each definition are direction-defined total class differentials. For the standard model in the second panel of Table 3, the primary effects would then be,

$$\begin{aligned}
PE_{S \text{ to } W} &= [.285 - .541] - (-.118) = -.138,\\
PE_{W \text{ to } S} &= [.541 - .285] - .095 = .161,\\
PE_{S \text{ to } I} &= [.406 - .541] - (-.034) = -.101,\\
PE_{I \text{ to } S} &= [.541 - .406] - .033 = .102,\\
PE_{I \text{ to } W} &= [.285 - .406] - (-.070) = -.051,\\
PE_{W \text{ to } I} &= [.406 - .285] - .061 = .060,
\end{aligned} \tag{10}$$

where the differences in brackets are total class differentials extracted from the diagonal of the panel and the values of the secondary effects are taken from equation (8).

Under these definitions of primary and secondary effects, the model allows for six class-referenced comparisons of the relative sizes of primary and secondary. For example, the observed total class differential for upward movement from the working class to the salariat is .256, which can be decomposed into a primary effect of .161 and a secondary effect of .095. In the opposite direction, the total class differential is –.256, which can be decomposed into a primary effect of –.138 and a secondary effect of –.118. Thus, each pairwise class comparison is decomposable in two ways, depending on the direction of the estimate of the secondary effect.

The extant literature on these effects does not discuss why these two decompositions differ. The reason should be clear from the definitions in equations (2) through (7). The secondary effect estimates are averages of individual-level effects over the performance distribution in the class from which counterfactual movement originates. Because the existence of primary effects generates class differences in performance, the direction of the calculation of the secondary effects will generate differences in the magnitudes of estimated secondary effects.

Odds-based estimates. Before carrying on to analyze the weaknesses of the underlying model that generates these estimates, it is necessary to consider how primary and secondary effects are defined and calculated in the work of Erikson et al. (2005), Jackson et al. (2007), and others following their lead. Although not too dissimilar from what has just been introduced, there are differences of note that result from (a) shifting away from probability differences to log odds ratios and (b) averaging further to produce combined estimates of primary and secondary effects that ignore the direction of counterfactual movement.

To make the differences clear, consider an equivalent demonstration of the calculations for movements between the salariat and the working class (again based on Model 3, using the four probabilities in the corners of the second panel of Table 3: .541, .423, .380, and .285). The relevant four steps, as outlined most clearly in Jackson et al. (2007), are the following:

1. Convert the predicted probabilities to odds:

$$\text{Observed odds for salariat:} \quad \frac{.541}{1-.541} = 1.179,$$

$$\text{Odds for salariat if working:} \quad \frac{.423}{1-.423} = .733,$$

$$\text{Odds for working if salariat: } \frac{.380}{1-.380} = .613,$$

$$\text{Observed odds for working: } \frac{.285}{1-.285} = .399.$$

2. Calculate three log odds ratios:

$$\text{Observed log odds ratio for salariat vs. working: } \ln\left(\frac{1.179}{.399}\right) = 1.083,$$

$$\text{What}-\text{if log odds ratio for salariat: } \ln\left(\frac{1.179}{.733}\right) = .475,$$

$$\text{What}-\text{if log odds ratio for working: } \ln\left(\frac{.613}{.399}\right) = .429,$$

where the order of division for the what-if log odds ratios is chosen to result in positive log odds ratios only.

3. Calculate the relative importance of primary and secondary effects as two ratios of the log odds ratios calculated in the last step:

$$\text{Relative importance for salariat to working: } \frac{.475}{1.083} = .439, \tag{11}$$

$$\text{Relative importance for working to salariat: } \frac{.429}{1.083} = .396. \tag{12}$$

4. Take the average of these two relative importance measures:

$$\frac{.439 + .396}{2} = .418. \tag{13}$$

The secondary effect of class origins in a comparison of the salariat to the working class is then equal to 41.8 percent of the total class differential in the log odds ratio. That which is left over, in this case 58.2 percent, is the primary effect. Following this procedure for all comparisons of classes would result in three measures of average relative importance (by equation (13)), which themselves would be averaged versions of six underlying relative importance measures calculated from log odds ratios (as in equations (11) and (12)).

Notice that, in Step 4, two nominal relative importance measures are averaged together in pursuit of a single number summary for the whole class comparison. The difference in the sizes of the two measures of relative importance in equations (11) and (12) receives no attention, even though the differences are produced by genuine individual-level heterogeneity that is then averaged across alternative distributions of performance that exist in the two reference classes.

Should probability-based effects be preferred?. In comparison to the odds-ratio-based measures of primary and secondary effects, the probability-based primary and secondary effects (equations (2) through (7) as well as equation (9)) have an advantage of transparency. The causal effects are defined as differences in probabilities, and the direction of counterfactual movement is clearly denoted by sign. Probability differences are more natural than odds ratios when rates of entry are being calculated, as is commonly the case for models of college entry in the United States. Probabilities have clear frequency interpretations, which allow for straightforward interpretations as expected shifts in portions of the relevant subpopulation. The probability-based effects also make clear that secondary effects are averages of individual-level effects that are weighted by the distribution of performance in the class from which counterfactual movement originates.

The odds-ratio-based effects have one clear advantage. They can be brought more easily into dialogue with classic results from social mobility studies, where odds ratios are often decomposed in intergenerational models of class entry. Other advantages may exist, but these have not been fully articulated by the developers of these methods.

On balance, a case can be made that the probability-based effects should be preferred, at least for college entry in the United States. For this reason, probability-based effects are offered in more detail in this article.

Causal Identification of the Standard Model

Beyond the choices of the measure for class and log-odds representations of counterfactual rates, there are more fundamental challenges to these methods *if* the estimated effects are meant to be imbued with causal interpretations that yield policy implications (as they have been in the past). As mentioned earlier, these challenges may be more acute when the methods are applied to educational transitions in the United States.

Return to Figure 1b and the Assumptions 1 through 4 that justify it as a causal model of primary and secondary effects. Is there any support for the Assumptions 2, 3, and 4 that stipulate that the primary and secondary effects can be isolated from each other and that confounding common causes do not exist?

First, consider Assumption 2, which states that P and U are independent conditional on *Class*. The recent literature on primary and secondary effects has challenged this assumption. Erikson et al. (2005:9733) recognized that the causal interpretations of their results rest on the tenability of the ''implicit potentially causal model'' that is depicted in this article as Figure 1a. They then noted that a ''more realistic model has an unobserved component, early choice,'' and they then offered a causal graph, for discussion, that is similar to what is presented in Figure 2a in this article.

