Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools

Joseph G. Altonji

Yale University and National Bureau of Economic Research

Todd E. Elder

University of Illinois at Urbana-Champaign

Christopher R. Taber

Northwestern University and National Bureau of Economic Research

In this paper we measure the effect of Catholic high school attendance on educational attainment and test scores. Because we do not have a good instrumental variable for Catholic school attendance, we develop new estimation methods based on the idea that the amount of selection on the observed explanatory variables in a model provides a guide to the amount of selection on the unobservables. We also propose an informal way to assess selectivity bias based on measuring the ratio of

We thank Timothy Donohue and Emiko Usui for excellent research assistance. We received helpful comments from Glen Cain, Tim Conley, Thomas DeLeire, Steve Levitt, Lars Hansen, James Heckman, Robert LaLonde, Jean-Marc Robin, George Jakubson, and especially Derek Neal and the anonymous referees, as well as participants in seminars at the American Economic Association winter meetings (January 2000), Boston College, Cornell University, CREST-INSEE, Duke University, Institute for the Study of Labor (Bonn), Johns Hopkins University, Harvard University, Massachusetts Institute of Technology, Northwestern University, Princeton University, University of Chicago, University College London, University of Florida, University of Maryland, University of Missouri, University of Rochester, and University of Wisconsin–Madison. We are grateful for support from the National Science Foundation grant 9512009 (Altonji), the National Institute of Child Health and Development grant R01 HD36480-03 (Altonji and Taber), and the Institute for Policy Research, Northwestern University.

[*Journal of Political Economy,* 2005, vol. 113, no. 1]

selection on unobservables to selection on observables that would be required if one is to attribute the entire effect of Catholic school attendance to selection bias. We use our methods to estimate the effect of attending a Catholic high school on a variety of outcomes. Our main conclusion is that Catholic high schools substantially increase the probability of graduating from high school and, more tentatively, attending college. We find little evidence of an effect on test scores.

I. Introduction

Distinguishing between correlation and causality is the most difficult challenge faced by empirical researchers in the social sciences. In most cases, doubts remain about the validity of the identifying assumptions and the inferences that are based on them. The challenge of isolating causal effects is particularly difficult for the question addressed in our paper: Do Catholic high schools provide better education than public schools? This question is at the center of the debate in the United States over whether vouchers, charter schools, and other reforms that increase choice in education will improve the quality of education. It is also relevant to the search for ways to improve public schools.

Simple cross tabulations or multivariate regressions of outcomes such as test scores and postsecondary educational attainment typically show a substantial positive effect of Catholic school attendance.[1] However, many prominent social scientists, such as Goldberger and Cain (1982), have argued that the positive effects of Catholic school attendance may be due to spurious correlations between Catholic school attendance and unobserved student and family characteristics. In the absence of experimental data, the challenge in addressing this potentially large bias is finding exogenous variation that affects school choice but not outcomes. Most student background characteristics that influence the Catholic school decision, such as income, attitudes, and education of the parents, are likely to influence outcomes independently of the school sector because they are likely to be related to other parental inputs. Characteristics of private and public schools that influence choice, such

[1] Influential early examples include Coleman, Hoffer, and Kilgore (1982) and Coleman and Hoffer (1987). Recent studies include Tyler (1994), Evans and Schwab (1995), Neal (1997), Figlio and Stone (1999), Grogger and Neal (2000), Sander (2000), and Jepsen (2003). Murnane (1984), Chubb and Moe (1990), Witte (1992), Cookson (1993), and Neal (1998) provide overviews of the discussion and references to the literature. Grogger and Neal (2000) provide citations to a small experimental literature, which for the most part has found positive effects of Catholic schools.

as tuition levels, student body characteristics, or school policies, are likely to be related to the effectiveness of the schools.[2]

In this paper we present new estimation strategies that may be helpful when strong prior information is unavailable regarding the exogeneity of either the variable of interest or instruments for that variable. We view this to be the situation in studies of Catholic school effects and in many other applications in economics and the other social sciences. We then use our strategies to assess the effectiveness of Catholic schools. The formal econometric theory is presented in Altonji et al. (2002*b*).

