Jeongeun Kim
Jiyun Kim
Stephen L. DesJardins
Brian P. McCall

Completing Algebra II in High School: Does It Increase College Access and Success?

Noting the benefits of mathematics in students' future educational attainment and labor market success, there is considerable interest in high school requirements in terms of course-taking in mathematics at the national, state, and school district level. Previous research indicates that taking advanced math courses in high school leads to positive college outcomes. However, these studies often fail to account for the self-selection of students into curricular pathways that may result in biased estimates of the effect of course-taking on subsequent educational outcomes. Applying an instrumental variable (IV) approach, we investigate how the level of math courses a student completes in high school differently affects their chances of attending and completing postsecondary education. Using longitudinal student unit record data from Florida, our results indicate that a statistical model that does not account for students' self-selection produces results different from a technique that corrects for this potential source of bias. Specifically, completing Algebra II significantly increases the probability of attending college, particularly two-year colleges, but has no significant effect on degree attainment.

Keywords: high school mathematics; high school course-taking; Algebra II; college attendance; degree attainment; instrumental variable

Introduction

There is broad consensus that mathematics skills are critical for one's future educational attainment and labor market success. Prior research

---

Jeongeun Kim is a Ph.D. candidate at the Center for the Study of Higher and Postsecondary Education, University of Michigan; okuje@umich.edu. Jiyun Kim is the Deputy Director of Korean Ministry of Education, Science, and Technology. Stephen L. DesJardins is Professor at the Center for the Study of Higher and Postsecondary Education and Ford School of Public Policy, University of Michigan. Brian P. McCall is Professor at the Center for the Study of Higher and Postsecondary Education, Ford School of Public Policy, and Department of Economics, University of Michigan.

The Journal of Higher Education, Vol. 86, No. 4 (July/August)

indicates that students who complete more and higher levels of mathematics courses in high school are more likely to pursue higher education and to have higher earnings later in life (Adelman, 2006; Altonji, 1995; Rose & Betts, 2004). Despite the importance of math skills, the United States has consistently ranked low among Organization for Economic Cooperation and Development (OECD) countries in international comparisons of math literacy, and the average math score of fifteen-year-old students in the U.S. has fallen significantly below the OECD average (OECD, 2013).

Increasing mathematics curriculum standards in high school has therefore received particular attention as a means to strengthen the rigor and raise the expectations for what counts in the awarding of a high school diploma. The curricular improvement envisioned in *A Nation at Risk* (1983) urged states to raise academic requirements for high school graduation, recommending that all students complete three courses in math and science to graduate from high school (Teitelbaum, 2003). The federal No Child Left Behind (NCLB) Act of 2001 prompted increased accountability at the state level, particularly in the specification of mathematics curriculum requirements (Reys, 2006). In response to these national policy changes, many states have increased the minimum number of credits (and years) of mathematics required to graduate from high school, and states are also specifying particular types of mathematics courses that students must complete (Federman, 2007; Reys, Dingman, Nevels, & Teuscher, 2007).

There is some evidence that these statewide efforts to increase high school graduation requirements encourage students to earn more mathematics credits in high school (Teitelbaum, 2003). However, what remains unclear is whether and how the benefits of completing more advanced courses in general, and the completion of specific types of advanced courses (e.g., mathematics) in high school translates into expanded access to and successful completion of postsecondary education, especially in terms of bachelor's degree attainment.

To answer this question, we examined the effects of completing more rigorous courses in high school on access to and graduation from college, focusing on the mathematics curriculum. We investigated how a specific level of math course preparation in high school (completing Algebra II) may differentially affect a student's chances of attending and completing postsecondary education.[1] We used Florida's longitudinal student unit record data to examine the impact of math preparation in high school for six cohorts of students who were enrolled in the 7th through 12th grades in the 1995–96 school year and retrospectively tracked through the 2005–06 academic year. Florida has changed math course requirements for high school graduation (Reys, et al., 2007) but estimating the impact of this state-level policy change is beyond the scope of this study. Rather, herein we examine the relationship between high school mathematics course completion and subsequent educational outcomes, and provide evidence that failing to account for nonrandom assignment into high school courses may provide incorrect information to policymakers about this important relationship.

Literature Review

Many studies of access to college take high school preparation and graduation as given and studies of persistence and success in college tend to focus on what happens during the college years. However, it is likely that what happens in high school is inextricably linked to one's chances of entering and being successful in college. There is considerable evidence that academic preparation and performance in high school–typically measured in terms of high school course completion patterns, grades, and/or college entrance exams–is positively related to not only graduation from high school, but also enrollment in postsecondary institutions (Berkner & Chavez, 1997; Choy, 2002; Goldrick-Rab, Carter, & Wagner, 2007; Lee, Croninger, & Smith, 1997). Students who take a more rigorous high school curriculum, such as Advanced Placement (AP), honors, and other "college prep" courses (especially in math), are more likely to attend college and receive a postsecondary degree (Adelman, 2006; Altonji, 1995; Attewell & Domina, 2008; Horn & Kojaku, 2001; Klopfenstein & Thomas, 2009; Rose & Betts, 2001). Among college preparatory curricula, the empirical evidence indicates the role of math course-taking in promoting access to a bachelor's degree relative to other subjects. St. John and Chung's (2006) analysis using the National Education Longitudinal Study 1988 (NELS:88) reveals that taking advanced math courses significantly increases high school graduates' probability of attending a four-year college. Klopfenstein and Thomas (2009) investigate the effect of AP course-taking on early college grades and retention. They control for non-AP curricular experiences that existing studies often failed to consider (e.g., Dougherty, Mellor, & Jian, 2005; Geiser & Santelices, 2005) and find no significant effect of taking AP core courses on first-semester college GPA and retention to the sophomore year. These findings suggest that improved educational outcomes by students taking AP courses may be a consequence of self-selection, rather than the direct result of AP course completion. However, these results do not shed light on the effect of the high school curriculum taken by most students. Using data from the High School & Beyond surveys (HS&B:80), Rose and Betts (2001) explore the relationship between math courses completed in high school and college graduation. They find that completing higher-level math courses during high school significantly increases the probability of earning a bachelor's degree.

Many scholars argue that important barriers to higher education are erected by unequal opportunities for college preparatory classes while in high school and that these academic barriers are greater for minority and low-income students (Kao & Thompson, 2003; Terenzini, Cabrera, & Bernal, 2001). Research demonstrates that socioeconomically advantaged students are more likely to take rigorous math and science courses in high school (Cavanagh, Schiller, & Riegle-Crumb, 2006). On the contrary, access to college-preparatory courses (e.g., advanced math) may be restricted for low-income and underrepresented minority students because their academic opportunities are often structurally constrained by poor school conditions that impede the provision of rigorous and high quality academic courses and learning resources (Perna, 2004; St. John & Chung, 2006).

Importantly, much of the education research that examines the effects of high school academic preparation on college-level outcomes fails "to be sufficiently critical of issues of causality and selection bias, and methodological solutions to these concerns are often underutilized" (Goldrick-Rab, Carter, and Wagner, 2007, p. 2471). Although high school curriculum choices are likely to be endogenously related to post high-school education outcomes, many of the existing studies do not adequately account for such nonrandom selection of students into different high school curricular pathways (B. T. Long, 2007). Failure to account for this self-selection of students into curricular pathways may result in biased estimates of the effect of high school course taking on subsequent educational outcomes, thereby misleading policymakers.

Several studies have sought to account for the potential bias noted above by employing quasi-experimental methods such as a propensity score matching (Attewell & Domina, 2008) and instrumental variable (IV) approaches (Klopfenstein & Thomas, 2009; Rose & Betts, 2001). Attewell and Domina (2008) account for the nonrandom selection of curricular choices by employing a propensity score matching technique that attempts to ensure the equivalence between the "treated" (students taking a college-preparatory curriculum) and "control" groups (those who did not take a college prep curriculum) based on observable characteristics. Their results indicate that taking a more demanding high school curriculum is associated with greater math and reading achievement in high school, higher SAT test scores, and greater access to and graduation from college (relative to students who took a less rigorous curriculum). Their findings are consistent with the historical research documenting the positive effects of rigorous course completion on college enrollment, but the smaller effects they find suggest that studies that fail to control for observed differences between those who take a more rigorous curriculum and those who do not tend to produce upwardly biased estimates.