This augmented causal graph in Figure 2a includes a third path, $Class \rightarrow AD \rightarrow P \rightarrow G$. For Erikson et al. (2005), *AD* is an unobserved anticipatory decision to go on to the next level of education, but one that is taken some years prior to the actual educational transition. As a result, the anticipatory decision structures P in the run up to the educational transition, which they note will lead to an underestimate of the secondary effect ($Class \rightarrow G$) if the anticipatory decision is unobserved. Their point, expressed with Figure 2a, is that the secondary effect of stratification now operates through two causal pathways, $Class \rightarrow G$ and $Class \rightarrow AD \rightarrow P \rightarrow G$. If *AD* is unobserved, then the pathway $Class \rightarrow AD \rightarrow P \rightarrow G$ is absorbed into $Class \rightarrow P \rightarrow G$. The effect of *Class* on P via *AD* is then implicitly attributed to the direct effect of *Class* on P, which shifts some of the true secondary effect to the estimated primary effect. As a consequence, by their reasoning, the estimate of the secondary effect is downwardly biased and the estimate of the primary effect is upwardly biased. Jackson et al. (2007) and Erikson and Rudolphi (2010) offer similar interpretations.

This conclusion is not necessarily incorrect, but it is certainly sanguine. Based on the literature on college entry in the United States, I would argue instead that the causal graph in Figure 2a is too simple as well.[10] I would include exogenous variables that determine *Class*, U, and G and then also allow the intermediate unobserved variable U to be more inclusive than just an anticipatory decision or an unobserved choice-based mechanism.

Accordingly, consider the more general causal graph presented in Figure 2b, which is an elaboration of the causal graph in Figure 1b after Assumptions 2 and 3 are abandoned in relation to *Class* and U. In addition to the two causal pathways $Class \rightarrow U \rightarrow G$ and $Class \rightarrow P \rightarrow G$ carried forward from Figure 1b, the causal graph in Figure 2b now includes additional paths through the variables U and X. The first, $Class \rightarrow U \rightarrow P \rightarrow G$, is analogous to the path $Class \rightarrow AD \rightarrow P \rightarrow G$ proposed by Erikson et al. (2005). Five additional backdoor paths from $Class$ to G also have arisen: $Class \leftarrow X \rightarrow G$, $Class \leftarrow X \rightarrow P \rightarrow G$, $Class \leftarrow X \rightarrow U \rightarrow G$, $Class \leftarrow X \rightarrow U \rightarrow P \rightarrow G$, and $Class \leftarrow X \rightarrow P \leftarrow U \rightarrow G$. These five backdoor paths through X generate supplemental dependence among $Class$, U, and G.[11]

Consider some of the variables that the literature suggests belong in the background cause X in relation to some of the variables presumed to belong in U. It is often argued that, for many educational transitions in many contexts, race serves as a variable in X while perceptions about the opportunity structure serve as a variable in U. Accordingly, race would be a common cause of both $Class$ and beliefs about the opportunity structure that are now embedded in U. The backdoor paths $Class \leftarrow X \rightarrow U \rightarrow G$ and $Class \leftarrow X \rightarrow U \rightarrow P \rightarrow G$ then follow from the positions in the literature that perceptions of the opportunity structure, now in U, determine G directly and also determine G indirectly via prior preparation P. Furthermore, the backdoor path $Class \leftarrow X \rightarrow P \rightarrow G$ is supported by the literature that argues for racial bias in performance evaluations, whether generated by tests of dubious quality or by teachers with biased expectations.

Although race is perhaps the most obvious variable to include in X, at least for educational transitions in the United States, other variables that signify categorical distinctions may also belong in X, such as place of residence in contexts where neighborhood disadvantage leads to forms of social isolation. Moreover, these variables in X then entail mechanistic structural variables in U that carry their causes on to college entry. These new variables in U would include institutional features of schooling, such as between-school differences in instructional quality and within-school curriculum differentiation. These variables in U then must have effects on both prior performance, P, as well as direct effects on college entry, G. Such causal pathways cannot be considered mechanistic elaborations of Boudon's choice-theoretic conception of the secondary effects of stratification. They are instead a separate component of the net association between class and college entry that is best attributed to a broad structural interpretation.

Now consider the even more subtle consequences of abandoning Assumption 4, as is the case for the causal graph presented in Figure 2c that includes backdoor paths from U and P to G through a new set of variables in V. For context, suppose that V includes a set of institutional features of schooling, such as between-school differences in instructional quality, that do not have an unconditional association with *Class*, as might be the case if they were exogenously determined by idiosyncratic policy differences across school districts in the United States. As such, V structures the graph in Figure 2c by generating supplemental associations between U, P, and G that are, at first glance, unrelated to *Class*, and presumably then to the primary and secondary effects generated by class.

To see the complications that now arise, note first that this causal graph includes two causal pathways $Class \rightarrow U \rightarrow G$ and $Class \rightarrow P \rightarrow G$ carried forward from Figure 1b as well as the path $Class \rightarrow U \rightarrow P \rightarrow G$ that is akin to the additional path suggested by Erikson et al. (2005) for their ''more realistic model.'' However, there are now three new paths between *Class* and G that are generated by the common-cause subpaths $U \leftarrow V \rightarrow G$ and $P \leftarrow V \rightarrow G$. These three paths are $Class \rightarrow U \leftarrow V \rightarrow G$, $Class \rightarrow P \leftarrow V \rightarrow G$, and $Class \rightarrow U \rightarrow P \leftarrow V \rightarrow G$. Note that these paths are not backdoor paths from *Class* to G. None includes a common cause that determines *Class*, as was the case with X in Figure 2b. Moreover, each of these three paths is blocked because P and U serve as colliders on these paths, preventing the paths from generating a supplemental association between *Class* and G.[12] In this sense, the dependencies imposed on U, P, and G by the new variable V may seem innocuous.

In practice, however, problems immediately arise from the fact that U and V are unobserved. If one follows the traditional strategy of conditioning on P to partial out the primary effect to develop an estimate of the secondary effect that operates by way of U, then the residual association between *Class* and G reflects no less than three net associations that cannot be separated from each other. The first, $Class \rightarrow U \rightarrow G$, is the secondary effect as defined in this literature. The other two associations arise from formerly blocked paths that connect *Class* and G: $Class \rightarrow P \leftarrow V \rightarrow G$ and $Class \rightarrow U \rightarrow P \leftarrow V \rightarrow G$. Because P is a collider on these two paths, conditioning on P opens these paths by generating new associations between *Class* and G (via V and U) within strata defined by P. These associations then contribute to the net association between *Class* and G in ways that cannot be attributed to choice-based secondary effects.