Our approach uses the degree of selection on observables as a guide to the degree of selection on the unobservables. Researchers often informally argue for the exogeneity of an explanatory variable by examining the relationship between the variable and a set of observed characteristics or by assessing whether point estimates are sensitive to the inclusion of additional control variables.[3] In this paper, we show that such evidence can be informative in some situations. More important, we provide a way to quantitatively assess the degree of omitted variables bias.[4]

In Section II, we set the stage for the development and application of our econometric methods by providing a standard multivariate analysis of the Catholic school effect using the National Educational Longitudinal Survey of 1988 (NELS:88). The descriptive statistics show huge Catholic high school advantages in high school graduation and college attendance rates and smaller ones in twelfth grade test scores. However, the evidence across the wide range of observables suggests fairly strong positive selection into Catholic schools. We also find that the link between observables and Catholic high school attendance is much weaker among children who attended Catholic eighth grade and that public school eighth graders almost never attend Catholic high school. These facts suggest that we can improve comparability of the "treatment" and

[2] Several recent studies, including Evans and Schwab (1995), Neal (1997), Figlio and Stone (1999), Grogger and Neal (2000), and Altonji, Elder, and Taber (2002*a*), use various exclusion restrictions to estimate the Catholic school effect on a variety of outcomes, including religious affiliation, geographic proximity to Catholic schools, and the interaction between them. Altonji et al. (2002*a*) raise doubts about the validity of all these instruments. Grogger and Neal (2000) also raise questions about proximity measures, and Ludwig (1997) raises serious doubts about the validity of Catholic religion as an instrument.

[3] See, e.g., Bronars and Grogger (1994), Poterba, Venti, and Wise (1994), Currie and Thomas (1995), Engen, Gale, and Scholz (1996), Udry (1996), Angrist and Evans (1998), Angrist and Krueger (1999), Jacobsen, Pearce, and Rosenbloom (1999), or Cameron and Taber (2004). Wooldridge's (2000) undergraduate textbook contains a computer exercise (15.14) that instructs students to look for a relationship between an observable (IQ) and an instrumental variable (closeness to college).

[4] Two precursors to our study are Altonji's (1988) study of the importance of observed and unobserved family background and school characteristics in the school-specific variance of educational outcomes and Murphy and Topel's (1990) study of the importance of selection on unobserved ability as an explanation for industry wage differentials.

"control" groups by focusing on the Catholic eighth grade sample. By doing so we also avoid confounding the effect of attending Catholic high school with the effect of Catholic elementary school. We find a small positive effect on twelfth grade math scores and a zero effect on reading scores. However, our estimates point to a very large positive effect of 0.15 on the probability of attending a four-year college two years after high school and 0.08 on the probability of graduating from high school. The insensitivity of the results to a powerful set of controls and the "modest" association between the observables that determine the outcome and Catholic high school suggest that part of the educational attainment effect is real.

In Sections III and IV, we present and apply our methods for using the degree of selection on observables to provide better guidance about bias from selection on unobservables. We discuss a condition that formalizes the idea that "selection on the unobservables is the same as selection on the observables." Roughly speaking, this condition states that the part of an outcome (such as high school graduation) that is related to the observables has the *same* relationship with Catholic high school attendance as the part related to the unobservables. It requires some strong assumptions, including the assumptions (1) that the set of observed variables is chosen at random from the full set of variables that determine Catholic school attendance and high school graduation and (2) that the number of observed and unobserved variables is large enough that none of the elements dominates the distribution of school choice or graduation. We argue that these assumptions are no more objectionable than the assumptions needed to justify the standard ordinary least squares (OLS) or univariate probit requirement that the index of unobservables that determine graduation has no relationship with Catholic school attendance. However, we also argue that for the decision to attend Catholic school, selection on the unobservables is likely to be less strong than selection on the observables. Operationally this means that we can obtain a lower-bound estimate of the Catholic school effect by estimating joint models of school choice and the outcome model subject to the restriction that selection on unobservables and observables is equal. The OLS or probit models assume that selection on the unobservables is zero and provide an upper-bound estimate. The estimate of the effect of Catholic school on high school graduation declines from the univariate estimate of about 0.08, which we view as an upper bound, to 0.05 when we impose equal selection, which we view as a lower bound, although sampling error widens this range. The estimate of the effect on college attendance declines from the univariate estimate of 0.15 to 0.03 or 0.02, depending on the details of the estimation method.

