Chapter Ten

Class Origins, High School Graduation, and College Entry in the United States

Stephen L. Morgan
Cornell University

Michael W. Spiller
Cornell University

Jennifer J. Todd
Cornell University

Forthcoming in 2012 in

Determined to Succeed? Performance, Choice and Education

Edited by Michelle Jackson

Stanford University Press

Class Origins, High School Graduation, and College Entry in the United States[1]

Stephen L. Morgan, Michael W. Spiller and Jennifer J. Todd

In sociological studies of intergenerational mobility in the United States, models of educational attainment have received more attention than the analysis of early child development, the transition to work, career advancement processes, or the direct inheritance of capital. This focus was encouraged by the attention of early scholars of inequality in the United States, such as Pitirim Sorokin who wrote that

> … the school is primarily a testing, selecting, and distributing agency. … [O]ut of the many pupils who enter the door of the elementary school only an insignificant minority reach the stage of university graduation. The great majority … are eliminated, not only from school, but automatically thereby from climbing up this ladder to high social positions. (1927:188-89)

The subsequent concurrence of Talcott Parsons (1953, 1959) that schooling is central to stratification processes was then decisive in establishing models of schooling as a core concern of research on mobility processes.

The models of primary and secondary effects of social class origins on educational attainment that are developed and deployed in this volume do not currently have a prominent place in empirical scholarship in the United States. In this chapter, we hope to demonstrate why this is the case, after first reviewing the educational system in the United States and the sociological literature that has examined pathways through it in recent decades. We then provide empirical results on primary and secondary effects in the United States, offering analysis of high school graduation and college entry, with data from the 2002 through 2006 Education Longitudinal Study (ELS). We conclude with a discussion of the future utility of models of primary and secondary effects, considering some of the unresolved challenges that these models must confront in establishing warranted causal assertions about the relative sizes of primary and secondary effects.

Education in the United States in Historical Context

In this section, we present the modal pattern of age-grade progression through the educational system in the United States. Then, we discuss the evolution of this educational system, focusing on the most important transitions for students.

Basic Contours of the Current Education System

In the United States, children begin compulsory schooling at the age of 6 when they enter the first grade, although Kindergarten is now nearly universal for 5 year-olds. Compulsory schooling continues until students reach the age of 16, at which time they may discontinue schooling in most states. During the period of compulsory attendance, most students enroll in three schools: an elementary school (typically grades 1 through 6), a junior high/middle school (typically grades 7 and 8), and a high school (typically grades 9 through 12). Because schooling is compulsory through the age of 16, most adolescents continue their formal education through the tenth grade, unless they have repeated a grade, in which case they may reach age 16 in an earlier grade.[2]

The first crucial educational transition in the United States is high school graduation, which is achieved for most students at the age of 18 after completing grade 12.[3] High school graduates may then continue their education by attending one of many postsecondary possibilities, including trade schools, community colleges, liberal arts colleges, and comprehensive universities. Most of these institutions accept all applicants who pay tuition and fees, but there is a stratum of elite institutions to which access is restricted.

Historical Development

Educational institutions in the United States are diverse in size and structure because of a legacy of local control that has allowed for the emergence of a differentiated and loosely coordinated system of institutions (see Tyack 1974). This system has comparatively weak constraints on enrollment levels, and it has expanded rapidly over the past 150 years in response to the increasing demands of the economy and the citizenry (see Goldin and Katz 2008). Formal barriers to entry for each level of education are also comparatively weak, but the quality of education experienced can differ dramatically depending on the particular schools that students attend. The following brief historical narrative captures the essence of the conventional wisdom on the institutional development of education in the United States.

From its earliest days as an independent nation, and in particular through the Land Ordinance Act of 1785, the federal government vested responsibility for education in local communities. Townships were granted land, which they could use or sell, to support the schooling of their children. Most local governments then established taxes to fund the operating expenses of local schools, which allowed communities to provide elementary schooling without charging fees.

Over the course of the nineteenth century, compulsory schooling laws proliferated, and by the 1870s free elementary education was provided for all youth. Tuition-funded high schools continued to prepare elite youth to attend college, but free public high schools began to emerge. Between 1889 and 1939, the number of students in public high schools then increased from 6.7

percent to 73.3 percent, propelled by increases in the demand for education from parents seeking to educate their children for the prized positions in the rapidly expanding industrial sector (Trow 1961). During this period, the character of the American high school changed substantially. The majority of preexisting high schools had been preparing the children of the elite to attend universities. The main objective of many new public high schools in the early twentieth century was to provide a general education for students who were not destined to attend college but who instead needed skills necessary for employment in the expanding industrial sector.

The expansion of higher education institutions was similarly rapid, following the massive growth of high school education. Provincial colleges that had been established to train youth from elite families were gradually transformed into modern universities with national constituencies (see Karabel 2005). These institutions were joined by rapidly growing state universities. Following World War II, enrollments diversified and grew in response to the Servicemen's Readjustment Act of 1944 (commonly known as the "G.I. Bill"), which provided financial support for World War II veterans to attend institutions of higher education (Bound and Turner 2002). Then, as part of the "Space Race" with the Soviet Union, the federal government invested heavily in its higher education system, introducing an array of programs between 1958 and 1973 that provided financial support to institutions and directly to individual students. College attendance rates then continued to increase through the end of the twentieth century, spurred on by increasing labor market returns to college degrees. In total, the number of students enrolled in colleges and universities increased from 2.7 million students in 1950 to 14.8 million students in 2000 (Snyder, Dillow, and Hoffman 2009).

During this expansion, the stratification of higher education institutions changed in character. At the top, elite private institutions formerly attended only by the wealthy became more firmly established as national universities for students from all backgrounds (at least nominally, since these institutions were opened only to those who performed well on standardized tests, a capacity that is substantially shaped by the quality of earlier schooling that parents are able to purchase for their children; see Lemann 1999). State-sponsored institutions came to serve the vast majority of college students, providing educations that many argue vary too widely in quality. At the bottom, a patchwork of community colleges emerged to offer two-year associate's degrees and many forms of vocational training. These open-access community colleges proved particularly attractive to less academically prepared, lower socioeconomic status, and first-generation college students, in part because of their low costs. Credits from community colleges can be transferred to most bachelor's degree institutions, but the rate of bachelor's degree completion among those who initially enter community colleges remains low (see Rosenbaum, Deil-Amen, and Person 2006; Bowen, Chingos, and McPherson 2009). Community colleges have not been able to shake the characterization of Clark (1960) that they are institutions that serve a "cooling-out" function in society, wherein the aspirations of ambitious but disadvantaged students with weak preparation for college are convinced to abandon higher education in order to pursue non-professional careers.