Overall, this consideration of the consequences of variables X and V, as well as the many causal pathways that they introduce, should be sobering for those who might consider adopting these methods for educational transitions in the United States. For completeness, Figure 2d presents a full causal graph that combines the graphs in Figures 2b and 2c. The very simple causal graphs in Figure 1 hardly seem to give a solid foundation to the causal

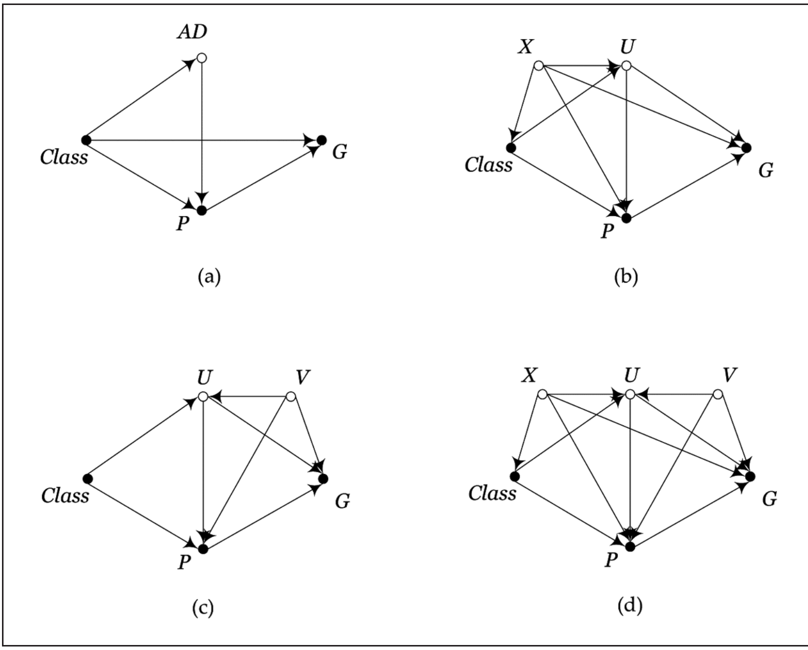

Figure 2. Alternative causal models for the putative primary and secondary effects of stratification

interpretations typical of current modeling practices, were they to be offered for college entry patterns in the United States.

Using standard methods would be tantamount to replacing the many pathways through X and V with a simple stand-in path $Class \rightarrow U \rightarrow G$. In this case, U would be regarded as an unobserved intervening variable that serves no purpose other than transmitting the net direct effect of *Class* to G. Treating U in this way would probably lead to overestimation of the secondary effects of stratification, as structural effects would then be let in through the backdoor and attributed by assumption to Boudon's choice-based secondary effects. At the same time, the estimates of the primary effects, $Class \rightarrow P \rightarrow G$, would then be confounded by similar additional pathways, although in less clear ways than for the simple case of a sole anticipatory decision considered by Erikson et al. (2005). When interpreted as estimates of causal effects, the biases in estimates of primary and secondary effects from standard estimation practices would therefore contain many countervailing sources, making a priori claims about the direction of total bias impossible.

A Simpler Associational Model

When one steps back from strong causal interpretations, a natural question arises: Should an even simpler descriptive approach to estimation and interpretation be embraced? Two points support this position. First, class-referenced differences based on Table 3 invite unwarranted causal interpretations from readers, both because of the influence of past analyses and because the differences are ineluctably tied to particular observed individuals in particular classes. Efforts to avoid ''what-if'' counterfactual language when considering such differences are unnatural, and such language is used throughout Erikson et al. (2005), Jackson et al. (2007), Erikson and Rudolphi (2010), and other articles following in this tradition. Second, the differences estimated (e.g., in equations (8) and (10)) cannot be easily mapped back to classes themselves, instead only in pairs for directional movement between classes. When more than three classes are considered, the interpretive advantage of having a smaller number of class-specific descriptive parameters grows.[13]

From Marginal Means to Associational Effects

A simpler associational analysis strategy is presented in Table 4. For three separate models, the table reports marginal means, which are the mean college entry rates for the full sample *as if* the full sample transitioned through the college entry regime of a particular class, under the specification of a particular model. In this subsection, I interpret only the models that are labeled the ''null model'' and the ''standard model.''

Consider first the null model of college entry where the sole predictors in the logit are two dummy variables for class (i.e., Model 1 from Table 2). The marginal mean entry rate is the predicted college entry rate for each class implied by the model if all members of the sample were in that class. Since the model includes no additional information beyond class, the marginal mean entry rates are equal to the observed college entry rates for the three classes.

Now consider the standard model of college entry for classes, where the standardized measure of performance in the senior year of high school and its interactions with class are added to the null model (which is then Model 3 from Table 2). For this model, the marginal means are calculated in the same way as for the null model, but now the model incorporates the performance characteristics of all individuals as well as the class-specific relationship between performance and college entry. Accordingly, the marginal mean for each class is the predicted college entry rate of *all* individuals in the sample if they kept their performance levels but transitioned through the relevant college-entry regime of the referenced class.

Associational primary effects can be calculated using the marginal means from the null and standard models. For the salariat, the marginal mean for college entry under the standard model is .476.[14] Rather than 54.1 percent of salariat entering college, as in the null model, the standard model predicts that only 47.6 percent of the salariat would enter college if the salariat had the performance distribution of the full sample rather than its own observed performance distribution. The difference between these two numbers is therefore a meaningful measure of an associational primary effect because it reflects the difference in the college entry rate that can be attributed to the particular performance distribution acquired by the class. Under this logic, class-specific associational primary effects can be calculated by subtracting marginal means of the standard model from the marginal means of the null model:

$$\begin{aligned} \text{Associational } PE_S &= .541 - .476 = .065, \\ \text{Associational } PE_I &= .406 - .442 = -.036, \\ \text{Associational } PE_W &= .285 - .367 = -.082. \end{aligned} \tag{14}$$

Associational secondary effects can then be formed by taking differences between the marginal means from the standard model, since the remaining class differences reflect net direct associations with class after the model adjusts for performance differences. Associational secondary effects for class comparisons are then,

$$\begin{aligned} \text{Associational } SE_{S \text{ to } W} &= .367 - .476 = -.109, \\ \text{Associational } SE_{W \text{ to } S} &= .476 - .367 = .109, \\ \text{Associational } SE_{S \text{ to } I} &= .442 - .476 = -.034, \\ \text{Associational } SE_{I \text{ to } S} &= .479 - .442 = .034, \\ \text{Associational } SE_{I \text{ to } W} &= .367 - .442 = -.075, \\ \text{Associational } SE_{W \text{ to } I} &= .442 - .367 = .075. \end{aligned} \tag{15}$$

Table 4. Predicted Marginal Means for College Entry by Class Origins

	Null model (Model 1 from Table 2) Marginal mean entry rate	Standard model (Model 3 from Table 2) Marginal mean entry rate	Augmented model: Standard model + expectations, immediate plans, and significant others' influence Marginal mean entry rate
Salariat	0.541 (0.011)	0.476 (0.010)	0.455 (0.010)
Intermediate	0.406 (0.014)	0.442 (0.014)	0.445 (0.012)
Working	0.285 (0.011)	0.367 (0.012)	0.403 (0.012)

Source: See note to Table 1.
N = 11,400 for all models. Standard errors in parentheses.