Instrumental variable (IV) approaches have also been used to study the effect of high school course completion on educational outcomes. An IV is a variable that is (conditionally) correlated with the treatment (e.g., highest level of high school math course completed) but uncorrelated with the error term in the outcome of interest (e.g., college attendance) equation. The IV approach can be used to isolate, net of other observed and unobserved factors, the effect of the treatment on the outcome. Altonji (1995) used the average number of courses taken by students in their high school as an instrument to examine the effect of an additional year of high school coursework on years of college education and postcollegiate wages. Using the National Longitudinal Survey of Youth (NLSY: 1979), his research indicates that an additional year of science, foreign language, and math during high school increases years of postsecondary education and future earnings. His results also reveal that the IV estimates are smaller than a "naïve" statistical model that inadequately accounts for self-selection into high school courses.

Rose and Betts (2001) use Altonji's instrument to investigate the effect of the *types* of math courses students take in high school on college graduation and labor market earnings. Their results indicate that the math curriculum has a strong effect: all types of math courses have statistically significant positive effects on college graduation, but completing more advanced math courses has a larger impact (than completing less advanced courses) on bachelor's degree completion and labor market outcomes.

Using data from the NLSY and the High School & Beyond surveys, Levine and Zimmerman (1995) examine the effect of math and science courses on college and labor market outcomes. They find that completing additional math courses during high school increases the chances of attending college (for men) and increases education and labor market outcomes for female college graduates, whereas the science curricula has very little impact for either gender. However, the effects are inconsistent across the two datasets, raising doubts about the existence of a common curricular impact on both education and labor market outcomes. Also, they find that the parameter estimates of the IV regression are not statistically significantly different than the "naïve" ordinary least squares (OLS) results. They acknowledge, however, that this finding is of little substance because the IV estimates are imprecisely estimated with standard errors substantially larger than the standard errors produced by the OLS models.

These studies are notable because they employ quasi-experimental methods to examine the effects of high school course taking on college and labor market outcomes. However, the instruments used in some of these studies (e.g., Altonji, 1995; Levine & Zimmerman, 1995) may not be valid. For example, Altonji's (1995) instrument is likely to be conditionally correlated with unobserved school characteristics that influence post-high school education outcomes. Moreover, these studies often focus on labor-market outcomes only, rather than on the educational processes related to high school completion, college access, and college completion (e.g., Joensen & Nielsen, 2009; Rose & Betts, 2004).

Notwithstanding the research mentioned above, there has been a dearth of evidence with regard to the direct effects of high-school curricular choices on postsecondary enrollment and degree completion. Our research begins to fill this gap by examining how the level of the math course completed in high school (e.g., Algebra II) influence students' chances of attending and graduating from college in the state of Florida. And we do so while accounting for the potential endogeneity of math course selection in high school on these postsecondary outcomes (Heckman, 1979). A better understanding of this relationship will be informative for state policymakers interested in promoting college access and success through curriculum reforms. If differences in high school curricular choices really do improve postsecondary outcomes, then state policies aimed at encouraging students to take a more rigorous curriculum may be an effective way of increasing access to and success in college, potentially improving students' labor market outcomes (Rose & Betts, 2001).

High School Math Requirements in Florida

In most states, the number and rigor of the mathematics courses completed in high school has been an important component of their high school graduation requirements. In recent years there has been a particular interest in Algebra II as a benchmark as a course that all students must complete. For example, in 19 states students must complete Algebra II for high school graduation.[2] And states that employ the Federal government's Common Core standards require all students to meet learning objectives at what is generally considered the Algebra II level, in order to be equipped with "algebraic thinking" skills.

The state of Florida is one such state that has increased its standards regarding math course completion in high school. Math is a major component of the requirements to receive a high school diploma, and the subject is particularly important in order to receive a special diploma establishing one's readiness for college. Since the 1997–98 academic year, the number of math credits required for high school graduation has increased (3 credits for the entering cohort of 1997–98 and thereafter, and 4 credits for the entering cohort of 2010–11 and thereafter). The required course level was also increased from Algebra I to Algebra I and Geometry (in the academic year of 2010–11), and to Algebra II (in the academic year of 2012–13) (Florida Department of Education, 2013).[3] Meanwhile, Algebra II was included as a requirement for the College Ready Diploma (College Preparatory Program) and qualification for state university admission (Florida Statutes, 1997; 2003; Dounay, 2006).

In April 2013, the Florida Senate approved Senate Bill 1076, which allows students to substitute courses, including Algebra II, with industry certifications. This decision provoked a debate about the math course requirements, particularly about the role of Algebra II. While similar debate is in progress in other states (e.g., Texas), knowing more about the impact of Algebra II on students' college attendance and degree attainment may be able to provide input to these important curriculum decisions happening in states.

Theoretical Framework

One theory that helps explain the relationship between high school course-taking and postsecondary outcomes is human capital theory. Human capital theory rests primarily on the hypothesis that more schooling increases the ability, productivity and, hence, wages of students who will enter into the labor market (Becker, 1993; Cohn & Geske, 1990). Completing an advanced high school curriculum may directly improve an individual's labor productivity. It is also likely to have an indirect impact by making further increases in labor productivity possible by improving a student's chance of being admitted to a four-year college and eventually earning a bachelor's degree (Rose & Betts, 2001). Once students are admitted to and attend college, students who completed higher-level courses in high school are less likely to be placed in remedial courses (Roth, Crans, Carter, Ariet, & Resnick, 2001). Avoiding developmental courses may keep students on track in terms of earning credits that count toward, and a GPA sufficient to graduate (Long & Boatman, 2013), thereby increasing the probability of obtaining a college degree (Berry, 2003).

Although human capital theory offers an explanation for how completing more rigorous courses in high school may increase college access and success, it is difficult to test empirically because individuals with higher unmeasured ability may simply be more likely to complete higher-level math courses in high school, attend a four-year college, and complete a bachelor's degree than individuals with low unmeasured ability. Untangling the actual contribution of a rigorous curriculum on postsecondary outcomes from that of students' innate ability (or other possibly confounding factors) depends on how well researchers account for the (potential) endogeneity between high school course completion and subsequent educational outcomes. Recognizing the challenges in estimating the effect of the high school curricular choices on college access and success, next we discuss the empirical approach we use to attempt to mitigate such self-selection (i.e., nonrandom assignment) problems.

Methodology

Research Questions

Expanding this line of research, we examined how the math courses students complete in high school affect their chances of going to college and obtaining a degree. Specifically, we intended to answer the following questions:

1. Does completing Algebra II in high school affect whether, and if so the type of college that a student attends?
2. Does completing Algebra II in high school affect whether, and if so the type of degree a student attains?

Data and Variables

In this study, we used data from Florida's student unit record (SUR) system, obtained from the Florida Department of Education's Educational Data Warehouse (EDW). This SUR data includes students' demographic, academic, and socioeconomic information as well as measures of the social, economic, financial, and institutional contexts these students faced during the observation period. Grade 7–12 and college enrollment information is available for individuals who have stayed in the Florida public education system. For individuals who work in Florida after leaving high school there is also information about their labor force participation and earnings. Although the coverage of the EDW is extensive with regard to the public education system, it does not include students who attend private institutions or postsecondary institutions outside of Florida, so they are not included in our data.

Our analytic sample included six cohorts of public school students (N = 758,456) who were enrolled in the 7th−12th grades during the

1995–1996 academic year. These students were (retrospectively) followed for up to 10 years in order to assess their educational progress over time. Students in 12th (7th) grade in 1995–1996 could have graduated from high school as early as 1996–1997 (2001–2002), and received their bachelor's degrees beginning in 1999–2000 (2004–2005).

In terms of educational outcomes, we examined both enrollment in and graduation from college, but different subsamples were used to study each of these outcomes. First, we analyzed the impact of completing Algebra II on the college enrollment margin (0 = no college, 1 = attended a two-year college, 2 = attended a four-year college). We also estimated the impact of completing Algebra II on the degree attainment margin (highest degree attained; where 0 = no degree, 1 = received an associate's degree, 2 = received a bachelor's degree) conditional on college attendance. Two different samples were used to study each of these outcomes. When estimating enrollment in college, the sample was restricted to students with high school records in the 12th grade (N = 615,185); when estimating college graduation, we further restricted the sample to students who attended college (N = 427,845). Because these samples were very large, missing values comprised a very small proportion of the data,[4] and appeared to be quite randomly distributed based on observable factors, we used listwise deletion.

The variable of interest was the level of mathematics courses completed in high school. This variable is binary and indicates whether a student passed Algebra II or higher (Trigonometry, pre-Calculus, and Calculus) (or not) during high school. We controlled for students' i) demographic (race/ethnicity, gender, free or reduced-priced lunch status, and language spoken at home) and financial characteristics (the total amount of financial aid an individual student received during college; included only in the degree attainment model); ii) academic experience (the number of AP and IB credits) and academic ability (SAT scores, high school GPA); and iii) cohort and school district differences, and institutional differences (in the degree attainment model). Descriptive statistics for each of the dependent and explanatory variables are included in Table 1.