In summary, educational has expanded dramatically over the past 150 years in the United States. Because of the rapidity and demand-from-below nature of this expansion, the overall relationship between students' origins and high school completion has weakened over time (Hout and Dohan 1996). Now that high school attendance rates have reached near saturation, and college entry rates may be shaped more by student interest and preparation than by rigid structural barriers, scholars are beginning to argue that the goal of providing basic access to educational opportunity through high schools and colleges is nearing completion. All agree that, holding these possible advances aside, a current and future challenge is to ensure that students who begin high school and college are able to stay on course to attain the degrees that they initially desire.

Models of Primary and Secondary Effects in the United States

Against this background of educational expansion, social stratification researchers in the United States have devoted considerable effort to modeling the effects of social background on educational attainment. The types of models considered in the present volume – which are typically attributed to Boudon's (1974) conception of primary and secondary effects, but which have been developed in the empirical work of Erikson and Jonsson (1996), Erikson, Goldthorpe, Jackson, Yaish, and Cox (2005), and Jackson, Erikson, Goldthorpe, and Yaish (2007) – have not received much attention. This is somewhat surprising, since Boudon (1974), when developing his distinction between primary and secondary effects, drew inspiration from the structural theory of Keller and Zavalloni (1964) that was based on observations of educational inequality in the United States.

Several factors account for this dearth of attention. Perhaps most important, the sociology of education in the United States has had its own variants of primary and secondary effects modeling, dating from the 1950s. Parsons (1953) wrote that many working class students, regardless of inherent ability, are likely to eliminate themselves from further schooling because of their adherence to a "security" rather than a "success" goal. His student Kahl (1953) then emphasized the extent to which working class families have reason to consider the costs and benefits of education differently. These ideas were progressively incorporated into the status attainment tradition, best represented by Sewell, Haller, and Portes (1969) and Sewell, Haller, and Ohlendorf (1970). Here, the consideration of goals and the differential weighting of costs and benefits were mostly lost, as status attainment researchers subordinated such processes to aspiration and motivation-based social psychological mechanisms.[4]

Nearly as important of an explanation for the lack of attention to Boudon's work in the United States is the scathing critique offered by Robert Hauser in an 18-page review essay published in 1976 in the *American Journal of Sociology*. His overall summary of Boudon's book was issued early:

> The premises of Boudon's thesis prove to be sociologically naïve, and the argument lacks cogency. The relationship between evidence and conclusions is often weak, is sometimes artifactual, and in a few instances is contradictory. The analytical and observational evidence is frequently flawed by errors of fact, of method, and of logic. (Hauser 1976:912-913).

Hauser then argued that Boudon's assumptions about class differences in the relative costs and benefits of educational attainment were incoherent because they were premised on the denial of direct class effects on adult social standing. Hauser then argued that Boudon had no evidence to deny the validity of value theories of educational inequality, and he demonstrated that Boudon's reanalysis of Kahl's 1953 article was incorrect. It stands to reason that few *AJS* subscribers then carried on to read the remaining 13 pages of Hauser's review or, more relevant here, Boudon's book itself.[5]

For reasons that are less clear, Boudon's distinction has not been used in the United States in response to its revival in European sociology in the 1990s. We have found no references to the distinction in pieces where one would most expect it, such as in the choice-sympathetic theoretical work of Morgan (2005), Hout (2006), and Lucas (2009).[6] Moreover, methodological and empirical work has similarly proceeded without its consideration, as in

Breen and Jonsson (2000, 2005), Lucas (2001), and Roksa, Grodsky, Arum, and Gamoran (2007).

This state of affairs is regrettable, not just because of a lost opportunity for cross-Atlantic cooperation. Primary and secondary effects should be considered in the United States because, as Jackson et al. (2007) demonstrate, their measurement is relevant to concerns in the United States about lost talent in response to college affordability.

In the United States, the federal government's role in promoting access to higher education has been vigorously debated in the past two decades. To some extent, this renewed attention is a direct result of the concern over escalating tuition costs at both public and private universities. Scholars do not agree on the extent to which short-term credit constraints prevent students who would otherwise attend college from attending (Kane 1999; Cameron and Taber 2004). Some argue that college entry rates increased at the same time that college costs increased, and many of the prospective students who are eligible for Pell grants – the federal government's main tuition grant for low-income college students – do not even apply for them. Others have countered that these surprisingly low take-up rate are not evidence that low-income students have sufficient access to credit on their own, but rather that the application forms for federal grants are too complex, intimidating, and burdensome for families with little past experience with higher education (King 2006; Dynarski and Scott-Clayton 2006). Thus, although various scholars estimate that a $1000 increase in financial aid is associated with an increase in the rate of attending college of between 3.5 and 6 percent, it is not clear how to effectively deliver this aid to students in the United States (Leslie and Brinkman 1988; Dynarski 2003).

More generally, in recent years, two opposite positions have crystallized in the debate over how to decrease the impact of social origins on educational attainment and social standing (see Heckman and Krueger 2003). One position calls for the expansion of government programs for adolescents and young adults, including the strengthening of vocational high school training programs, displaced worker training programs, and subsidies to relieve short-term credit constraints that prevent able students from attending college. The other position, in contrast, maintains that policies targeted at adolescents and young adults have low social returns because academic and career successes are determined earlier in the lifecourse. If the goal is to increase the skills and attainments of the workforce in the long run, and to break down rates of intergenerational transmission of poverty, policies should provide more support for programs for young children.

How is the literature on primary and secondary effects relevant to this debate? Jackson et al. (2007) argue that the existence of meaningful secondary effects has direct policy implications. These effects provide support for the argument against the increasingly dominant position in the United States' policy dialogue that investments in early childhood education have a high social rate of return but policies that promote access to higher education have a comparatively low social rate of return. Jackson et al. (2007:225) argue further that "the possibility must be recognized that any gains that may be made in reducing primary effects via interventions in early years – at a probably high cost – will be at least in some degree offset in so far as the further problem of secondary effects is not addressed." The literature on primary and secondary effects literature is therefore at the frontier of policy debate, helping to determine, in the best traditions of European welfare states research, how to develop the optimal mix of development subsidies from cradle to grave.[7]

In the remainder of this chapter, we estimate primary and secondary effects for a three-class model of social stratification, following the basic framework worked out by Erikson et al. (2005) and Jackson et al. (2007).[8] We consider the two most important transitions discussed in the literature on educational attainment in the United States: high school graduation and college entry. In conclusion, we offer plausible interpretations of the results, and we raise some methodological concerns that this strand of literature should confront in order to gain the attention of a wider range of scholars of inequality and education.