Table 5. Means and Standard Deviations of Additional Variables Utilized for the Augmented Model

	Salariat class		Intermediate class		Working class		Full sample	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Added for "augmented model"								
Educational expectations (years)	16.823	1.968	16.234	2.094	15.896	2.167	16.448	2.091
Immediate plans (ref: No further schooling)								
Go to four-year college	0.740		0.638		0.550		0.667	
Go to two-year college	0.203		0.259		0.313		0.245	
Mother's expectations (years)	16.765	2.088	16.324	2.331	16.005	2.476	16.465	2.275
Father's expectations (years)	16.734	2.164	16.346	2.299	15.867	2.448	16.415	2.303
Significant others' influence	0.746	0.319	0.706	0.327	0.638	0.364	0.708	0.337

Source: See note to Table 1.
N = 11,400. A modest amount of data were imputed with best subset regression methods.

The symmetry across pairs of these associational secondary effects is generated by the method of calculating marginal means, where the entire sample is moved through each of the three class-specific college entry regimes. All of the underlying individual-level associations are averaged over the full distribution of performance, not the class-specific distributions of performance for the differences presented earlier in equations (2) through (7). As a result, the associational secondary effects for class comparisons differ only in sign within pairs, and these signs indicate only the direction of movement. Differences between pairs are attributable solely to the class-specific index functions implied by the estimated coefficients of the logit model.

Finally, if one wishes to then have class-specific associational secondary effects, which can be compared to class-specific primary effects, then these can be formed by taking the difference between the marginal mean for each class from the standard model and appropriately chosen reference values for each class. In descriptive analysis, where the goal is simply to summarize the data in an efficient and organized way, alternative types of reference values could be chosen. I would argue that it makes sense to favor one more than others: class-sized-weighted averages of the marginal means of the other classes, as in,

$$\begin{aligned}
\text{Associational } SE_S &= .476 - \left[\left(\frac{.197}{.197+.300}\right)(.442)+\left(\frac{.300}{.197+.300}\right)(.367)\right] \\
&= .476 - .397 = .079, \\
\text{Associational } SE_I &= .442 - \left[\left(\frac{.503}{.503+.300}\right)(.476)+\left(\frac{.300}{.503+.300}\right)(.367)\right] \\
&= .442 - .435 = .007, \\
\text{Associational } SE_W &= .367 - \left[\left(\frac{.503}{.503+.197}\right)(.476)+\left(\frac{.197}{.197+.503}\right)(.442)\right] \\
&= .367 - .466 = -.099. \qquad (16)
\end{aligned}$$

where, as shown in Table 1, .503, .197, and .300 are the probability masses associated with the relative sizes of the three classes. With these chosen reference values, the associational secondary effect of a class is then the predicted probability of college entry if all members of the sample were in that class minus the predicted probability of college entry if no member of the sample were in that class but instead were in the alternative classes in proportion to the observed sizes of those alternative classes.

A benefit of having class-specific primary and secondary effects, as in equations (14) and (16), is that relative sizes of these effects can then be constructed. For the salariat, the associational primary effect is about the same size as the associational secondary effect, at .065 and .079, respectively. For the intermediate class, the associational primary effect is smaller and negative at –.036, and the associational secondary effect is close to 0 at .007. For the working class, the associational primary effect and the associational secondary effects are both relatively large and negative, at –.082 and –.099, respectively. In combination, the associational primary and secondary effects of the working class are more negative than the combined effects are positive for the salariat.

Of course, offering interpretations labeled *associational* primary and secondary effects is itself awkward and unnatural. This will be a legacy of this research tradition as long as the word *effect* continues to be used in a general way. Semantic clarity would be well served by using the word *effect* only for contrasts that have a strong causal warrant, but it is perhaps too late for this particular set of models.

Toward Causal Models of Primary and Secondary Effects

Having laid out a simple method for estimating associational primary and secondary effects in the last section, this section has two goals. First, a demonstration is offered on how to use standard conditioning techniques to attempt to develop more credible estimates of causal primary and secondary effects in the United States. The chief constraints on this strategy, as is shown, are the dearth of data on students' choice behavior as well as the additional assumptions that must be introduced (and which are unlikely to be regarded as uncontroversial by researchers with opposing theoretical orientations). Second, an explanation is offered for why research on primary and secondary effects will be especially valuable when patterns of underlying heterogeneity can be effectively analyzed. These patterns are central to an understanding of the consequences of hypothetical interventions on the cost of higher education, which has been of particular interest in both the primary and secondary effects literature in Europe and in the college entry literature in the United States.

Evaluating Alternative Strategies for Causal Identification

The causal identification issues discussed earlier are challenging, but they are not fundamentally insurmountable. The first step to address them is to understand the predictive power of other variables that fall along the causal pathways that emanate from social class. This is part and parcel of an attempt to understand whether backdoor associations that confound estimates of genuine choice-based secondary effects can be partialed out of the analysis.

The final column of Table 4 presents additional marginal means based on a logit model that includes additional predictors alongside and interacted with class. For this augmented model, six additional variables are added to the standard model, resulting in a logit model for college entry predicted from performance, educational expectations, immediate plans, and significant others' influence. As summarized in Table 5, the three educational expectations variables are for high school seniors' own expected years of education as well as the expectations maintained for them by their mothers and fathers. The two immediate plans variables signify whether seniors expect to attend either a four-year college or a two-year college immediately following high school (in comparison to either a vocational school or no further education). The significant others' influence variable is a composite of whether a respondent reports, as of the senior year, that his or her parents, peers, and teachers expect that the respondent will go to college after high school. All six variables are standard predictors of college entry in sociology from the status attainment tradition (see Sewell, Haller, and Portes 1969 and citations to it).

For the augmented model, class differences in the marginal means decline substantially, with, for example, the marginal mean for the salariat decreasing from .476 to .455, while the marginal mean for the working class increases from .376 to .403. How should one interpret the narrowing of these marginal means? Does the augmented model suggest that one can pursue estimation of causal secondary effects after conditioning on these additional variables? One could certainly substitute the marginal means for the augmented model for those derived from the standard model into the same difference-based estimators of *associational* primary and secondary effects in equations (14) through (16). Associational interpretations can then be offered.

But can one go further to (a) generate predicted values for counterfactual movement, as in Table 3, and then (b) offer causal primary and secondary effects by recalculating primary and secondary effects with analogous estimators to those used in equations (8) and (10) for the standard model?

The answer to this question requires taking a position on the location of the variables for expectations, immediate plans, and significant others' influence in the causal model in Figure 2d. Because these variables are used in empirical models derived from many alternative theoretical traditions (see Morgan 2005), one could argue for at least two positions for models of college entry in the United States.

The most common interpretation in the sociology of education would follow directly from the status attainment tradition. Expectations and plans would be considered characteristics of individuals that are shaped by socialization processes (measured by significant others' influence), which then crystallize in early adolescence and remain stable through early adulthood. From this perspective, educational expectations, plans, and significant others' influence are components of U that are determined by both $Class$ and X (and perhaps with significant others' influence as a mediating variable that transmits the effects of $Class$ and X to expectations and plans). Moreover, because these variables are not explicit components of a choice process, they are not the choice component of U that is thought to generate the causal secondary effects suggested by Boudon.