The Statistical Model

Given the categorical nature of the outcomes, we estimated the effects of completing Algebra II on enrollment in and graduation from college using multinomial logistic regression (MNL). Employing MNL allowed us to investigate how each covariate, and in particular the Algebra II variable, affects the odds of choosing a particular outcome category. The general model is formally described in (1) below:

TABLE 1

Descriptive Statistics

Variable	All Students N (%)	Completed Alg II N (%)	Did Not Complete Alg II N (%)	Description
Female	318,758 (51.81)	192,482 (54.62)	126,276 (48.06)	1 if female, 0 if male
Race				1 if White/African American/Hispanic/Asian/Native American/Other, 0 otherwise
White	363,780 (59.13)	227,613 (64.59)	136,167 (51.82)	
African American	141,718 (23.04)	64,782 (18.38)	76,936 (29.28)	
Hispanic	92,138 (14.98)	46,644 (13.24)	45,494 (17.31)	
Asian	15,300 (2.49)	11,986 (3.40)	3,314 (1.26)	
Native American	1,307 (0.21)	823 (0.23)	484 (0.18)	
Other	942 (0.15)	573 (0.16)	369 (0.14)	
Language at home				1 if language spoken at home is English/Spanish/Haitian/Other/Missing, 0 otherwise
English	500,823 (81.41)	291,129 (82.61)	209,694 (79.80)	
Spanish	69,875 (11.36)	34,282 (9.73)	35,593 (13.55)	
Haitian	13,377 (2.17)	5,802 (1.65)	7,575 (2.88)	
Other	12,953 (2.11)	9,212 (2.61)	3,741 (1.42)	
Missing	18,157 (2.95)	11,996 (3.40)	6,161 (2.34)	
Reduced lunch	23,277 (3.78)	12,069 (3.42)	11,158 (4.25)	1 if received reduced-priced lunch during high school, 0 otherwise
Free lunch	147,090 (23.91)	60,498 (17.17)	86,592 (32.95)	1 if received free lunch during high school, 0 otherwise
High school GPA	2.67 (.65)	2.96 (.53)	2.34 (.60)	High School Grade Point Average
Attempted AP Credits	.39 (1.06)	.68 (1.36)	.06 (.33)	Attempted Advanced Placement credits during high school
Attempted IB Credits	.10 (.80)	.18 (1.07)	.01 (.22)	Attempted International Baccalaureate credits during high school
SAT Score				

(continued)

TABLE 1 *(continued)*

Descriptive Statistics

Variable	All Students N (%)	Completed Alg II N (%)	Did Not Complete Alg II N (%)	Description
SAT Score				
No SAT score	381,231 (61.97)	148,021 (42)	233,210 (88.75)	1 if no SAT score, 0 otherwise
400−599	1,339 (0.22)	598 (0.17)	741 (0.28)	1 if SAT score is in the range, 0 otherwise
600−799	16,724 (2.72)	10,987 (3.12)	5,737 (2.18)	
800−969	57,511 (9.35)	47,225 (13.40)	10,286 (3.91)	
970−1099	66,831 (10.86)	59,755 (16.96)	7,076 (2.69)	
1100−1269	62,304 (10.13)	57,818 (16.41)	4,486 (1.71)	
1270−1399	22,087 (3.59)	21,067 (5.98)	1,020 (0.39)	
1400−1600	7,158 (1.16)	6,950 (1.97)	208 (0.08)	
Total Amount Financial Aid ($1,000)[a]	5.204 (4.677)	5.773 (4.796)	2.871 (3.233)	Total amount of financial aid received during college
PSE attendance				
None	187,340 (30.45)	69,883 (19.83)	117,457 (44.70)	1 if never attended a postsecondary education institution, 0 otherwise
2-year institutions	249, 066 (40.49)	128,582 (36.49)	120,484 (45.85)	1 if ever attended two-year institution, 0 otherwise
4-year institutions	178,779 (29.06)	153,956 (43.69)	24,823 (9.45)	1 if ever attended four-year institution, 0 otherwise
Degree attainment[a]				
None	283,354 (66.23)	158,957 (56.26)	124,397 (85.61)	1 if never attained a postsecondary degree, 0 otherwise
Associate degree	42,345 (9.90)	33,060 (11.70)	9,285 (6.39)	1 if ever attained an Associate degree, 0 otherwise
Bachelors' degree	102,146 (23.87)	90,521 (32.04)	11,625 (8.00)	1 if ever attained a Bachelor's degree, 0 otherwise
Cohort				
7th grade	103,156 (16.78)	60,221 (17.10)	42,935 (16.35)	1 if in the grade in 1997−1998 academic year, 0 otherwise
8th grade	102,901 (16.74)	59,395 (16.87)	43,506 (16.57)	
9th grade	106,308 (17.29)	61,238 (17.39)	45,070 (17.17)	
10th grade	103,731 (16.87)	57,219 (16.25)	46,512 (17.71)	
11th grade	99,210 (16.14)	52,773 (14.99)	46,437 (17.69)	
12th grade	99,417 (16.17)	61,318 (17.41)	38,099 (14.51)	
Total	615,185	262,764	352,421	

[a]Conditioning on attending any postsecondary institutions ($N = 427{,}845$)

$$\ln \frac{\Pr(Y_i = \mathrm{m}|\mathrm{x})}{\Pr(Y_i = \mathrm{b}|\mathrm{x})} = \alpha_{\mathrm{m|b}} + \beta_{m|b}X_i + \gamma_{m|b}A_i + \delta_{m|b}Z_i + \varepsilon \text{ for m} = 1 \text{ to j} \tag{1}$$

where m indexes the j = 3 outcome categories for the enrollment and graduation margins; $\Pr(Y_i = \mathrm{b}|\mathrm{x})$ is the probability that student *i* will choose the baseline category (not attending college in the access model; no postsecondary degree in the attainment model); $\Pr(Y_i = \mathrm{m}|\mathrm{x})$ represents the probability that student *i* will choose one of the other categories (e.g., two-year or four-year for the college attendance margin; associate's degree or bachelor's degree and above for the degree attainment margin) compared to the baseline category.

The vector X_i represents a set of demographic and financial characteristics, such as gender, race, economic status (i.e., receiving free or reduced-priced lunch in secondary school), language spoken at home, and the total amount of financial aid received during college (included only in the degree attainment model). Because income data was not available, we used whether students receive free or reduced-priced lunch as a proxy for family income. As an indicator of one's socioeconomic status, researchers have traditionally used direct measures of parental income, education, or occupational status. However, the eligibility for a free or reduced-priced lunch may also be employed as an indirect indicator of socioeconomic status because eligibility for the free/reduced lunch program is determined by the federal poverty guideline based on family's gross income and the number of people in the family (Dynarski & Scott-Clayton, 2008; Heller & Rogers, 2006). We further distinguish students receiving *free* lunch from those students receiving a *reduced-priced* lunch (relative to students who do not participate in these programs).

The vector A_i represents an individual's academic experiences during high school including course-taking behavior (i.e., the number of AP and IB credits and highest level of mathematics course completed [1 = Algebra II or higher, 0 = less than Algebra II]) and academic ability (i.e., high school GPA and SAT score). The vector Z_i represents cohort and school district fixed effects to control for time-related and unobserved/unmeasured differences in the school district, and (in the degree attainment model) this vector also includes fixed effects to control for the college a student attended. The parameter estimates associated with *X, A,* and *Z* are δ, β, and γ, respectively.

The multinomial logit model can be derived from an additive random utility model (Cameron & Trivedi, 2005). In this model, the error terms associated with the utility for each choice are assumed have type 1 extreme value distributions that are independent of each other and independent of the all the explanatory variables. However, the Algebra II regressor is likely to be correlated with these errors since a student with high unmeasured ability may be more likely to complete Algebra II and receive different utility from post high-school educational alternatives than a student with low unmeasured ability. That is, after conditioning on their other characteristics, students' selection of high school math courses may not be independent of subsequent educational outcomes (i.e., enrollment in and graduation from college). Unless this endogenous relationship is adequately accounted for, the coefficient estimate on the Algebra II variable may be biased.

To mitigate this problem we employed an instrumental variable (IV) approach (Angrist & Krueger, 2001; Bielby, House, Flaster, & DesJardins, 2013; Stock & Trebbi, 2003). An instrument is a variable that is unrelated to the error term (ε) in (1) but, conditional on the other variables, is related to the endogenous variable (e.g., whether a student took Algebra II or higher in high school). A valid instrument identifies a source of exogenous variation and uses this variation to determine the impact of a treatment (e.g., Algebra II) on an outcome (e.g., college enrollment/graduation). IV estimation allows a researcher to make rigorous claims about the effect of the treatment on the outcomes (Angrist & Pischke, 2009) by minimizing bias due to endogeneity. Yet, the "cure can be worse than the disease" when the instrument is only weakly correlated with the endogenous variable (Bound, Jaeger, & Baker, 1993; 1995). Weak instruments may result in tests of significance with incorrect size and inaccurate confidence intervals, thereby (potentially) leading to incorrect statistical inferences. Weak instruments may also lead to estimates that are not consistent (Chao & Swanson, 2005) or are biased in the same direction as the naïve statistical approach that does not correct for nonrandom assignment issues.