Methods

Data and Analysis Sample

Data were drawn from the 2002 base-year through 2006 second follow-up 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 of the U.S. Department of Education. We restricted the analysis to respondents who participated in all three waves of the survey, using a weight to account for sample attrition between the sophomore year (in 2002) and the two follow-up surveys in 2004 and 2006. The resulting analysis sample includes 12,546 students for the high school graduation models. Then, for the college entry models, we dropped students who did not complete high school, resulting in an analysis sample of 11,400 students for the college entry models. Substantial detail on the definition of the analysis sample, procedures for developing weights, and decisions for handling missing data are available in the online Appendix.

Measures

High School Graduation and College Entry. Our respondents were sampled in 2002 as enrolled high school sophomores (i.e., second-year high school students, typically aged 15 or 16). We coded high school graduates as those who obtained high school diplomas before the second follow-up in 2006. We did not include GED recipients among our high school graduates.[9] Of the 12,546 students who had complete data for high school graduation and social class, 89.8 percent graduated from high school. Approximately 2 percent of these high school graduates experienced a delay in their graduation, defined as receiving a diploma between the summer of 2004 but before data collection in 2006.

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

Social Class. Following established procedures for estimating primary and secondary effects in the sociology of education, we adopt the three-tiered Erikson-Goldthorpe (EG) class schema of salariat, intermediate, and working classes. First, we coded the broad occupational categories of fathers and mothers (or male and female guardians) into the nine-category EG schema. Then, we collapsed these nine-category classes to the three tiers of salariat, intermediate, and working, after which we then further collapsed father's and mother's class position into an overall family measure. In this family class coding, we very slightly privileged father's class, as explained in the online Appendix.

Academic 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. Our sophomore year academic performance measure is a standardized composite of 2002 math and reading IRT-estimated scores. Our 2004 academic performance measure is a standardized composite of 2004 math and 2002 reading IRT-estimated scores. 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. Prior work shows considerable stability in relative rankings throughout high school for reading achievement, and this is 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. In contrast, math achievement continues to evolve as a function of social class, mostly because of greater differentiation of coursetaking through the final two years of high school.

Modeling Strategy

We first estimate logit models for the two educational transitions of high school graduation and then college entry. We do not employ the smoothing techniques developed by Erikson et al. (2005) and instead estimate models directly on the observed data. For each transition, we estimate pooled logit models for all three classes. To calculate primary and secondary effects, we then estimate models that include academic performance interacted with the indicator variables for class. In so doing, we create parameter estimates that are equivalent to what could be obtained by class-specific logit models. We then calculate marginal observed and "what-if" transition rates across classes, which Erikson et al. (2005) and Jackson et al. (2007) label "counterfactual," but which are generally labeled "synthesized" in this volume (see Chapter 2). These rates are equal to the sort of class-by-class analysis used in this research tradition, with the added benefits that no smoothing assumptions are introduced and the calculation of standard errors is straightforward. (In the online Appendix, we also estimate the same models for three education groups and for nine class-education groups.)

Each of these logit models is weighted by the estimated inverse probability of being in the analysis sample, multiplied by the panel weight developed by the data distributors (F2BYWT). To sharpen counterfactual projections, we estimated two sets of inclusion probabilities for the college entry models, as fully explained in the online Appendix. The first set of probabilities is estimated with a logit model that includes only those who had graduated from high school. The resulting weight generates college entry rates that can be generalized only to high school graduates. Since this group is not selected at random from the full population of high school students who could *ex ante* enter college, we also estimated a second set of probabilities derived from a broader baseline logit that includes even those who do not graduate from high school. These probabilities are then used to generate weights that deliver college entry rates that can be generalized to all high school sophomores, as if they had all graduated from high school by 2006 and were available in theory to enter college immediately. As expected, we will show later that this broader set of estimated college entry rates is generally lower, but the overall pattern is the same as when the weights only permit generalizations to the selected group of students who are observed to have graduated from high school.

Models of Primary and Secondary Effects

Table 1 presents observed transition rates for both high school graduation and college entry, separately by class and for the full sample. Within each panel, means and standard deviations of the performance measure are presented, separately by class and for the full sample. The results show clearly that class in the United States is a strong baseline predictor of both transition rates and performance in high school, as measured by standardized achievement tests.

Table 2 presents six logit models in three separate pairs. In the first column of each pair, the educational transition under consideration is modeled solely as a function of social class origins. In the second column of each pair, primary and secondary effects are modeled by estimating coefficients for social class, performance, and the interaction between performance and class. Because of the parameterization chosen, wherein class is represented by two dummy variables, each of which is interacted with performance, these models produce the same predicted probabilities of making the educational transitions that could be obtained by three class-specific logit models with performance as the sole predictor variable in each.

For the high school graduation model, class and performance both predict whether or not sophomores complete high school. The association between performance and the probability of graduating high school appears to be slightly weaker for lower classes, although this difference is modest at best and well within sampling error. The predicted probabilities from this model are then presented in panel (a) of Table 3, and they are reported as estimated true and estimated "what-if" rates (labeled "synthesized" rates elsewhere in this volume). The nine cells of the panel give the estimated probability for each of the three classes that would prevail if (1) they kept their own performance distribution but (2) they went through the estimated transition regime for each of the three classes. The diagonal of the panel includes simple estimated probabilities of graduating high school for each class, which are .933 for the salariat class, .886 for the intermediate class, and .848 for the working class. [10] The values off the diagonal are estimated what-if transition rates. For example, the value .901 in the upper right-hand cell of the panel is the estimated probability that the salariat class would graduate from high school, assuming the salariat class retains its distribution of performance but moves through the transition regime of the working class. Rather than moving through a logit model with an index function of 2.75 + (.97)*Performance* as its argument, they move through a logit model with an index function of (2.75-.5) + (.97-.09)*Performance* as its argument. For every value of performance, the predicted probability is then lower. Similar comparisons across columns reveal a strict ordering of transition rates, as determined by the distribution of performance and the estimated logit parameters. Higher classes have lower transition rates when they pass through the transition regimes of lower classes, and vice versa.