Under this interpretation of where the additional variables belong in the causal graph, one could then reestimate the predicted probabilities in Table 3 and then assert that resulting causal secondary effects akin to those defined in equations (2) through (7) are closer to the true causal secondary effects than those offered in equation (8). One could then calculate estimates of refined causal primary effects, as in equations (9) and (10). At a minimum, the resulting causal secondary effects could then be regarded as upper-bound estimates of the true secondary effects, since there may still be some additional backdoor confounding produced by other causal pathways through X and U (and perhaps as induced through V when conditioning on P).

A permissible alternative interpretation would suggest that this conclusion is mistaken because expectations, plans, and significant others' influence should be considered indicators of the anticipatory decision that Erikson et al. (2005) argue may be an essential component of secondary effects. By conditioning out these effects, the augmented model artificially biases the secondary effect estimates toward zero. Under this second interpretation, the prior secondary effect estimates presented earlier in equation (8) based on Table 2 would be preferred, and might even be regarded as lower-bound estimates for the reasons stated earlier in the discussion of Figure 2a.[15]

The lesson that is suggested by these results is therefore quite sobering for the prospects of standard conditioning strategies: It is unlikely that efforts to model backdoor paths with existing national data in the United States (or perhaps in any country) will deliver clear identification of net direct effects that can then be imbued with choice-based causal interpretations. For most data with which these models have been estimated in the past, rich sets of adjustment variables are not available at all. Even when they are available, as is the case for the ELS data analyzed in this article, it is unclear how one should interpret direct effects net of variables whose measurement is motivated by other theoretical traditions. As a result, scholars are unlikely to agree that any particular set of available conditioning variables will be effective at removing the desired portion of the association between class and college entry for each of the classes, leaving a net effect of class that is widely agreed can be interpreted as the sole product of choice.

Because the data that currently exist are ill-suited for efforts to purge non-choice-based contamination from educational choice processes, it would seem that a necessary step in an analysis of causal primary and secondary effects is to directly model the choice mechanism that constitutes the genuine secondary effect of interest. If what is desired in this research tradition is an estimate of the number of students in each class who choose not to enter postsecondary education even though they are prepared to do so, then research needs to shift toward models of the causal pathways that are embedded in the implicitly assumed choice processes, as in Gabay-Egozi, Shavit, and Yaish (2010) and Stocké (2007) for regional samples in Israel and Germany, respectively. There is much sociology to be completed in this effort. We still have a poor understanding of how information about costs and benefits is differentially available and differentially utilized by students and parents from alternative locations in the structure of social advantage in the United States (see Dominitz and Manski 1996; Avery and Kane 2004; Grodsky and Jones 2007).

Implied Heterogeneity

One reason to pursue causal identification within a primary and secondary effects framework is that such identification will permit the examination of important heterogeneity of causal effects, as discussed in this subsection.

Recall that when discussing the estimates of causal secondary effects, based on the predicted values from Table 3 and presented in equation (8), the magnitudes of these effects were not equal within pairs. For example, the value for $SE_{S\text{ to }W}$ was –.118 while the value for $SE_{W\text{ to }S}$ was .095.

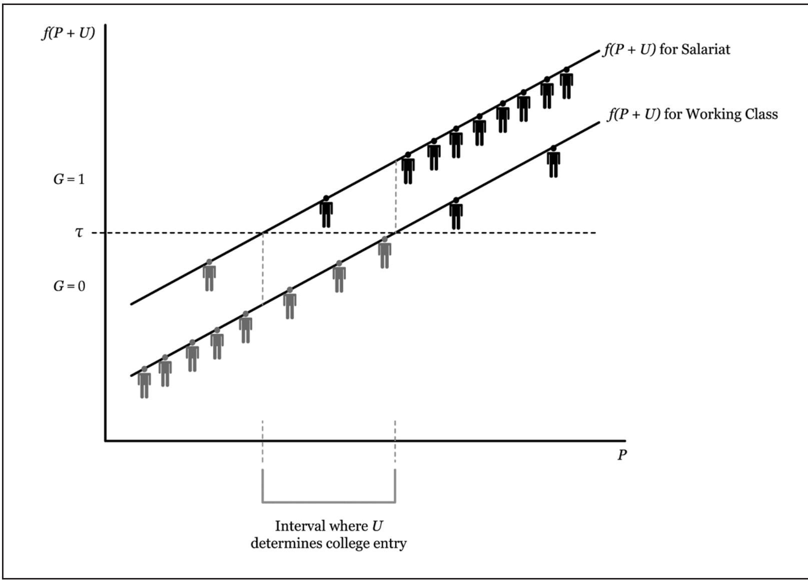

Figure 3. A simplified model of the relationships between performance (*P*), an unobserved variable (*U*), and college entry (*G*) for the salariat and the working class

As stated earlier, such differences are attributable to the definitions of the estimators in equations (2) through (7), which weight individual-level effects over class-specific distributions of performance. This explanation then begs a question: Why should this matter?

Recall the formalization of the standard model presented earlier in equation (1). For this model, the parameter τ is a fixed threshold for enrollment, based in the choice-theoretic tradition on the relative costs and benefits of education. Now consider a graphical depiction of this basic model for the salariat and the working class, presented in Figure 3, that strips away the randomness that occurs in real contexts, assumes linearity of $f(.)$ that is positive in P and U, assumes that U varies by class by a structural constant, and assumes that U does not vary directly with P beyond the constant class difference. In this scenario, the uppermost line in Figure 3 depicts for the salariat only the expected relationship between $f(P+U)$ and college entry. The parallel line just below it is the same relationship for the working class, with the space between the two lines representing the assumed structural constant difference in U between the two classes.

Note how students are distributed along these lines, assuming each class has 10 students within it. Because of the primary effects, the average level of performance P is higher for the salariat. As a result, students from the salariat are, on average, farther to the right-hand side in the figure.

Now consider college entry patterns. Students shaded black enter college because they are above the threshold τ, and students shaded gray do not enter college because they are below the threshold τ. Accordingly, 9 of the 10 students from the salariat enter college, and only 2 of the 10 students from the working class enter college. These differences in college entry are produced solely by class differences in both P and U.

The heart of the explanation for the lack of equality within pairs of secondary effects can be found in the interval around the threshold τ, delineated by dashed vertical gray lines. Note first that a counterfactual shift that moves students from one class to the other class is equivalent in this simplified model to moving the students vertically from their line to the line of the alternative class. Under this operation, consider first the students from the salariat to the upper right of the delineated interval. If all eight of these students fell straight down to the line for the working class, they would still remain above the threshold and hence still enter college. In effect, each student would swap her or his prior value of U for the lower value of U that characterizes working-class students with the same level of performance. For these students, the swap would be inconsequential. Similarly, the student in the salariat who is already to the lower-left of the delineated interval would still not go to college if she or he fell to the working class line. Only the sole student from the salariat who is within the delineated interval would cross the threshold in the counterfactual scenario. Thus, applying the expression in equation (2), we have $.8 - .9 = -.1$ for $SE_{S \text{ to } W}$.