We also present a closely related but more informal way to use the relationship between the observables and Catholic high school attendance as a guide to endogeneity bias. We measure the amount of selection on the index of observables in the outcome equation and then calculate a ratio of how large selection on unobservables would need to be in order to attribute the entire effect of Catholic school attendance to selection bias. We find that selection on unobservables would need to be 3.55 times stronger than selection on observables in the case of high school graduation, which seems highly unlikely. It would have to be 1.43 times stronger to explain the entire college effect, which is also unlikely. However, more modest positive selection on the unobservables could explain away the entire Catholic school effect on math scores.

Our main conclusion is that Catholic high school attendance substantially boosts high school graduation rates and, more tentatively, college attendance rates. In Section V we obtain larger univariate effects for urban minorities but also stronger evidence of selection. We conclude the paper in Section VI.

II. Preliminary Analysis of the Catholic School Effect

A. Data

Our data set is NELS:88, a National Center for Education Statistics (NCES) survey that began in the spring of 1988. A total of 1,032 schools contributed up to 26 eighth graders to the base year survey, resulting in 24,599 participants. The NCES attempted to contact subsamples consisting of 20,062 base year respondents in the 1990 and 1992 follow-ups and 14,041 in the 1994 survey. Additional observations are lost as a result of attrition.

The NELS:88 contains information on a wide variety of outcomes, including test scores and other measures of achievement, high school dropout and graduation status, and postsecondary education (in the 1994 survey only). Parent, student, and teacher surveys in the base year provide a rich set of information on family and individual background, as well as pre–high school measures of achievement, behavior, and expectations of success in high school and beyond. Each student was also administered a series of cognitive tests in the 1988, 1990, and 1992 follow-ups. We use the eighth grade test scores as control variables and twelfth grade reading and math tests as outcome measures.

The high school graduation variable HS is equal to one if the respondent graduated from high school by the date of the 1994 survey, and zero otherwise. The college attendance indicator COLL is one if the respondent was enrolled in a four-year college at the time of the

1994 survey and zero otherwise.[5] The indicator variable for Catholic high school attendance, CH, is one if the current or last school in which the respondent was enrolled was Catholic as of 1990 (two years after the eighth grade year) and zero otherwise.[6]

We estimate models using a full sample, a Catholic eighth grade sample (hereafter, the C8 sample), and various other subsamples. We always exclude approximately 400 respondents who attended non-Catholic private high schools or eighth grades. Observations with missing values of key eighth grade or geographic control variables (such as distance from the nearest Catholic high school) are dropped. Sample sizes vary across dependent variables because of data availability and are presented in the tables. The sampling probabilities for the NELS:88 follow-ups depend on choice of private high school and the dropout decision, so sample weights are used to avoid bias from a choice-based sample. Unless noted otherwise, the results reported in the paper are weighted.[7] Details regarding construction of variables and the composition of the sample are provided in Altonji et al. (2002*b*, apps. A, B).