What about comparisons across rows within columns? For example, consider the difference between .933 in the upper left-hand cell of the panel and .892 in the lower left-hand cell. This difference is produced by the alternative distributions of performance in the salariat and the working classes, as would be observed if they both passed through the transition regime of the salariat class. Likewise, the difference between .901 in the upper right-hand cell and .848 in the lower right-hand cell is produced by the same difference in distributions of performance between the salariat and the working classes, but in this case the comparison holds under the alternative scenario where both classes move through the transition regime of the working class.

Figure 1 plots all of these predicted probabilities by observed levels of performance, separately by class. Because the logit functional form is constrained to have its largest marginal effect where its *S*-curve crosses the value of .5, and because probabilities of graduating high school in the United States are quite high (for samples of sophomores), the predicted values over the observed data range only reveal the upper tail of the underlying *S*-curve. Each of the point values in panel (a) of Table 3 corresponds to one of nine separate points that lie along the dashed lines of predicted probabilities in Figure 1. The values in the first column of panel (a) are the corresponding points on the three dashed lines that intersect a vertical line that could be drawn through mean of the performance distribution of the salariat class. The values in the second and third columns are then three points on the dashed lines that intersect vertical lines that could be drawn through the mean of the performance distributions for the intermediate and working classes, respectively.

Returning to Table 2, the models for college entry are presented in the last four columns, in pairs separately for the two weighting schemes discussed earlier. For the first pair, college entry rates are estimated by logit models that are weighted toward the distribution of high school graduates. For the second pair, the same logit models are weighted toward the distribution of all high school sophomores. The logit coefficients suggest that the college entry rate is lower for all three classes when weighted toward all high school sophomores, as indicated for the first model in each pair by the decline in the intercept from .16 to .07 as well as declines in the class dummies from .-54 to -.58 and -1.08 to -1.12. When performance is added to the predictor variables, these differences persist, although the coefficients for performance do not vary substantially across models.

As shown in panels (b) and (c) of Table 3, college entry rates are considerably lower than high school graduation rates in the United States. The probability that a high school graduate will enter college within six months of high school graduation is only .541, .406, and .285 for the salariat, intermediate, and working classes, respectively. These rates are slightly lower at .518, .375, and .259 when weighted toward the characteristics of high school sophomores. Similar to Figure 1, the full set of predicted probabilities from the college entry models akin to those in panels (b) and (c) are presented in Figures 2 and 3. The predicted probabilities are lower on average than is the case for high school graduation, and thus most of the corresponding logistic

S-curves are observed over the data range of performance. Moreover, performance more strongly predicts college entry than was the case for high school graduation. When these two factors are combined, the predicted probabilities of college entry increase from less than .1 for all classes to greater than .8 for all classes.

Taken together, and with reference to the established literature on primary and secondary effects, as well as new contributions in this volume, four basic patterns are clear from the analysis:

1. Class and performance predict educational transitions in an orderly fashion in the United States, similar to what is observed in other industrial and post-industrial societies when a similar modeling framework is deployed.
2. The absolute rate of high school graduation is quite high in the United States (for a sample of sophomores), in comparison to analogous rates in other countries that have more finely differentiated forms of secondary education.
3. The absolute rate of entry into four-year colleges in the United States is substantially lower than the high school graduation rate in the United States.

In results available in the online Appendix, we offer results that support a fourth conclusion:

4. Family class and family education have separable associations with both high school graduation and college entry. The associations cumulate in complementary fashion, but family education is a slightly stronger predictor of both high school graduation and college entry than is social class.

Beyond the empirical summary of our findings just offered, we are less comfortable making additional comparative statements about the relative sizes of primary and secondary effects, both across transitions in the United States and across countries. We leave these comparisons to the Editor for the concluding chapter, which utilizes comparisons of log odds ratios drawn from our results. In the next section, we conclude our chapter with a discussion of our hesitancy, as a set of general concerns that may be particularly pronounced in the stratification system of the United States.

Discussion

We are comfortable stating that primary and secondary effects are surely both present in the United States, as in other countries. We support the basic position that students make choices that determine educational attainment, at least in some fashion and to some extent. Thus, a secondary effect exists, under our implicitly assumed theoretical model. Moreover, since we also see no reason to suspect that such a choice process can explain anything more than a portion of the association between class origins and educational attainment, we must also assume that a primary effect exists.

Nonetheless, we have some general concerns that suggest a reorientation of these methods of analysis, especially when applied to educational attainment in the United States. We conclude with this cautionary perspective, which we hope reads as an appeal for further work within this mode of analysis rather than as a reason to abandon it (see also Morgan 2012).

To fix ideas for this concluding discussion, we will define primary and secondary effects for the college entry decision, using graphs as in Erikson et al. (2005). We consider the primary effect of stratification to be the effect of class origins (*Class*) on college entry (*G*) that operates indirectly via academic performance (*P*) in prior schooling. The secondary effect is then the remaining direct effect of class origins on college entry. The causal graph in Figure 4(a) is consistent with these definitions, and it is equivalent in structure to Figure 2(a) of Erikson et al. (2005).[11] The primary and secondary effects of stratification are the causal pathways $Class \rightarrow P \rightarrow G$ and $Class \rightarrow G$, respectively.

As shown in the causal graph, the secondary effect is a traditional net direct effect. A causal graph with the same basic content is presented in Figure 4(b). For this augmented graph, the primary and secondary effects of stratification are defined as two separable causal pathways. As before, the primary effect is $Class \rightarrow P \rightarrow G$. Now, the secondary effect is $Class \rightarrow U \rightarrow G$, where *U* can be thought of as the choice mechanism that leads students of different class origins to choose differently. The graph in Figure 4(a) is best regarded as a simplification of the graph in Figure 4(b), which is permissible because *U* is unobserved.[12]

The recent literature on primary and secondary effects suggests that the causal model in Figure 3(b) is too simple. Erikson et al. (2005) extend their simple "potentially causal model," which is represented by the graph in Figure 4(a), to an elaborated one that is closer to the graph presented in Figure 4(c). They label this graph a "more realistic model," and it includes a path that is equivalent to the path $Class \rightarrow AD \rightarrow P \rightarrow G$ in Figure 4(c). For Erikson et al. (2005), *AD* is an unobserved anticipatory decision that is taken some years prior to the actual educational transition (and which they label an "unobserved early choice component"). The anticipatory decision structures *P* in the run up to the educational transition. The secondary effect of stratification then operates through two causal pathways, $Class \rightarrow G$ and $Class \rightarrow AD \rightarrow P \rightarrow G$, the second of which is mixed in with what remains the primary effect conceptually, $Class \rightarrow$

P → *G*, if *AD* is unobserved. As a result, standard methods attribute the effect of *Class* on *AD* to the association between *Class* and *P,* which shifts some of the true secondary effect to the estimated primary effect. The result is that the estimate of the secondary effect is downwardly biased, and the estimate of the primary effect is upwardly biased. Or, stated in terms of our results, the difference across columns in the what-if transition rates in Table 3 are too small because the differences across rows are too large.