The instrumental variable we used was the unemployment rate in the county where students lived during the 9th grade, information that was obtained from the Bureau of Labor Statistics' Local Area Unemployment Statistics. We also included the interaction of the unemployment rate with race and free/reduced-priced lunch status, which allowed the influence of the unemployment rate on Algebra II completion to differ by race and free- or reduced-priced lunch status. Our selection of this instrument was based on conceptual and empirical grounds. Conceptually, the selection of this instrument was based on a simple two-period model of time allocation. We assumed that students allocated their time between school, work, and leisure while enrolled in high school (period one)[5] and between work and leisure once they were in the post-education labor force (period two). We also assumed that students who allocate more time to schooling while in high school take more difficult courses.

In the following equations, h_s^t denotes how much time a student devotes to schooling in a period, h_w^t denotes how much time is allocated to work in period *t*, l^t denotes how much time is devoted to leisure activities in period *t*, and *t* is either 1 or 2 depending on whether we are referencing time period one or time period two.

To simplify matters, we assumed that school, work, and leisure comprise all of a student's time when individuals are in high school (period one, T^1), and work and leisure take up all of their time when individuals were in the post-education labor force (period two, T^2). So:

$$h_s^1 + h_w^1 + l^1 = T^1 \tag{2}$$

and

$$h_w^2 + l^2 = T^2 \tag{3}$$

We also assumed that a student's overall utility or well-being (U) was a function of their consumption of goods, *c*, and their leisure time, *l*. The amount of consumption in period *t* depended on their earnings in period t, equal to their hourly wage, w^t, multiplied by their hours of work:

$$h_w^t \times w^t = c^t \tag{4}$$

The human capital model implies that wages will depend on schooling, so we also assumed that wages in period two increase with the amount of time allocated to schooling in period one according to $w_h^2 = f(h_s^1)$. So a student will choose their hours of work and time devoted to schooling in such a way as to attempt to maximize their utility, *U,* across both periods:

$$max_{h_w^1 h_s^1 h_w^2} U\left(w^1 \times h_w^1, T^1 - h_w^1 - h_s^1\right) + \beta \times U\left(f\left(h_s^1\right) \times h_w^2, T^2 - h_w^2\right) \tag{5}$$

Here, β represents a discount factor that depreciates the value of future utility relative to current utility. Therefore, individuals attempt to obtain the highest overall combined utility in period one and discounted utility in period two. However, the utility obtained in period two depends on the discount factor and the amount of time allocated to schooling in period one.

An individual's time allocation in period one is not only determined by their personal discount factor, but also by exogenous factors such as the availability of work, which is reflected by w^1. Local labor market conditions when students are in high school may affect their college preparation decisions and in so doing affect college attendance and completion. For example, in a weak (strong) local labor market students may allocate less (more) time to work and more (less) time to study by increasing (decreasing) the quantity or difficulty of the courses that they take in high school.

Furthermore, we used the five assumptions associated with a valid IV as proposed by Angrist, Imbens, and Rubin (1996) to evaluate our selection of this instrument: 1) stable unit treatment value assumption, 2) random assignment, 3) exclusion restriction, 4) nonzero average causal effect of the instrument on the treatment, and 5) monotonicity. The *stable unit treatment value assumption* (SUTVA) requires that an individual student's Algebra II completion does not influence other students' college outcomes (e.g., spillover effects). It is unlikely that one student completing Algebra II (or not) will impact another student's decision to attend or complete college. However, students in the unit record data are clustered within high schools, so it is worth addressing possible threats to this assumption. Students who do not complete Algebra II might be affected by a college going/prep-culture, sharing instructors, by interactions among Algebra II instructors and other teachers, or by Algebra II content being covered in lower-level math classes. Schools with a college going culture may offer more academic and administrative supports for students, which may induce more students to take advanced courses and prepare for college (Klugman, 2012). Yet, students' decision for college attendance starts at early stage, often before sophomore year, followed by enrollment in college-bound curriculum (Cabrera & Nasa, 2000; Kinzie et al., 2004). Therefore, students who did not complete Algebra II still could be encouraged to pursue postsecondary education, but those will be a very small number. Furthermore, it is hard to imagine that one student's algebraic knowledge is transmitted to another through peer interaction. In addition, high school curricula are hierarchically structured, particularly for mathematics, and teaching Algebra II without prerequisite material would be difficult. Thus, Algebra II content is less likely to be covered in lower-level classes, and this practice would not significantly vary across teachers. Thus, we believed the risk of violating this assumption was low. This assumption also required that the treatment was consistent across all treated groups. Given that Algebra II content/curricula may vary from school district to school district or year-by-year, we included school district fixed effects by year to address this concern.

The second assumption is about *random assignment* which required that the distribution of the instrumental variable across individuals be comparable to what would be the case under random assignment. This means that any individual in the sample must have an equal probability of having any level of the instrumental variable, which may not be the case if students/their families move to find lower unemployment rates. We believe our IV satisfies this assumption for following reasons. First, 9th graders are highly unlikely to travel across county lines for employment or change their residence to obtain employment in a county with a lower unemployment rate which would invalidate the use of a home address-based unemployment rate as an IV. It is also unlikely that parents will choose to move their students when they are in 9th grade: residential mobility across counties decreases as children get older (L. H. Long, 1972). Second, the geographical characteristics of the state provide indirect evidence to support our argument. Florida's counties are large (average land area is 804 sq. miles), and none of Florida's four largest cities (Miami, Tampa, Orlando, and Jacksonville) are located at a county border, although Miami is about 30 miles from Ft. Lauderdale, which is in Broward County. Finally, local labor market conditions (and other factors related to residential choice) change independently of students' course taking/completion, college going, and degree completion behavior.

The *exclusion restriction assumption* ensures that an IV affects the dependent variable (e.g., enrollment/graduation from college) only through its relationship with the endogenous independent variable of interest (e.g., Algebra II). To satisfy this assumption, the instrument must not be correlated with the error term in (1). Our instrument variable satisfies this assumption in two ways. First, we used county level unemployment rates when students were in the 9th grade. It is arguable that the local unemployment rate changes while students are in high school, affects students' and parents' financial status and ability to save, and thus affects students' decision about attending college and persisting through graduation. However, financial conditions for high school students tend to be stable. Local unemployment rates when students are in 9th–12th grades are highly correlated (r = 0.895). For the years that are included in our analyses (1997–2005), the state unemployment rate in Florida was stable, ranging from 3.8% in 2001 to 5.7% in 2003. Therefore, students in our sample did not face financial circumstances that made them likely to change their academic plans dramatically. Nonetheless, to remove any potential correlation between the 9th grade unemployment rate and outcomes, we also included the unemployment rate at the time when a student was in 12th grade as a control in the outcome equation. Second, we conducted an over-identification test providing evidence that the second-stage residuals are uncorrelated with the IV. This test assumes that the first IV (county-level unemployment rate) is properly excluded from the second-stage equation. Thus, satisfying this assumption means failing to reject the null hypothesis that instruments are correctly excluded from the second stage estimation of the dependent variable. The Hansen J statistic[6] results indicate that the null hypothesis is rejected both for the college attendance equation (x^2 = 26.46, p = 0.0002) and for the degree attainment equation (x^2 =17.92, p = 0.0064). These tests are valid, however, only when the treatment has a homogenous effect on outcomes, which is unlikely to be the case in this study.[7] We also tested the redundancy of the IVs by comparing the results when each instrument is used separately versus the results when all IVs are included.[8] If the parameter estimates using different instruments differ appreciably and significantly from one another, the validity of the instruments is suspect. If all of the estimates are consonant with a single interpretation of the data, the credibility for instruments is enhanced (Murray, 2006). We estimated the attendance and graduation models using only one instrument–9th grade unemployment rate. The estimated results were similar for both outcomes compared to our specification that included 9th grade unemployment rate and its interactions with race and free-and reduced-priced lunch status.[9]

The fourth assumption is *nonzero average causal effect of the instrument on the treatment.* This assumption requires that there be a strong relationship between the instrumental variable(s) and the endogenous regressor (Algebra II). Although there is no formal statistical test to verify the quality of the IV, some test statistics can provide guidance. Often, first stage F-statistics and the partial R^2 have been used (Bound, Jaeger, & Baker, 1995); However, there are tests now available that are improvements on these statistics (see Baum, 2008; see also Angist & Pischke, 2009) and are more appropriate for judging whether there is a weak instrument problem. For example, the Cragg-Donald Wald F-statistic test is now often employed to assess whether there are violations to this assumption. However, this test assumes independent and identically distributed standard errors (the i.i.d assumption) and is invalid when the analyst employs cluster-robust standard errors, as is the case in the models we estimated. Thus, we employed the Kleibergen-Paap (K-P) test, which provides evidence about the weak instrument issue when the i.i.d. assumption is violated (see Baum, 2008). Although there is no formal critical value associated with the K-P test, both the college attendance (F = 167.16) and degree attainment model (F = 56.05) K-P test results exceed the rule-of-thumb critical value of 10 (Baum, 2008) and the maximum value of 19.86 (Stock & Yogo, 2005) for models like ours. Thus, we believe any bias that could be introduced by weak/invalid instruments is mitigated.