B. *Characteristics of Catholic and Public High School Students*

In table 1 we report the weighted means by high school sector of a set of family background characteristics, student characteristics, eighth grade outcomes, and high school outcomes. We report results separately for students for the C8 sample and for the full sample.[8] Catholic high school students are far less likely to drop out of high school than their public school counterparts (0.02 vs. 0.15) and are almost twice as likely to be enrolled in a four-year college in 1994 (0.59 vs. 0.31). Differences in twelfth grade test scores are more modest but still substantial—about 0.4 of a standard deviation higher for Catholic high school students. In the C8 sample the gap in the dropout rate is also very large (0.02 vs. 0.12), as is the gap in the college attendance rate (0.61 vs. 0.38). In contrast, the gap in the twelfth grade math score is only about 0.25 standard deviations. Table 2 shows that the gaps in school attainment and test scores are even more dramatic for minority students in urban schools.

[5] Our major findings are robust to whether or not college attendance is limited to four-year universities, full-time vs. part-time attendance, or enrollment in college "at some time since high school" or at the survey date.

[6] Bias from the fact that students who started in a Catholic high school and transferred to a public school prior to the tenth grade survey would be coded as attending a public high school (CH = 0) is likely to be very small; see Altonji et al. (2002*b*, 10n11) for details.

[7] See Altonji et al. (2002*b*, 10n12) for details on the sampling scheme.

[8] In tables 1 and 2 the outcome variables are weighted with the same weights used in the regression analysis. All other variables are weighted using NELS:88 second follow-up panel weights.

Tables 1 and 2 also show that the means of favorable family background measures, eighth grade test scores and grades, and positive behavior measures in eighth grade are substantially higher for the students who attend Catholic high schools. The large discrepancies raise the possibility that part or even all of the gap in outcomes may be a reflection of who attends Catholic high school. However, the gap is much lower for most variables in the C8 sample. For example, the gap in log family income is 0.49 for the full sample but only 0.19 for the C8 sample. The high school sector gap in parents' educational expectations for the child is substantially larger in the full sample than in the C8 sample, and the difference in the student's expected years of schooling is 0.72 in the full sample but only 0.40 in the C8 sample.[9] The discrepancy in the fraction of students who repeated a grade in grades 4–8 is −0.05 in the full sample and only −0.01 in the C8 sample, and the gap in the fraction of students who are frequently disruptive is −0.05 in the full sample and zero in the C8 sample; both of these variables are powerful predictors of HS. Finally, the gaps in the eighth grade reading and math scores are 3.86 and 3.44, respectively, in the full sample, but only 1.47 and 1.09, respectively, in the C8 sample.

The fact that observable differences by high school sector are smaller for Catholic eighth graders than for public eighth graders is consistent with the presumption that since the parents of eighth graders from Catholic schools have already chosen to avoid public school at the primary level, other more idiosyncratic factors drive selection into Catholic high schools from Catholic eighth grade. Intuitively, it seems likely that these factors could lead to less selection bias than in the full sample, although the overwhelming evidence based on a very broad set of eighth grade observables is that selection bias is positive in both samples.[10] These considerations, concerns about selection bias that arise from the fact that only 0.3 percent of public school eighth graders in our effective sample go to Catholic high school, and the desire to avoid confounding the Catholic high school effect with the effect of Catholic elementary school lead us to focus on the sample of Catholic eighth graders in most of our analysis.[11]

[9] See the notes to table 3 for a complete list of the variables used in our multivariate models. Some are excluded from tables 1 and 2 to keep them manageable. The expectations variables in tables 1 and 2 are excluded from our outcome models because if Catholic school has an effect on outcomes, this may influence expectations.

[10] In unreported results, the pattern of positive selection on the eighth grade variables changes little after conditioning on the set of family background, demographic, and geographic variables in tables 1 and 2.

[11] This percentage is unweighted, with the corresponding weighted percentage being 0.8 percent. The percentage is 0.3 percent in the sample of 16,070 individuals for whom information on sector of eighth grade and sector of tenth grade is available. It is only 0.7 percent among children whose parents are Catholic.