Although the reasoning of Erikson et al. (2005) on anticipatory decisions is correct, the full set of causal identification challenges is more extensive. Indeed, Erikson et al. (2005) also allow for the direct effect of *AD* on *G* in their "more realistic model," and that creates complications for the conclusion that the anticipatory decision necessarily biases downward estimates of the secondary effect.

We would argue for a baseline model that is more complex than those in Figure 4(a) through 4(c). We see it as essential to include exogenous variables summarized by *X* and to also allow the intermediate unobserved variable *U* to encompass more than just an anticipatory decision. This baseline model is presented in Figure 4(d), and we would argue that the burden of effort, in justifying present methods and direct legacies of them, should be to make the case that some of the arrows in the model in Figure 4(d) should be deleted.

Consider some of the variables that the literature on educational attainment in the United States 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* (and thereby broadening its default definition from simply an omnibus choice-theoretic mediator). Accordingly, race is a common cause of both class position and beliefs about the opportunity structure. The back-door paths *Class* ← *X* → *U* → *G* and *Class* ← *X* → *U* → *P* → *G* then follow from the positions in the literature that perceptions of the opportunity structure, *U*, determine *G* directly and also determine *G* indirectly via prior preparation *P*. Furthermore, the back-door path *Class* ← *X* → *P* → *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. In addition, institutional features of schooling, such as between-school differences in instructional quality and within-school curriculum differentiation, create similar problems.

Accordingly, the causal graph in Figure 4(b) does not appear to give a solid foundation to current modeling practices when applied to the educational system in the United States (nor, possibly, in other countries). Using standard methods is tantamount to replacing the many pathways through *X* to a simple stand-in path *Class* → *U* → *G*. In this case, *U* is 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 will probably lead to overestimation of the secondary effects of stratification, as structural effects will then be let in through the back-door and attributed (by nothing more than an assumption of simple choice-driven behavior) to choice-interpreted secondary effects. At the same time, the estimate of the primary effect, usually represented by *Class* → *P* → *G*, is then 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). Overall, the bias in estimates of primary and secondary effects from standard estimation practices will contain many countervailing sources, making *a priori* claims about the direction of biases impossible. Thus, for example, stating that the differences across columns in the what-if transition rates in Table 3 are too small because of omitted variables is only true if an anticipatory decision is the only omitted variable that generates additional dependence in the true causal graph.

Overall, we regard the causal identification challenges just discussed to be sufficient grounds for caution in interpreting differences between what-if transition rates from these models as causal quantities that reflect choice-based secondary effects. Yet, these are identification challenges not barriers, and we look forward to further empirical and theoretical work designed to overcome them. The first step is to model additional causal pathways that emanate from social class, so as to partial out back-door associations that comingle with narrower forms of primary and secondary effects. The second step is to directly model the choice mechanism that constitutes the conjectured secondary effect, considering 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.

Endnotes

---

[1] We thank Theo Leenman and Kelly Andronicos for excellent research assistance as well as the stimulating comments of Michelle Jackson, Louis-André Vallet, CIQLE seminar participants at Yale, and sociology seminar participants at Nuffield College, Oxford.

[2] However, see Oropesa and Landale (2009) for evidence that a non-negligible proportion of immigrant youth never enroll in a school after entering the United States.

[3] Following Mare (1981), high school graduation is considered a school continuation decision in the United States (see also Mare 1995). Models of primary and secondary effects have focused more frequently on transitions between levels of education, as in our college entry models. We will therefore focus most of our explanatory attention on the college entry decision because it is considered by all traditions of analysis to be an important school continuation decision and an important educational transition.

[4] Still, in the broader status attainment research tradition, the effects of family background on educational attainment are routinely decomposed into indirect effects that operate through mediating variables such as academic performance as well as net direct effects that cannot be attributed to the same mediators.

[5] Both Boudon's book and Hauser's review covered more ground than primary and secondary effects, and the contention here is simply that Hauser's review prompted many readers to ignore the primary and secondary effects distinction.

[6] It is notable that other recent contributions to theories of educational inequality have drawn inspiration from it, particularly Goldthorpe (1996, 2000), Breen and Goldthorpe (1997), and the work that has developed to consider their models of relative risk aversion (see Goldthorpe 2007 for a review). This theoretical work has also had more influence in European sociology than in the United States.

[7] And here, there are complementarities that need to be assessed. For example, existing evidence suggests that the benefits of early childhood programs for children living in poverty may evaporate as these children enter relatively poor performing public schools (McKey et al. 1985; Puma, Bell, Cook, and Heid 2010). This erosion of a promising early treatment effect suggests that broad swaths of the public school system may need to be fixed before the benefits of early childhood programs can be sustained until students approach the college entry decision.

[8] For comparison with other contributions to the volume, we also estimate primary and secondary effects for education groups and for nine class-education groups. These additional results are presented in the online Appendix.

[9] The GED is a test that, if passed, certifies that an individual has academic skills at the level of a high school graduate. Although general equivalency diploma (or degree) is the typical long form of GED, the acronym is technically short for General Educational Development.