Finally, *monotonicity* assumes that the IV has a unidirectional effect on the endogenous variable. That is, the relationship between the endogenous independent variable (Algebra II) and the instrumental variable (local unemployment rate) should be either positive or negative for all students in the sample. In our context this meant that increases in the unemployment rate should never result in decreases in math course taking. Our instrument is unlikely to satisfy this assumption fully, as there may be students who choose not to take Algebra II when unemployment rates are rising because they feel they need to devote additional time to searching for work when unemployment rates are high. However, we believe this set of students is likely to represent a very small fraction of our sample, and any presence of these "defiers" simply places an upward bound on our estimate of the treatment effect (Angrist & Pischke, 2009; Porter, 2012). Because defiers act in contradiction to the expected influence of the instrument, the estimated relationship between the instrument and the endogenous variable was expected to be in the opposite direction to that of compliers (i.e., negative influence of unemployment on Algebra II). Mathematically, if we were to combine the estimated effects for each individual, the opposing signs would simply push the average effects of the instruments toward zero. As long as compliers outnumber defiers in our sample, we are able to obtain at least a lower bound estimate of the causal effect of mathematics course taking on postsecondary enrollment and graduation. Moreover, nonparametric estimates we conducted of the effect of the unemployment rate on the Algebra II variable were generally positive throughout the range of unemployment rates, providing indication of monotonicity.

It is important to note that the IV estimates are local to the students whose math course choices are shifted by the unemployment rate in the county where students lived during the 9th grade (a "Local Average Treatment Effect" or LATE). Therefore, the analysis and the interpretation of the findings hold for students in the sample who were induced to complete Algebra II by a change in the local unemployment rate. Those students were likely to be individuals at the margin of ether working or taking/completing Algebra II and were then induced to take Algebra II by an increase in the unemployment rate. Such students were more likely to come from working class families (Kane & Rouse, 1999; Betts & McFarland, 1995) compared to their counterparts from high-income families who were likely to be “always-takers” in the sense that they would take Algebra II regardless of the level of the local unemployment rate.

We implemented the IV method discussed above in two stages. In the first stage we regressed a binary variable Algebra II (T) on a full set of covariates (X) as well as the 9th grade county level unemployment rate and its interactions with race and free lunch status (Z), using a linear probability specification.

$$T = \gamma + \beta X + \theta Z + \omega \tag{6}$$

From our estimates of (6) we calculated and stored the residuals $\hat{\omega}$.

In stage two we estimated a multinomial logit model where the categories for the attendance and enrollment margins were the same as in (1) on a full set of controls (X; discussed above) and the residuals from the stage one equation (6) that accounts for the endogeneity of Algebra II in this structure.

$$\ln \frac{\Pr(Y_i = \mathrm{m} \mid \mathrm{x})}{\Pr(Y_i = \mathrm{b} \mid \mathrm{x})} = \alpha_{m|b} + \beta_{m|b} X_i + \gamma_{m|b} A_i + \delta_{m|b} Z_i + \rho_{m|b} \hat{\omega}_i \tag{7}$$

Correcting for endogeneity using this “control function” approach should result in the estimate of δ that more accurately approximates the causal influence of T on Y than when employing the “naïve” statistical model that does not account for self-selection [as formalized in (1) above].[10] In addition, we also bootstrapped the entire two-step process 200 times to account for the estimation of the residuals in stage one and account for any uncertainty introduced by this two-stage estimation process. For example, bootstrapping this two stage structure only one time would not account for sources of uncertainty introduced by the estimation of the first stage. Failure to account for such uncertainty may results in underestimates of the coefficient standard errors, overestimates of the significance of the regressors, and incorrect inferences about their effects on enrollment and graduation.

To test for any differences in the naïve and IV model we estimated (1) and compared the results to those obtained from (7). In particular we were interested in any differences in the effect of the policy variable of interest, Algebra II, on college enrollment and graduation. To ease comparison of the estimates of this important variable across model specifications, the Algebra II coefficients were reported as average marginal effects (AMEs). All other estimates were reported as odds ratios because it was computationally prohibitive (due to the large sample sizes), to bootstrap the marginal effects for all the other regressors included in the enrollment and graduation models.

Limitations

This study has a number of limitations. First, although the postsecondary attainment process begins when students make a transition to college and is over when they graduate from college, we considered college enrollment and completion as distinct outcomes and model these outcomes separately. College students make transitions across a set of discrete states, and these states include being enrolled in one's institution, interrupting one's enrollment in the institution (i.e., stopout), dropout, and graduation. Future research is thus warranted to jointly account for this interrelationship in students' academic pathways using longitudinal modeling methods.

Second, we exclusively focused on how mathematics course completion in high school was related to postsecondary outcomes. Given the variety of curricular choices other than math curriculum open to high school students it would be interesting to investigate the association between other subjects (e.g., science, English) and postsecondary outcomes. We intend to do so at a later date.

Ideally, we would have preferred to estimate two-stage models where both the selection (equation 6) and outcome (equation 7) regressions include a multinomial outcome. We simplified the selection equation by dichotomizing the math course treatment into whether a student completed Algebra II or not. We acknowledge that this strategy resulted in a loss of information which might introduce a potential bias in our findings. Although we treated the effect of the different levels of math courses to be the same within the Algebra II (= 1) or not (= 0) groups, the effects may differ within these groups. For example, the effects of math course completion may "flatten out" at certain math course levels (MacCallum, Zhang, Preacher, & Rucker, 2002). By treating students who completed Algebra II *only* and those who completed Algebra II *or higher* the same, we may have introduced a potential upward bias in our estimates.

In our data we have the full range of high school math courses students take so we could estimate the first stage using these multiple categories as the dependent variable. However, we focused on Algebra II for two reasons. First, the national policy debate on math course taking has centered on the effect of completing Algebra II, rather than what course level matters for subsequent educational outcomes. The goal of this study was to test the former and not necessarily the latter. Second, to our best knowledge, there is no statistical program that will permit the estimation of both the first- and second-stage with multicategorical dependent variables. We are currently working to develop statistical software that will enable us to estimate such a model.

Results

Course Completion and the Type of College Attended

As noted above, we first estimated a "naïve" MNL regression of college attendance on completing Algebra II. Although this regression also controlled for factors such as demographic and academic characteristics of students, it ignored the possible endogeneity between students' math course selection and college attendance. We compared all other results produced by the two-stage IV method describe above to this model in order to ascertain the extent of any bias due to not accounting for the (potential) endogeneity of Algebra II on enrollment in college.

The results of the naïve model, depicted in columns (1) through (3) in Table 2 indicate that completing Algebra II increases the chances of attending a postsecondary institution, in particular of attending a four-year college. Holding all other characteristics constant, completing Algebra II increases the probability of attending a four-year college by 20.6%. For students who completed Algebra II, the probability of not going to college (attending a two-year college) is 18.5% (2.1%) lower than their counterparts who did not complete Algebra II. Again, the coefficients produced by this naïve model may be biased due to nonrandom selection into math courses in high school.