An equivalent set of counterfactual moves from the working class line to the salariat line would result in three students crossing the threshold for college entry. The two students to the upper right of the interval would still enter college after shifting to the salariat line, and the five students to the lower left of the interval would still not enter college after an equivalent shift. However, each of the three students within the interval would cross the threshold when moving to the salariat. Thus, applying the expression in equation (3), we have $.5 - .2 = .3$ for $SE_{W \text{ to } S}$. Overall, for this simple example, $SE_{S \text{ to } W}$ does not equal $SE_{W \text{ to } S}$, as was suggested in the discussion earlier of equations (2) and (3).

Now consider what would happen if, under the same exact scenario, τ increased as in Figure 4. Fewer student would enter college, and the relative magnitude of $SE_{S \text{ to } W}$ and $SE_{W \text{ to } S}$ would flip signs. The value for $SE_{S \text{ to } W}$

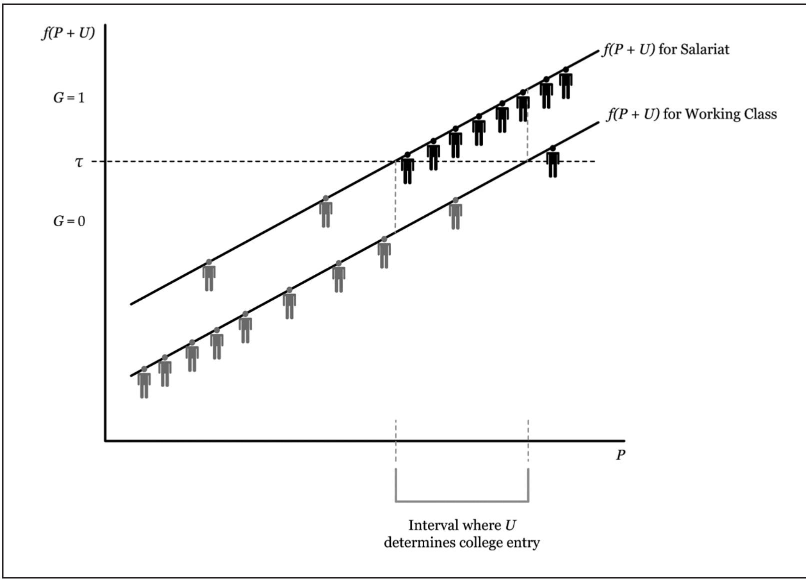

Figure 4. The same model as in Figure 3 but with a higher threshold for college entry

would be $.2 - .8 = -.6$, while $SE_{W \text{ to } S}$ would be $.2 - .1 = .1$. The result would be $SE_{S \text{ to } W} + SE_{W \text{ to } S} = -.5$, not .2 as for Figure 3.

As Figures 3 and 4 suggest, one could build a general model for how these patterns of heterogeneity would unfold under alternative fixed thresholds and under alternative specific functions for $f(.)$ in $f(P, U)$. Such a model will not be pursued here because the goal of this section is more modest. I only wish to make two simple points. First, if primary effects exist, then the distribution of P will vary by class. This alone will ensure that different proportions of the classes will almost always fall on either side of the threshold that determines who moves on to the next higher level of education (i.e., except for a single threshold value, unlikely ever to obtain, that would perfectly balance the asymmetry conditional on U). Second, because primary effects will necessarily generate such an asymmetry, directional shifts will not generally be equal in expectation.

These asymmetries are meaningful and highly relevant if, as recent scholarship on primary and secondary effects has contended, this scholarship is meant to be relevant to the evaluation of policies that promote access to higher education. Policy interventions that shift values for U will have differential effects for different classes, and these are the expected effects of informational changes that were advocated by Jackson et al. (2007) and Erikson and Rudolphi (2010) in Great Britain and Sweden, respectively. Likewise, shifts in the costs of higher education, which is another common policy proposal, would shift τ downward, inducing more students to enter college, but differentially so for different classes depending on the proportions of the classes in the interval across which the threshold traverses. All of these effects will vary with the strength of primary effects that structure P differentially for classes. To the extent that primary effects are powerful, classes will approach the college entry decision from very different profiles of observed characteristics, and the expected effects of interventions will then vary accordingly.

Overall, the current literature on causal primary and secondary effects does not attempt to account for these heterogeneous effects, even though the framework is ideally suited for pursuing their estimation. This is an area of as-yet-unfulfilled promise for these models, and it will be greatly aided if causal identification of the secondary effects can be achieved, as outlined in the prior subsection.

Conclusion

Boudon's model of primary and secondary effects of stratification on educational transitions is a promising conceptual framework for separating the effects of students' own choices from the standard baseline family background and preparation effects that are known to also determine patterns of educational attainment in the United States. This article has demonstrated that the empirical methods that have been used recently in the sociology of education to examine educational transitions in many European societies, and which are inspired by Boudon's distinction, do not yield sufficiently strong conclusions about the sizes of these causal effects in the United States, and perhaps elsewhere. Instead, the models yield contrasts that can be misleading because they invite unwarranted causal interpretations.

As an alternative, this article proposes a simple model of associational analysis that is less prone to unwarranted causal interpretation and also yields straightforward descriptive decompositions of estimated primary and secondary effects (albeit ones that permit only associational interpretations). In addition, the article demonstrates that backdoor conditioning with available national data is unlikely to lead to estimates of primary and secondary effects that have clear and sufficiently convincing causal interpretations.

It is likely that direct examination of the choice-based mechanism itself is the most fruitful way forward for estimating secondary effects. But, here, we await future data collection, since no known national data appear able to furnish the desired information for deeper modeling in any country. Accordingly, this article can be read as yet another appeal for better data on students' beliefs about their alternative futures and on their subsequent choices, for without them only marginal progress in the causal modeling of educational transitions is likely.

Appendix

Additional Details of the ELS Data

Data were drawn from the 2002 through 2006 waves of the Education Longitudinal Study (ELS), which is a nationally representative sample of students in public and private high schools collected by the National Center for Education Statistics (NCES) of the U.S. Department of Education. Respondents were sampled as enrolled high school sophomores in 2002, and follow up surveys were also conducted in both 2004 and 2006. The analysis sample for this article was first restricted to respondents who participated in all three waves of the survey, using a weight to account for sample attrition between the sophomore and senior years. From among these students, high school graduates were then selected for analysis in the college entry models (where high school graduates were de.ned as those who obtained high school diplomas, thereby excluding GED recipients). The resulting analysis sample includes 11,400 respondents.

College entry is defined strictly as having made the transition to a 4-year college within 6 months of high school graduation, regardless of whether high school graduation was delayed. Of the 11,400 students who had complete data for college entry and social class, 44.3 percent entered a four-year college within 6-months of high school graduation.