[10] These rates of high school graduation among high school sophomores are higher than many standard estimates of overall high school completion, and this is partly attributable to having estimated the high school graduation rate of high school sophomores only. Thus, students who drop out before the sophomore year, because they have obtained the age of 16 early (possibly because they were retained in grade earlier in their school careers) are not in our models. Even so, there is no clear baseline number against which to compare our rates. In part because of the constraints of administrative records, as well as debates about definitional issues, little consensus exists on the absolute levels and trends in high school completion since 1970. For example, estimates produced by the U.S. National Center for Education Statistics, using data and methods similar to ours, suggest that high school graduation rates have slowly increased since 1970 to a current rate of 85 percent (Laird, Cataldi, Kewal, Ramani, and Chapman 2008). By contrast, Swanson (2004) reports a decline in the rate to 68 percent, and Heckman and Lafontaine (2007) report a relatively stagnant rate of 75 percent.

[11] Erikson et al. (2005:9732-33) present graphs that they label "path diagrams," but which they also describe as a "potentially causal model" for their Figure 2(a) and a "more realistic model" for their Figure 2(b). Because their article uses counterfactual language throughout, and moreover cites the work of Rubin, Rosenbaum, and Holland in the section where these figures are introduced, we regard their path diagrams as putative causal models meant to open up a discussion of causal identification issues.

[12] For causal graphs, 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. In this tradition of analysis, the causal arrow in $Class \rightarrow G$ is given a choice-theoretic characterization, which justifies granting the pathway an omnibus mediating variable U in Figure 4(b). See Morgan (2012) for a more specific model for U.

References

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

Bound, John and Sarah Turner. 2002. "Going to War and Going to College: Did the G.I. Bill Increase Educational Attainment for Returning Veterans?" *Journal of Labor Economics* 20:784-815.

Bowen, William G., Matthew M. Chingos, and Michael S. McPherson. 2009. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton: Princeton University Press.

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

Breen, Richard and Jan O. Jonsson. 2000. "Analyzing Educational Careers: A Multinomial Transition Model." *American Sociological Review* 65:754-72.

—. 2005. "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility." *Annual Review of Sociology* 31:223-43.

Cameron, Stephen V. and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *Journal of Political Economy* 112:132-82.

Clark, Burton R. 1960. "The 'Cooling-out' Function in Higher Education." *American Journal of Sociology* 65:569-76.

Dynarski, Susan M. 2003. "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion." *American Economic Review* 93:279–88.

Dynarski, Susan M. and Judith Scott-Clayton. 2006. "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics." *National Tax Journal* 59:319-56.

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: Westview.

Goldin, Claudia Dale and Lawrence F. Katz. 2008. *The Race between Education and Technology*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Goldthorpe, John H. 1996. "Class Analysis and the Reorientation of Class Theory: The Case of

Persisting Differentials in Educational Attainment." *The British Journal of Sociology* 47:481-505.

—. 2000. *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press.

—. 2007. "The Theory Evaluated: Commentaries and Research." Pp. 73-100 in *On Sociology*, vol. 2. Stanford: Stanford University Press.

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

Heckman, James J. and Alan B. Krueger. 2003. *Inequality in America: What Role for Human Capital Policies?* Cambridge: MIT Press.

Heckman, James J. and Paul A. LaFontaine. 2007. "The American High School Graduation Rate: Trends and Levels." National Bureau of Economic Research, Working Paper No. 13670, Cambridge, Mass.

Hout, Michael. 2006. "Maximally Maintained Inequality and Essentially Maintained Inequality: Crossnational Comparisons." *Sociological Theory and Methods* 21:237-52.

Hout, Michael and Daniel P. Dohan. 1996. "Two Paths to Educational Opportunity: Class and Educational Selection in Sweden and the United States." Pp. 207-31 in *Can Education Be Equalized? The Swedish Case in Comparative Perspective*, edited by R. Erikson and J. O. Jonsson. Boulder: Westview.

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.

Kahl, Joseph. 1953. "Educational and Occupational Aspirations of Common Man Boys." *Harvard Educational Review* 23:186-203.

Kane, Thomas J. 1999. *The Price of Admission: Rethinking How Americans Pay for College*. Washington, D.C.: Brookings Institution Press.

Karabel, Jerome. 2005. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston: Houghton Mifflin.

Keller, Suzanne and Marisa Zavalloni. 1964. "Ambition and Social Class: A Respecification." *Social Forces* 43:58-70.

King, Jacqueline E. 2006. "Missed Opportunities Revisited: New Information on Students Who Do Not Apply for Financial Aid." American Council on Education, Washington, DC.

Laird, Jennifer, Emily F. Cataldi, Angelina Kewal Ramani, and Chris Chapman. 2008. "Dropout and Completion Rates in the United States: 2006." National Center for Education

Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Lemann, Nicholas. 1999. *The Big Test: The Secret History of the American Meritocracy*. New York: Farrar, Straus and Giroux.

Leslie, Larry L. and Paul Brinkman. 1988. *The Economic Value of Higher Education*. New York: Macmillan.

Lucas, Samuel R. 2001. "Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106:1642-90.

—. 2009. "Stratification Theory, Socioeconomic Background, and Educational Attainment: A Formal Analysis." *Rationality and Society* 21:459-511.

Mare, Robert D. 1981. "Change and Stability in Educational Stratification." *American Sociological Review* 46:72-87.

—. 1995. "Changes in Educational Attainment and School Enrollment." Pp. 155-213 in *State of the Union: America in the 1990s*, Vol. 1, edited by R. Farley. New York: Russell Sage Foundation.

McKey, Ruth H., Larry Condelli, Harriet Ganson, Barbara J. Barrett, Catherine McConkey, and Margaret C. Plantz. 1985. "The Impact of Head Start on Children, Families, and Communities: Final Report of the Head Start Evaluation, Synthesis, and Utilization Project." CSR Inc. and the U.S. Department of Health and Human Services, Washington, DC.

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

Morgan, Stephen L. 2012. "Models of College Entry in the United States and the Challenges of Estimating Primary and Secondary Effects." *Sociological Methods and Research*.

Oropesa, R. S. and Nancy S. Landale. 2009. "Why Do Immigrant Youths Who Never Enroll in U.S. Schools Matter? School Enrollment among Mexicans and Non-Hispanic Whites." *Sociology of Education* 82:240-66.

Parsons, Talcott. 1953. "A Revised Analytical Approach to the Theory of Social Stratification." Pp. 92-128 in *Class, Status, and Power: A Reader in Social Stratification, 1st Edition*, edited by R. Bendix and S. M. Lipset. Glencoe: Free Press.

—. 1959. "The School Class as a Social System." Pp. 129-54 in *Social Structure and Personality*, edited by T. Parsons. New York: Free Press.