To account for any potential endogeneity, the second set of results are from the "control function" (i.e., two stage) IV approach described above. Columns (4) through (6) in Table 2 indicate that the estimated effect of Algebra II produced by the IV approach is different from that of the naïve statistical model. Whereas the naïve model suggests a negative effect of completing Algebra II on two-year college enrollment, the IV model results indicate a positive effect. Holding other variables constant, completing Algebra II increases the probability of two-year college attendance by 27.6%. In addition, the estimated effect of Algebra II on four-year college attendance is about 0.3 percentage points lower for the IV than the naïve model results, and the effect is statistically insignificant. In terms of not attending college, completing Algebra II results

TABLE 2

Estimated Effects of Math Course Completion on College Attendance

Variables	Naive: No College (Base category) (1)	Naive: 2-year (2)	Naive: 4-year (3)	Control Function (IV): No College (Base category) (4)	Control Function (IV): 2-year (5)	Control Function (IV): 4-year (6)
Algebra 2 or higher		1.649***	2.973***		1.956	10.80***
		(0.034)	(0.102)		(0.920)	(6.532)
In marginal effects	−0.185***	−0.021**	0.206***	−0.479***	0.276***	0.203
	(0.006)	(0.007)	(0.005)	(0.183)	(0.088)	(0.163)
Female		1.543***	1.628***		1.525***	1.645***
		(0.016)	(0.020)		(0.017)	(0.020)
Race						
African American		0.858***	1.231***		0.855**	1.385***
		(0.030)	(0.073)		(0.049)	(0.106)
Hispanic		1.123***	1.207***		0.989	1.199***
		(0.022)	(0.037)		(0.025)	(0.030)
Asian American		1.121***	1.477***		1.149***	1.499***
		(0.033)	(0.047)		(0.035)	(0.052)
Other race		1.059	1.159*		1.015	1.147
		(0.061)	(0.087)		(0.075)	(0.112)
Reduced lunch		0.792***	0.603***		0.799***	0.587***
		(0.021)	(0.026)		(0.024)	(0.030)
Free lunch		0.665***	0.489***		0.677***	0.508***
		(0.016)	(0.018)		(0.029)	(0.035)
Language at home						

(continued)

TABLE 2 *(continued)*

Estimated Effects of Math Course Completion on College Attendance

Variables	Naive			Control Function (IV)		
	No College (Base category) (1)	2-year (2)	4-year (3)	No College (Base category) (4)	2-year (5)	4-year (6)
Language at home						
Spanish		1.086**	1.375***		1.016	1.489***
		(0.034)	(0.068)		(0.048)	(0.078)
Haitian		1.354***	1.530***		1.090	1.486***
		(0.094)	(0.126)		(0.102)	(0.121)
Other language		0.970	1.202***		0.837**	1.124
		(0.034)	(0.059)		(0.052)	(0.085)
Language missing		1.036	1.111		1.275	1.038
		(0.046)	(0.071)		(0.213)	(0.173)
High school GPA		1.031*	1.950***		1.012	1.375*
		(0.015)	(0.045)		(0.092)	(0.176)
Attempted AP credits		0.751***	0.845***		0.767***	0.856***
		(0.010)	(0.017)		(0.022)	(0.029)
Attempted IB credits		0.801***	0.904***		0.817***	0.904***
		(0.009)	(0.009)		(0.010)	(0.011)

TABLE 2 *(continued)*

Estimated Effects of Math Course Completion on College Attendance

Variables	Naive			Control Function (IV)		
	No College (Base category) (1)	2-year (2)	4-year (3)	No College (Base category) (4)	2-year (5)	4-year (6)
SAT score						
SAT score 400−599		4.692***	10.277***		4.319***	9.169***
		(0.479)	(1.077)		(0.440)	(0.992)
SAT score 600−799		4.062***	14.765***		3.618***	10.30***
		(0.205)	(0.888)		(0.461)	(1.714)
SAT score 800−969		3.744***	23.697***		3.286***	14.29***
		(0.169)	(1.118)		(0.584)	(3.252)
SAT score 970−1099		3.217***	34.915***		2.823***	20.58***
		(0.131)	(1.414)		(0.531)	(5.231)
SAT score 1100−1269		2.385***	32.703***		2.080***	19.75***
		(0.103)	(1.386)		(0.353)	(4.616)
SAT score 1270−1399		1.890***	24.248***		1.594***	15.31***
		(0.106)	(1.360)		(0.221)	(2.815)
SAT score 1400−1600		1.838***	11.058***		1.506**	7.427***
		(0.129)	(0.730)		(0.209)	(1.130)
Unemployment rate in 12th grade		1.017	1.011		1.064***	1.037*
		(0.010)	(0.012)		(0.018)	(0.017)
Number of observations	556,676			556,676		

Note. Exponentiated coefficients are reported; standard errors in parentheses. Cohort & District Fixed Effects included but not displayed. Clustered standard errors in outcome equations. Control Function model estimates are bootstrapped (200 reps).
$*p < .05$. $**p < .01$. $***p < .001$.

in a lower probability of not going to college by about 48%. The results indicate that for statistically similar students, completing Algebra II in high school dramatically increases their chances of college attendance in general and is also related to the *type* of postsecondary institutions they attend.

Course Completion and the Type of Postsecondary Degree Attained

Next we present results (see Table 3) illustrating whether the level of math courses students complete in high school affects their probability of obtaining (or not) different postsecondary (e.g., associate's or bachelor's) degrees. The naïve multinomial model results indicate that completing Algebra II is positively associated with degree attainment: holding other characteristics constant, completing Algebra II increases the probability of associate's and bachelor's degree attainment by 2.1 and 5.6%, respectively. However, students who complete Algebra II in high school have a lower probability of not receiving a postsecondary degree (about 8%), compared to their counterparts who did not complete Algebra II.

The IV results tell a different story. The positive association between Algebra II and associate's degree attainment found in the naïve model is insignificant when controlling for student self-selection into Algebra II. When employing the IV estimation method, completing Algebra II increases the probability of bachelors' degree attainment (by about 20%), but this result is not statistically significant at the .05 level. In addition, the significant effect of completing Algebra II on the probability of obtaining an associate degree or not obtaining a degree that was evident in the naïve model is not significant once selection into math courses is taken into account.

Overall, our results suggest that completing Algebra II in high school has a positive effect on postsecondary education attainment of any kind, but the effect is only significant for two-year attendance, rather than four-year college attendance. Conditioning on college attendance, we found no Algebra II effect on degree attainment. Finally, estimates from the naïve model diverge from those from the IV models in direction and magnitude, often overestimating the impact of Algebra II. This suggests the importance of accounting for selection bias in order to estimate the actual effect of Algebra II and to provide accurate policy implications.

TABLE 3

Estimated Effects of Math Course Completion on College Degree

	Naive			Control Function (IV)		
Variables	No Degree (Base category) (1)	AA (2)	BA (3)	No Degree (Base category) (4)	AA (5)	BA (6)
Algebra 2 or higher		1.479***	1.473***		0.413	2.565
		(0.040)	(0.033)		(0.570)	(4.216)
In marginal effects	–0.078***	0.021***	0.056***	–0.026	-0.174	0.200
	(0.005)	(0.003)	(0.004)	(0.173)	(0.152)	(0.164)
Female		1.099***	1.204***		1.072	1.231***
		(0.018)	(0.019)		(0.040)	(0.056)
Race						
African American		0.517***	0.866		0.526***	0.926
		(0.023)	(0.070)		(0.019)	(0.039)
Hispanic		0.963	0.980		0.943	1.034
		(0.035)	(0.037)		(0.034)	(0.033)
Asian American		0.922	1.039		0.957	1.057
		(0.038)	(0.060)		(0.047)	(0.054)
Other race		0.920	0.924		0.915	0.938
		(0.106)	(0.099)		(0.103)	(0.091)
Reduced lunch		0.899**	0.658***		0.888**	0.631***
		(0.034)	(0.039)		(0.036)	(0.025)
Free lunch		0.813***	0.632***		0.789***	0.618***

(continued)

TABLE 3 *(continued)*

Estimated Effects of Math Course Completion on College Degree

Variables	Naive			Control Function (IV)		
	No Degree (Base category) (1)	AA (2)	BA (3)	No Degree (Base category) (4)	AA (5)	BA (6)
Language at home						
Spanish		1.173*	1.225**		1.165**	1.296***
		(0.086)	(0.078)		(0.058)	(0.055)
Haitian		1.703***	1.644***		1.651***	1.807***
		(0.098)	(0.070)		(0.140)	(0.157)
Other language		1.276***	1.240***		1.286***	1.300***
		(0.077)	(0.074)		(0.092)	(0.091)
Language missing		1.166**	1.016		1.286**	0.947
		(0.055)	(0.046)		(0.103)	(0.121)
High school GPA		1.454***	2.976***		1.766**	2.475***
		(0.073)	(0.063)		(0.358)	(0.591)
Attempted AP credits		0.897***	1.036		0.921**	1.048
		(0.011)	(0.029)		(0.028)	(0.035)
Attempted IB credits		0.961*	1.061***		0.967*	1.054***
		(0.015)	(0.010)		(0.014)	(0.012)