Class and SES. To construct EG classes, broad occupational categories of fathers and mothers (or male and female guardians) were coded into the standard nine-category EG schema. Then, these nine-category classes were collapsed into the three tiers of salariat, intermediate, and working, after which father's and mother's class position were collapsed into an overall family measure. In this family class coding, father's class was slightly privileged, such that (1) families were coded as salariat if either parent had a salariat occupation, (2) families were coded as intermediate class if the father had an intermediate class occupation but the mother had an 35 intermediate class or working class occupation, and (3) families were coded as working class if the father had a working class occupation but the mother had either an intermediate class occupation or a working class occupation. This coding provides some consistency with definitions of class that use only father's occupation but also gives some weight to mother's class position as well.

The SES variable was constructed by the data distributors, which is a linear composite of mother's education, father's education, mother's occupational prestige, father's occupational prestige, and family income. The underlying data were drawn from both parent-provided information and student-provided information, as well as some imputed information developed by the data distributors.

Performance. Respondents were administered achievement tests in both the 2002 base year and the 2004 first follow-up waves. Base-year respondents were administered math and reading achievement tests, but only math tests were administered in 2004. The 2004 academic performance measure, which is used in this article, is a standardized composite of 2004 math and 2002 reading IRT-estimated scores. It should be noted that it would have been preferable to have a 2004 measure of academic performance that included 2004 results from a reading test, but this was not available. That such a measure is absent, however, should not be very consequential. One of the reasons that the U.S. Department of Education employs an updated math test for the 2004 ELS but not an updated reading test is that prior work shows considerable stability in relative rankings throughout high school for reading achievement. In contrast, math achievement is more dynamic, mostly because of greater curriculum differentiation across students.

College Entry Models with SES Groups

Given the strong preference for modeling educational attainment in the United States as a function of SES rather than class, Table A1 offers analogous results to those in Table 2 but with two indicator variables for SES groups rather than classes. (Models were also estimated with quartiles and quintiles of SES. These alternative models differed very little, merely creating more finely graded results.)

A comparison of Chi-squared test statistics between Tables 2 and A1 suggests that SES is a stronger predictor than class. In one sense, this is obviously true; the same number of parameters is .t to the data, and the

Chi-squared value is higher for Model 1 in Table A1. Moreover, Table 1 already showed a stronger baseline relationship between SES and college entry (i.e., entry rates of 54.1, 40.6; and 28.5 for the three classes but more widely dispersed entry rates of 65.6, 40.2, and 23.9 for SES tertiles). Yet, this comparison is somewhat unfair to class as a predictor, given that SES is a composite of five underlying components.

Table A1. Logit Models for College Entry Among High School Graduates, with SES Groups as An Alternative to Classes

	Model 1	Model 2	Model 3
Constant	.646	.119	.044
	(.051)	(.053)	(.058)
Socioeconomic status:			
× Middle tertile	−1.043	−.655	−.563
	(.060)	(.065)	(.071)
× Lower tertile	−1.807	−1.043	−.972
	(.068)	(.073)	(.075)
Performance		1.147	1.331
		(.039)	(.060)
× Middle tertile			−.268
			(.080)
× Lower tertile			−.296
			(.082)
Wald Chi-Squared	725.0	1,167.2	1,229.8
N	11,400	11,400	11,400

Source: See Table 1.
Notes: See Table 2.

Tables A2 and A3 are analogous to Tables 3 and 4. Similar broad patterns of predicted college entry rates prevail for SES groups, though the differences are slightly more substantial than for the class categories. Overall, it seems reasonable to conclude that, if one switched away from classes to SES groups, as would be the common practice in studies of college entry in the United States, some differences in the strength of patterns would result. Yet, the overall conclusions of the effort would be rather insensitive to such a change, suggesting that the reliance of Erikson et al. (2005), Jackson et al. (2007), and others on classes has not created too great of a disjuncture with the literature on college entry in the United States.

Table A2. College Entry Rates by SES Group, Estimated from the Models in Table A1

	Model 2 (from Table A1) What-If SES Transition Probability		
	Higher SES	Middle SES	Lower SES
Higher SES	.656	.529	.449
	(.011)	(.012)	(.014)
Middle SES	.533	.402	.329
	(.012)	(.011)	(.012)
Lower SES	.422	.301	.239
	(.012)	(.010)	(.010)
	Model 3 (from Table A1) What-If SES Transition Probability		
	Higher SES	Middle SES	Lower SES
Higher SES	.656	.522	.431
	(.011)	(.013)	(.016)
Middle SES	.519	.402	.321
	(.013)	(.011)	(.012)
Lower SES	.397	.306	.239
	(.014)	(.011)	(.010)

Source: See Table 1.
Notes: See Table 2.

Table A3. Predicted Marginal Means for College Entry by SES Groups

	Null Model (Model 1 from Table A1)	Standard Model (Model 3 from Table A1)	Augmented Model Standard Model + Expectations, Immediate Plans, and Significant Others' Influence
	Marginal Mean Entry Rate	Marginal Mean Entry Rate	Marginal Mean Entry Rate
Higher SES	.656	.530	.481
	(.011)	(.012)	(.011)
Middle SES	.402	.415	.428
	(.011)	(.011)	(.011)
Lower SES	.239	.335	.383
	(.010)	(.012)	(.013)

Source: See Table 1.
Notes: See Table 4.

Acknowledgments

I thank Theo Leenman for data analysis assistance. For their helpful and stimulating comments, I thank Michelle Jackson, Trey Spiller, Jenny Todd, CIQLE seminar participants at Yale, and sociology seminar participants at Nuffield College, Oxford.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. In addition to Erikson and Jonsson (1996), the renewed interest in Boudon's primary and secondary effects is also the result of the theoretical models proposed by Goldthorpe (1996) and Breen and Goldthorpe (1997). Boudon's book has been less influential in the United States, in part because of an influential review essay (Hauser 1976). Hauser argued that Boudon's assumptions about class differences in the relative costs and benefits of educational attainment were incoherent, and he demonstrated that Boudon's reanalysis of past data on educational plans in the United States was largely incorrect.
2. These conclusions were echoed most recently by Erikson and Rudolphi (2010:301), who argue, after estimating similar models with data from Sweden, that "it may be possible to motivate children [from the working class] to choose secondary training by giving full information on the benefits . . . of higher education." For other recent work, see Kloosterman et al. (2009), Jackson (2010), and Jonsson and Rudolphi (2011).
3. Erikson et al. (2005:9732-33) present graphs that they label "path diagrams." They label their graph in Figure 2(a) (which is equivalent to Figure 1a in the present article) as "the implicit potentially causal model" that underlies their empirical analysis. They then offer "a more realistic model" in their Figure 2(b), which will be discussed later in the present article.
4. One could elaborate this model by decomposing U into its components. One could also add random noise, either as an additional input in $f(.)$ or as a component of U. Such refinements are not necessary at this stage of the argument and would not alter the reasoning presented here.
5. From a rational choice perspective, τ can be regarded as a function of exogenous costs and benefits of college enrollment (see Manski 1989; Breen and Goldthorpe 1997; Morgan 2005). Any variation across individuals in the

relative attractiveness of a college education can then be embedded in U and hence treated as an individual-level determinant of college entry.