Puma, Michael, Stephen Bell, Ronna Cook, and Camilla Heid. 2010. "Head Start Impact Study: Final Report." Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Roksa, Josipa, Eric Grodsky, Richard Arum, and Adam Gamoran. 2007. "United States: Changes in Higher Education and Social Stratification." Pp. 165-91 in *Stratification in Higher Education: A Comparative Perspective*, edited by Y. Shavit, R. Arum, A. Gamoran, and G. Menahem. Stanford: Stanford University Press.

Rosenbaum, James E., Regina Deil-Amen, and Ann E. Person. 2006. *After Admission: From College Access to College Success*. New York: Russell Sage.

Sewell, William H., Archibald O. Haller, and George W. Ohlendorf. 1970. "The Educational and Early Occupational Status Attainment Process: Replication and Revision." *American Sociological Review* 35:1014-24.

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

Snyder, Thomas D., Sally A. Dillow, and Charlene M. Hoffman. 2009. *Digest of Education Statistics 2008*. Washington, DC: National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education.

Sorokin, Pitirim. 1927. *Social Mobility*. New York: Harper & Brothers.

Swanson, Christopher B. 2004. "Who Graduates? Who Doesn't? A Statistical Portrait of Public High School Graduation, Class of 2001." The Urban Institute, Washington, D.C.

Trow, Martin. 1961. "The Second Transformation of American Secondary Education." *International Journal of Comparative Sociology* 2:144-66.

Tyack, David B. 1974. *The One Best System: A History of American Urban Education*. Cambridge: Harvard University Press.

Table 1. Means and Standard Deviations of Variables Used in the Analysis, Separately for All High School Sophomores and for All High School Graduates

	Salariat Class		Intermediate Class		Working Class		Full Sample	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	(a) All Sophomores (i.e., Second-Year High School Students)							
High School Graduation	.933		.886		.848		.898	
Performance in 10th grade (Standardized test)	.275	.967	-.141	.993	-.368	.917	.001	1.000
Number of Respondents	6,704		2,434		3,408		12,546	
	(b) High School Graduates							
College Entry	.541		.406		.285		.443	
Performance in 12th grade (Standardized test)	.331	.943	-.072	.986	-.296	.923	.076	.986
Number of Respondents	6,293		2,185		2,922		11,400	

Source: Education Longitudinal Study of 2002-2006

Notes: Means and SDS are weighted, but the cells for the Number of Respondents are the observed numbers of respondents in the data. For the proper weighted distribution of class, see the final column of Table S1 in the online Appendix. The mean and SD of performance in the twelfth grade are not 0 and 1 in the second panel because the variable was standardized across all 12,546 high school sophomores but then presented for this panel only for the 11,400 students from the high school graduate subsample.

Table 2. Logit Models for High School Graduation and College Entry

	High School Graduation		College Entry (Among HS Graduates)		College Entry (Weighted to the Distribution of all HS Sophomores)	
Constant	2.64	2.75	.16	-.25	.07	-.31
	(.07)	(.08)	(.04)	(.05)	(.04)	(.05)
Class:						
Intermediate	-.59	-.28	-.54	-.19	-.58	-.19
	(.10)	(.14)	(.06)	(.07)	(.06)	(.07)
Working	-.92	-.50	-1.08	-.54	-1.12	-.56
	(.09)	(.11)	(.06)	(.07)	(.06)	(.07)
Performance		.97		1.27		1.28
		(.07)		(.05)		(.05)
x Intermediate		-.06		.03		.04
		(.12)		(.09)		(.09)
x Working		-.09		-.14		-.13
		(.10)		(.08)		(.08)
Wald Chi-Squared	105.9	469.8	336.7	1,206.0	354.0	1,213.1
N	12,546	12,546	11,400	11,400	11,400	11,400

Source: See Table 1.
Notes: Standard errors in parentheses.

Table 3. High School Graduation and College Entry Rates by Class Origins, Estimated from the Models in Table 2 that Adjust for Prior Performance

(a) High School Graduation Rates

Observed Class	What-if Class		
	Salariat	Intermediate	Working
Salariat	.933	.917	.901
	(.004)	(.007)	(.006)
Intermediate	.906	.886	.866
	(.006)	(.009)	(.007)
Working	.892	.870	.848
	(.007)	(.010)	(.007)

(b) College Entry Rates (for High School Graduates Only)

	What-if Class		
	Salariat	Intermediate	Working
Salariat	.541	.507	.423
	(.009)	(.013)	(.013)
Intermediate	.439	.406	.336
	(.011)	(.012)	(.011)
Working	.380	.346	.285
	(.009)	(.012)	(.010)

(c) College Entry Rates (Weighted to the Distribution of all High School Sophomores)

	What-if Class		
	Salariat	Intermediate	Working
Salariat	.518	.484	.399
	(.009)	(.013)	(.013)
Intermediate	.409	.375	.307
	(.009)	(.012)	(.010)
Working	.352	.318	.259
	(.009)	(.012)	(.009)

Source: See Table 1.
Notes: Observed transition rates on the diagonal and what-if transition rates on the off diagonal. Standard errors in parentheses.

Figure 1. Smoothed Performance Distributions and Estimated High School Graduation Rates by Social Class