TABLE 3 *(continued)*

Estimated Effects of Math Course Completion on College Degree

Variables	Naive			Control Function (IV)		
	No Degree (Base category) (1)	AA (2)	BA (3)	No Degree (Base category) (4)	AA (5)	BA (6)
SAT score						
SAT score 400−599		1.203*	1.185		1.152	1.229
		(0.107)	(0.199)		(0.156)	(0.207)
SAT score 600−799		1.771***	2.321***		2.130**	2.159**
		(0.081)	(0.152)		(0.500)	(0.586)
SAT score 800−969		1.929***	3.259***		2.706*	2.856*
		(0.045)	(0.111)		(1.058)	(1.322)
SAT score 970−1099		1.641***	3.721***		2.398*	3.278*
		(0.066)	(0.203)		(1.042)	(1.686)
SAT score 1100−1269		1.317***	3.744***		1.892	3.364*
		(0.087)	(0.166)		(0.777)	(1.656)
SAT score 1270−1399		0.896	3.220***		1.225	2.975*
		(0.063)	(0.200)		(0.435)	(1.314)
SAT score 1400−1600		0.536***	2.062***		0.708	1.971
		(0.067)	(0.169)		(0.245)	(0.800)
Total amount financial aid (‘000)		1.021*	1.089***		1.022***	1.091***
		(0.009)	(0.019)		(0.005)	(0.008)
Unemployment rate in 12th grade		1.005	1.001		1.010	0.983
		(0.014)	(0.023)		(0.013)	(0.017)
Number of observations	186,598			186,598		

Note. Exponentiated coefficients are reported; standard errors in parentheses. Cohort & District Fixed Effects included but not displayed. Clustered standard errors in outcome equations. Control Function model estimates are bootstrapped (200 reps).

*$p < .05$. **$p < .01$. ***$p < .001$.

Discussion and Conclusions

The political debate over high school mathematics requirement has been focused on two main questions. First, "does completing Algebra II in high school matter for college access and success?" The particular interest in Algebra II is based on the assumption that it is a "gatekeeper" course in terms of college readiness and success not only because colleges require Algebra II, but also because algebraic thinking is thought to prepare students for college entrance and college level classes, which in turn increases the chances of obtaining a degree (Rose & Betts, 2001). Previous studies have supported this idea: all types of math courses have statistically significant positive effects on college enrollment (Altonji, 1995; Levine & Zimmerman, 1995) as well as college graduation, and the impact is larger for higher level math courses (Long et al., 2012; Rose & Betts, 2001). However, the results from our study suggest a somewhat different story. For students who graduated from high schools in the Florida public system between 1996–97 and 2001–02 and who are likely to make course-taking choices that are associated with local labor market conditions, completing Algebra II only matters for two-year college attendance, but not for four-year attendance and degree attainment. The finding is significant since students to whom the results apply are the very students we might be most interested in helping to access and complete college.

This finding is in line with the recent national trend in Algebra II completion and high school math achievement. Using the data from the national assessment of educational progress (NAEP), Loveless (2013) demonstrated that despite the increasing completion rates of Algebra II, the average math scores of 17 year olds who have taken Algebra II or Advanced Algebra fell by 10 points between 1992 and 2012. He concluded that about one-third of students who complete a college preparatory curriculum are ill-prepared for college course work, and this trend reflects the watered down effect of Algebra II over time. However, to fully understand the insignificant impact of completing Algebra II on four-year attendance and degree attainment, we need to know what is taught in Algebra II. Although graduation requirements in many states, including Florida, are increased to align with national standards (e.g., Common Core standards), having Algebra II requirements does not guarantee that all states, districts, or schools would cover the same level of concepts in Algebra II classes. If concepts that are supposed to be covered in Algebra I are taught as Algebra II in many districts, we may observe a diminishing impact of Algebra II on subsequent outcomes.

While the intensity and pattern of coursework completed in high school are considered to be the best indicators of academic performance and eventual graduation from college (Adelman, 1999), the insignificant effect of Algebra II on four-year attendance and degree attainment raises questions about what level of math courses *are* sufficient to equip students with the requisite knowledge and skills that improve their chances of attaining a degree. For example, Geometry may increase a student's chances of attaining an associate's degree, whereas Trigonometry or Calculus may increase the chances of bachelor's degree attainment. Therefore, the national attention to Algebra II needs to be expanded to consider the impact of different levels of math courses. To support this, future research needs to examine the effect of different levels of math courses on students' access to, choice, and completion of college. We intend to do just that, and develop a statistical program that will permit the estimation of a multicategorical outcome in an IV framework.

The second question raised in the policy debate is whether Algebra II should be required for all students. The major debate about this question is whether Algebra II has similar effects for college- and career-bound students. The result of our study suggests that the policy debate should also consider the difference between students who pursue associate's degrees and students who pursue bachelor's degrees. Given the increased interest in state policy to encourage transfer and bachelor's degree attainment (Wellman, 2002), the effect of high school math course completion for different postsecondary pathways also deserves an in-depth inquiry.

From a methodological standpoint, this study indicates that the failure to account for selection bias is one of the causes of the "fallacy of misplaced causation" (Baker, 2013). The comparison between the naïve and control function models indicates that without taking account of selection bias, we would overestimate the impact of advanced math courses on subsequent outcomes. Relying on those results may well then lead to an incorrect policy response.

Finally, future research needs to investigate the effect of completing advanced math courses in high school on college outcomes across racial and/or socioeconomic status (examine "heterogeneous treatment effects"). Our study results indicate that even if students of different races take advanced math courses in high school, there are differential effects on outcomes by race and socioeconomic background. Given that African Americans, Hispanics, and students with low-socioeconomic backgrounds tend to be concentrated in schools with lower levels of social, economic, and academic resources (Perna & Titus, 2005), even among those who complete advanced math courses the quality of the knowledge obtained may be lower for these underrepresented students compared to their counterparts from schools with more resources. Therefore, it may be beneficial to target academic support for math and other subjects to students in under-resourced high schools to maximize their access to college and degree attainment.

In addition, the effect of math course-taking on degree attainment may vary depending on what majors students take. For example, levels of math courses taken in high school may matter more for students in an engineering field compared to their counterparts in the liberal arts. Unfortunately, we are not able to test this proposition due to a lack of data regarding college majors in our Florida data set.

We hope that our study informs individuals responsible for instituting policies that promote postsecondary educational outcomes through high school curriculum reforms. One interesting avenue for future research is to examine whether or not differences in math curriculum choices in high school eventually affect individuals' labor market outcomes. This line of inquiry will be possible once linkages between the longitudinal student unit record data used for this study and employment data are obtained.

Notes

This research was supported by the Spencer Foundation's Education and Social Opportunities Major Grants Program (Grant Number 201000120).

[1] Our definition of level of math course completion is dichotomous. For example, "Algebra II" means the student completed Algebra II *or higher,* "Trigonometry" means the student completed Trigonometry *or higher,* etc.

[2] Arizona, Minnesota, New Mexico, North Carolina, Ohio, and Tennessee recently adopted and started implementing Algebra II requirement.

[3] The ordering of math courses in Florida is: pre-Algebra, Algebra I, Geometry, Algebra II, Trigonometry, pre-Calculus, and Calculus.

[4] The variable that has highest missing rate is the highest level of math courses completed (9.19%).

[5] Although compulsory schooling is required up to age 16, high school students are allowed to work up to 15 (age 14 & 15) to 30 (age 16 & 17) hours per week when school is in session according to the child labor law of the State of Florida.

[6] The Hansen J statistic is the most common test statistic for the validity of the instruments, testing whether the instruments are uncorrelated with the errors. The significant statistic indicates that one or more of our instruments are not valid (assuming that the model is otherwise correctly specified).

[7] Maintaining the hypothesis that all instruments in an overidentified model are valid, the traditional overidentification test statistic becomes a formal test for treatment-effect heterogeneity (Angrist & Pischke, 2009).

[8] This is based on tests of instrument redundancy available in Stata's ivreg2 command. The test is needed because including redundant instruments in overidentified

models (include more IVs than endogenous regressors), will tend to bias point estimates (Angrist & Pischke, 2009). The significant statistic indicates that the specified instruments are redundant.

[9] Specifically, the results were similar for the attendance equation; there were some differences in the estimated graduation (point estimate of BA degree became negative) but the effect was not statistically significant.

[10] The control functions (CF) approach is more robust than alternatives, such as using propensity scores (predicted values only) as controls in the outcome equation. The CF approach allows any unobservables affecting the dependent variable (Y) to be dependent on X (Algebra II) while controlling for both the covariates (Z) and any instrumental variables. Importantly, the CF approach explicitly models this dependence whereas alternative methods fail to incorporate such dependence (see Heckman & Navarro-Lozano, 2004, for details).