6. Here, the implicit assumption is that inequalities such as $E[P|S=1] \geq E[P|I=1] \geq E[P|W=1]$ are causal. This could be made more explicit using the $do(.)$ operator of Pearl (2009), as in $E[P|do(S=1)] \geq E[P|do(I=1)] \geq E[P|do(W=1)]$.
7. Moreover, background causes may exist that give rise to a distribution for G, but these background causes do not need to be explicitly represented in the causal graph that elaborates the effects of *Class* on G because Assumptions 3 and 4 stipulate that any such additional causes will be independent of *Class*.
8. In this tradition of causal analysis, such simplifications are permissible because we often think of causal arrows as black boxes that indicate the existence of an unobserved mechanism that, in theory, could be observed.
9. The estimated inclusion probabilities were extracted from a baseline logit that included 26 parameters for gender, race, region, urbanicity, school sector, family structure, parents' education, parents' occupation, family income, students' own educational expectations, parents' expectations, and significant others' influence.
10. In fact, Erikson et al. (2005) also allow for the direct effect of AD on G, and that creates unstated complications for the conclusion they offer in the same article that the anticipatory decision necessarily biases downward estimates of the secondary effect.
11. Note that, for simplicity, backdoor paths where P causes U, such as $Class \rightarrow P \rightarrow U \rightarrow G$ and $Class \leftarrow X \rightarrow P \rightarrow U \rightarrow G$, are not introduced in this causal graph. If they were included, the same basic conclusions would follow from this identification analysis, but the graph and the explanation of the conclusions would be more complicated.
12. See Pearl (2009) for a full discussion of the complications generated by collider variables. See Morgan and Winship (2007, chap. 3) for a gentle introduction.
13. There is little interpretive loss of reporting via class comparisons when only three classes are present, since only three comparisons are possible. Interpretive losses mount as more class categories are utilized. For a seven-category EG class schema (i.e., Classes I, II, IIIa, IIIb, IV, V, VI, and VIIa), standard methods would yield 21 measures of average relative importance (as in equation (13)), which themselves would be averaged versions of 42 underlying relative importance measures calculated from log odds ratios (as in equations (11) and (12)). Having only 7 rather than 21 measures of relative importance would seem to have advantages in descriptive representations, where parsimony is usually valued.
14. This number is interpretable as the weighted average of the three predicted entry rates in the first column of the second panel of Table 3, which are .541, .439, and .380, interpreted earlier as the college entry rates of the three classes if all members of each class kept their own performance levels but transitioned through the college entry regime of the salariat.

15. Finally, it is worth considering whether additional conditioning would result in the elimination of all class differences. If the conditioning set for the augmented model is expanded by including 15 additional variables for race, family structure, region of residence, urbanicity, and sector of school, the marginal means change little. The marginal means are .452, .444, and .398 for the salariat, intermediate, and working classes. It seems fair to conclude, therefore, that much of the backdoor associations generated by broader demographic factors and school sector differences are screened off by conditioning on the more proximate determinants of performance and choice in the augmented model offered in Table 4.

References

Avery, Christopher and Thomas J. Kane. 2004. ''Student Perceptions of College Opportunities: The Boston COACH Program.'' Pp. 355-91 in *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, edited by C. M. Hoxby. Chicago: University of Chicago Press.

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: John Wiley.

Breen, Richard and John H. Goldthorpe. 1997. ''Explaining Educational Differentials: Towards a Formal Rational Action Theory.'' *Rationality and Society* 9:275-305.

Dominitz, Jeff and Charles F. Manski. 1996. ''Eliciting Student Expectations of the Returns to Schooling.'' *Journal of Human Resources* 31:1-26.

Erikson, Robert, John H. Goldthorpe, Michelle Jackson, Meir Yaish, and D. R. Cox. 2005. ''On Class Differentials in Educational Attainment.'' *Proceedings of the National Academy of Sciences* 102:9730-33.

Erikson, Robert and Jan O. Jonsson. 1996. ''The Swedish Context: Educational Reform and Long-Term Change in Educational Inequality.'' Pp. 65-93 in *Can Education Be Equalized? The Swedish Case in Comparative Perspective*, edited by R. Erikson and J. O. Jonsson. Boulder, CO: Westview.

Erikson, Robert and Frida Rudolphi. 2010. ''Change in Social Selection to Upper Secondary School—Primary and Secondary Effects in Sweden.'' *European Sociological Review* 26:291-305.

Gabay-Egozi, Limor, Yossi Shavit, and Meir Yaish. 2010. ''Curricular Choice: A Test of a Rational Choice Model of Education.'' *European Sociological Review* 26:447-63.

Goldthorpe, John H. 1996. ''Class Analysis and the Reorientation of Class Theory: The Case of Persisting Differentials in Educational Attainment.'' *British Journal of Sociology* 47:481-505.

Grodsky, Eric and Melanie T. Jones. 2007. ''Real and Imagined Barriers to College Entry: Perceptions of Cost.'' *Social Science Research* 36:745-66.

Hauser, Robert M. 1976. ''On Boudon's Model of Social Mobility.'' *American Journal of Sociology* 81:911-28.

Jackson, Michelle. 2010. ''Investigating Inequalities in Educational Attainment.'' Pp. 163-77 in *Spatial and Social Disparities*, edited by J. Stillwell, P. Norman, C. Thomas, and P. Surridge. Dordrecht, Netherlands: Springer.

Jackson, Michelle, Robert Erikson, John H. Goldthorpe, and Meir Yaish. 2007. ''Primary and Secondary Effects in Class Differentials in Educational Attainment.'' *Acta Sociologica* 50:211-29.

Jonsson, Jan O. and Frida Rudolphi. 2011. ''Weak Performance—Strong Determination: School Achievement and Educational Choice among Children of Immigrants in Sweden.'' *European Sociological Review* 27:487-508.

Kloosterman, Rianne, Stijn Ruiter, Paul M. de Graaf, and Gerbert Kraaykamp. 2009. ''Parental Education, Children's Performance and the Transition to Higher Secondary Education: Trends in Primary and Secondary Effects over Five Dutch School Cohorts (1965–99).'' *British Journal of Sociology* 60:377-98.

Manski, Charles F. 1989. ''Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon.'' *Economics of Education Review* 8:305-12.

Morgan, Stephen L. 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford, CA: Stanford University Press.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. 2d ed. Cambridge, UK: Cambridge University Press.

Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. ''The Educational and Early Occupational Attainment Process.'' *American Sociological Review* 34:82-92.

Stocké, Volker. 2007. ''Explaining Educational Decision and Effects of Families' Social Class Position: An Empirical Test of the Breen-Goldthorpe Model of Educational Attainment.'' *European Sociological Review* 23:505-19.

Bio

Stephen L. Morgan, Ph.D, is Professor of Sociology and the Director of the Center for the Study of Inequality at Cornell University.