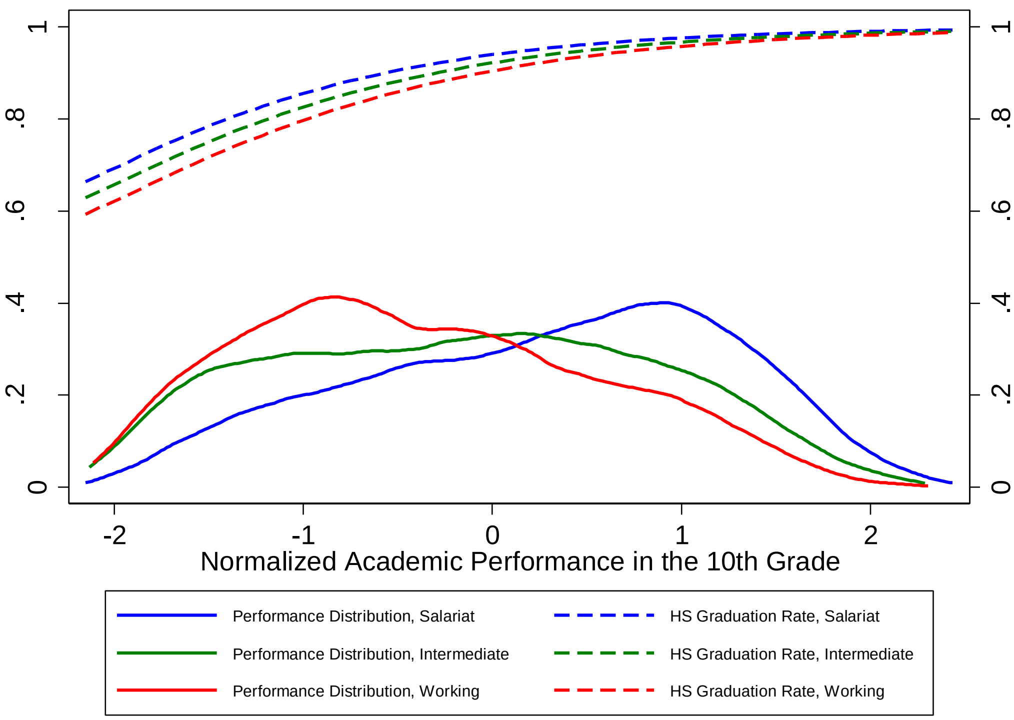

Figure 2. Smoothed Performance Distributions and Estimated College Entry Rates Among High School Graduates by Social Class

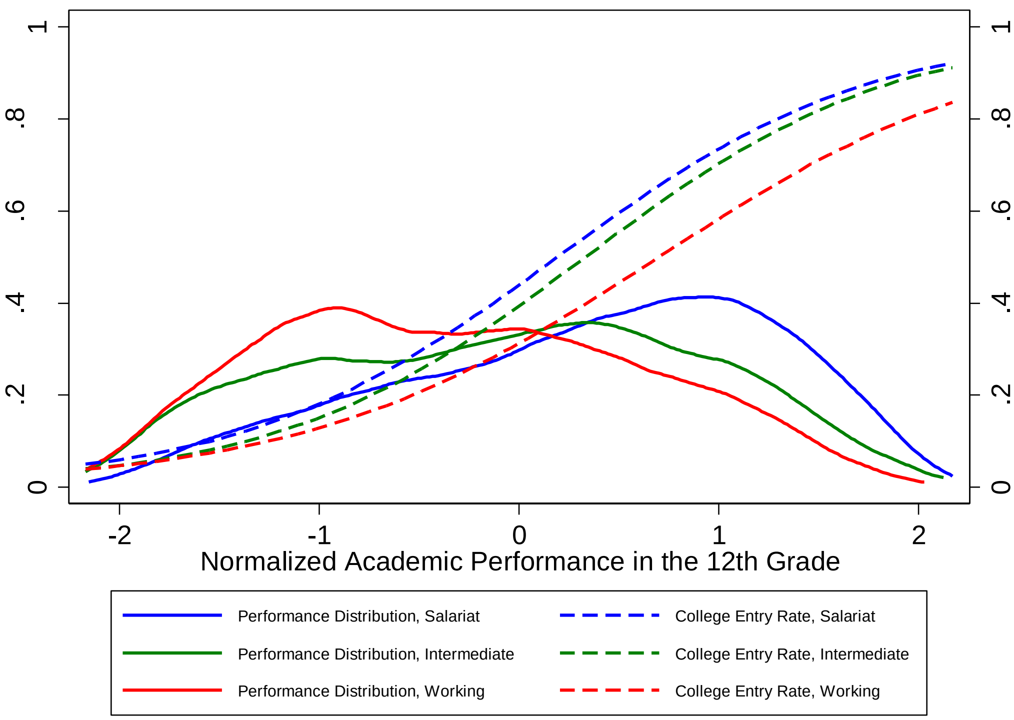

Figure 3. Smoothed Performance Distributions and Estimated College Entry Rates Among High School Graduates by Social Class, Weighted Toward the Full Distribution of High School Sophomores

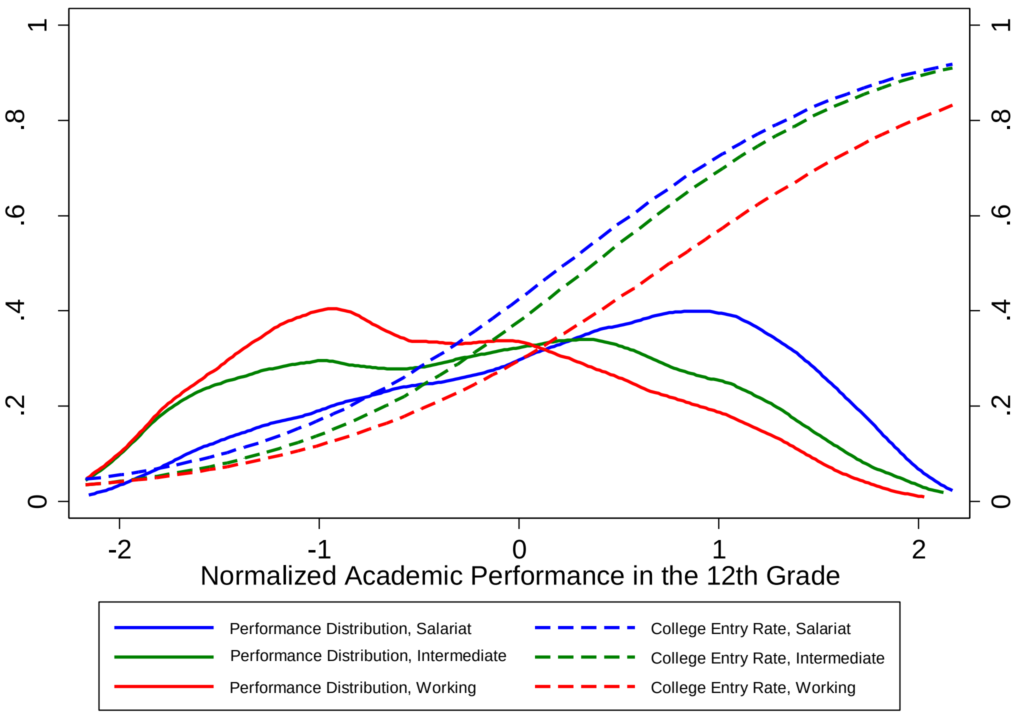

Figure 4. Alternative Causal Graphs of Primary and Secondary Effects