References

ACT. (2012). *The reality of college readiness: National.* Iowa City, IA: ACT, Inc.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, DC: U.S. Department of Education.

Altonji, J. G. (1995). The effects of high school curriculum on educational and labor market outcomes. *The Journal of Human Resources, 30*(3), 409–438.

American College Test (2004). *Crisis at the core: Preparing all students for college and work.* Iowa City, IA.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444–455.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives, 15*(4), 69–85.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics.* Princeton, NJ: Princeton University Press.

Attewell, P., & Domina, T. (2008). Raising the bar: Curricular intensity and academic performance. *Educational Evaluation and Policy Analysis, 30*(1), 51–71.

Baker, N. (2013). Wrong answer: The case against Algebra II. *Harper's Magazine,* September 2013. Retrieved from http://harpers.org/archive/2013/09/wrong-answer/

Baum, C. F. (2008, June). Instrumental variables and panel data methods in economics and finance. Proceedings from 2008 German Stata users group meeting. Berlin, Germany: STATA.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). Chicago: The University of Chicago Press.

Berkner, L., & Chavez, L. (1997). *Access to postsecondary education for 1992 high school graduates.* Washington, DC: U.S. Department of Education.

Betts, J. R., & McFarland, L. L. (1995). Safe port in a storm: The impact of labor market conditions on community college enrollments. *Journal of Human Resources, 30*(4), 741–765.

Berry, L. (2003). Bridging the gap: A community college and area high schools collaborate to improve student success in college. *Community College Journal of Research and Practice. 27*, 393–407.

Bielby, R. M., House, E., Flaster, A., & DesJardins, S. L. (2013). Instrumental variables: Conceptual issues and an application considering high school course taking. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 263–321). New York: Springer.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association, 90*(430), 443–450.

Cabrera, A. F., & La Nasa, S. M. (2000). Understanding the college-choice process. *New Directions for Institutional Research, 107*, 5–22.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications.* Cambridge, UK: Cambridge University Press.

Cavanagh, S. E., Schiller, K. S., & Riegle-Crumb, C. (2006). Marital transitions, parenting, and schooling: Exploring the link between family-structure history and adolescents' academic status. *Sociology of Education, 79*(4), 329–354.

Chao, J. C., & Swanson, N. R. (2005). Consistent estimation with a large number of weak instruments. *Econometrica, Econometric Society, 73*(5), 1673–1692.

Choy, S. P. (2002). *Access & persistence: Findings from 10 years of longitudinal research on students.* Washington, DC: American Council on Education.

Cohn, E., & Geske, T. G. (1990). *The economics of education* (3rd ed.). Oxford, UK: Pergamon Press.

Dougherty, C., Mellor, L., & Jian, S. (2005). *The relationship between Advanced Placement and college graduation.* Austin, TX: National Center for Educational Accountability.

Dounay, J. (2006). *Policy brief: High school – Advanced placement.* Denver, CO: Education Commission of the States. Retrieved from http://www.ecs.org/clearinghouse/67/44/6744.htm

Dynarski, S. and Scott-Clayton, J. E. (2008). Complexity and targeting in federal student aid: A quantitative analysis (NBER Working Paper No. 13801). Cambridge, MA: National Bureau of Economic Research.

Federman, M. (2007). State graduation requirements, high school course taking, and choosing a technical college major. *The B.E. Journal of Economic Analysis & Policy, 7*(1), 1–32.

Geiser, S., & Santelices, V. (2005). *The role of Advanced Placement and honors courses in college admissions.* Berkeley, CA: Center for Studies in Higher Education, University of California, Berkeley.

Goldrick-Rab, S., Carter, D. F., & Wagner, R. W. (2007). What higher education has to say about the transition to college. *Teachers College Record, 109*(10), 2444–2481.

Heller, D., & Rogers, K. (2006). Shifting the burden: Public and private financing of higher education in the United States and implications for Europe. *Tertiary Education and Management, 12*(2), 91–117.

Horn, L. J., & Kojaku, L. K. (2001). *High school academic curriculum and the persistence path through college: Persistence and transfer behavior of undergraduates 3 years after* *entering 4-year institutions.* Washington, DC: National Center for Education Statistics, Department of Education.

Joensen, J. S., & Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources, 44*(1), 171–198.

Kane, T. J., & Rouse, C. E. (1999). The community college: Educating students at the margin between college and work. *Journal of Economic Perspectives, 13*(1), 63–84.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology, 29*(1), 417–442.

Kinzie, J., Palmer, M., Hayek, J., Hossler, D., Jacob, S. A., & Cummings, H. (2004). *Fifty years of college choice: Social, political, and institutional influences on the decision-making process.* Indianapolis, IN: Lumina Foundation for Education.

Klopfenstein, K., & Thomas, M. K. (2009). The link between Advanced Placement experience and early college success. *Southern Economic Journal, 75*(3), 873–891.

Klugman, J. (2012). How resource inequalities among high school reproduce class advantages in college destinations. *Research in Higher Education, 53*(8), 803–830.

Lee, V. E., Croninger, R. G., & Smith, J. B. (1997). Course-taking, equity, and mathematics learning: Testing the constrained curriculum hypothesis in U.S. secondary schools. *Educational Evaluation and Policy Analysis, 19*(2), 99–121.

Levine, P. B., & Zimmerman, D. J. (1995). The benefit of additional high-school math and science classes for young men and women. *Journal of Business & Economic Statistics, 13*(2), 137–149.

Long, B. T. (2007). The contributions of economics to the study of college access and success. *Teachers College Record, 109*(10), 2367–2439.

Long, B. T., & Boatman, A. (2013). The role of remedial and developmental courses in access and persistence. In A. Jones and L. Perna (Eds.), *The state of college access and completion: Improving college success for students from underrepresented groups*. London: Routledge International Studies in Higher Education.

Long, L. H. (1972). The influence of number and ages of children on residential mobility. *Demography, 9*(3), 371–382.

Loveless, T. (2013). *The algebra imperative: Assessing algebra in a national and international context.* Washington, DC: Brookings. Retrieved September 1, 2013 from http://www.brookings.edu/research/papers/2013/09/04-algebra-imperative-education-loveless

Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *The Journal of Economic Perspectives, 20*(4), 111–132.

OECD (2013). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know.* Retrieved from: http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Perna, L. W. (2004). The key to college access:A college preparatory curriculum. In W. G. Tierney, Z. Corwin & J. Colyar (Eds.), *Preparing for college: Nine elements of effective outreach* (pp. 113–134). Albany: State University of New York Press.

Perna, L. W., & Titus, M. A. (2005). The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences. *The Journal of Higher Education, 76*(5), 485–518.

Porter, S. R. (2012). *Using instrumental variables properly to account for selection effects.* Unpublished manuscript. Retrieved from http://www.stephenporter.org/papers/Pike_IV.pdf

Reys, B. J. (Ed.). (2006). *The intended mathematics curriculum as represented in state-level curriculum standards: Consensus or confusion?* Greenwich, CT: Information Age Publishing, Inc.

Reys, B. J., Dingman, S., Nevels, N., & Teuscher, D. (2007). *High school mathematics: State-level curriculum standards and graduation requirements.* Colombia, MO: Center for the Study of Mathematics Curriculum.

Rose, H., & Betts, J. R. (2001). *Math matters: The links between high school curriculum, college graduation, and earnings.* San Francisco: Public Policy Institute of California.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics, 86*(2), 497–513.

Roth, J., Crans, G., Carter, R., Ariet, M., & Resnick, M. B. (2001). Effect of high school course-taking and grades on passing a college placement test. *The High School Journal, 84*(2), 72–87.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87*(3), 355–374.

St. John, E. P., & Chung, A. S. (2006). Access to advanced math. In E. P. St John (Ed.), *Education and the public interest* (pp. 135–162). Netherlands: Springer.

Stock, J. H., & Trebbi, F. (2003). Who invented Instrumental Variable Regression? *Journal of Economic Perspectives,* 17(3), 177−94.

Stock, J. H., and M. Yogo. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in Honor of Thomas Rothenberg,* (pp. 80–108). Cambridge, UK: Cambridge University Press.

Teitelbaum, P. (2003). The influence of high school graduation requirement policies in mathematics and science on student course-taking patterns and achievement. *Educational Evaluation and Policy Analysis, 25*(1), 31–57.

Terenzini, P. T., Cabrera, A. F., & Bernal, E. M. (2001). *Swimming against the tide: The poor in American higher education.* New York: College Entrance Examination Board.

Wellman, J. V. (2002). *State policy and community-baccalaureate transfer* (National Center Report No. 02–6). San Jose, CA: The National Center for Public Policy and Higher Education, Washington, DC: The Institute for Higher Education Policy.