TABLE 1
MEANS OF KEY VARIABLES IN HIGH SCHOOL AND EIGHTH GRADE SECTOR

Variable	Full Sample			Catholic 8th Grade		
	Public 10th (N=11,167)	Catholic 10th (N=672)	Difference	Public 10th (N=366)	Catholic 10th (N=640)	Difference
Demographics:						
Female	.52	.45	−.07**	.61	.50	−.11**
Asian	.03	.04	.01	.05	.05	.00
Hispanic	.09	.09	.00	.08	.09	.01
Black	.10	.09	−.01	.07	.11	.04
White	.78	.78	.00	.80	.74	−.06
Family background:						
Mother's education (years)	13.21	13.96	.75***	13.34	13.88	.54***
Father's education (years)	13.49	14.51	1.01***	13.39	14.38	.99***
Log of family income	10.23	10.72	.49***	10.47	10.66	.19***
Mother only in house	.14	.09	−.05***	.07	.09	.02***
Parent married	.79	.89	.10***	.90	.88	−.02
Parents Catholic	.28	.82	.54***	.84	.84	.00
Geography:						
Rural	.36	.03	−.33***	.13	.01	−.12**
Suburban	.45	.51	.06*	.40	.48	.08
Urban	.19	.46	.27***	.47	.51	.04
Distance to closest Catholic high school (miles)	22.16	2.97	−19.19***	6.91	2.37	−4.53***

Expectations:[a]						
Schooling expectations (years)	15.25	15.97	.72***	15.52	15.92	.40***
Very sure to graduate from high school	.84	.89	.05***	.84	.90	.06*
Parents expect at least some college	.89	.98	.09***	.94	.98	.04
Parents expect at least college graduation	.79	.92	.13***	.88	.91	.03
Student expects white-collar job	.47	.61	.14***	.55	.59	.04
8th grade variables:						
Delinquency index (range 0–4)	.64	.53	−.11*	.54	.46	−.08
Student got into fight	.24	.23	−.02	.20	.19	−.01
Student rarely completes homework	.19	.08	−.11***	.08	.06	−.01
Student frequently disruptive	.12	.08	−.05***	.08	.08	.00
Student repeated grade 4–8	.06	.02	−.05***	.03	.02	−.01
Risk index (range 0–4)	.69	.35	−.34***	.39	.39	.00
Grades composite	2.94	3.16	.22***	3.09	3.20	.11**
Unpreparedness index (0–25)	10.77	11.08	.31***	10.84	11.02	.17
8th grade reading score	51.19	55.05	3.86***	54.12	55.59	1.47
8th grade mathematics score	51.13	54.57	3.44***	52.89	53.98	1.09
Outcomes:						
12th grade reading standardized score	51.20	54.60	3.40***	53.25	54.70	1.45
12th grade math standardized score	51.20	55.54	4.34***	53.13	55.63	2.49***
Enrolled in four-year college in 1994	.31	.59	.28***	.38	.61	.23***
High school graduate	.85	.98	.13***	.88	.98	.10***

* Difference is statistically significant at the .1 level.
** Difference is statistically significant at the .05 level.
*** Difference is statistically significant at the .01 level.
[a] The expectations variables are not included in our empirical models.

TABLE 2
COMPARISON OF MEANS OF KEY VARIABLES BY SECTOR: NELS:88 URBAN MINORITY SUBSAMPLE

	Full Sample			Catholic 8th Grade		
Variable	Public 10th (N=700)	Catholic 10th (N=56)	Difference	Public 10th (N=15)	Catholic 10th (N=54)	Difference
Demographics:						
Female	.57	.57	.00	.60	.61	.01
Asian	.00	.00	.00	.00	.00	.00
Hispanic	.44	.49	.05	.34	.45	.11
Black	.56	.51	−.05	.66	.55	−.11
White	.00	.00	.00	.00	.00	.00
Family background:						
Mother's education (years)	12.61	13.27	.66	13.58	13.21	−.37
Father's education (years)	12.64	14.33	1.69***	12.66	14.36	1.70***
Log of family income	9.62	10.45	.83***	10.16	10.38	.22
Mother only in house	.29	.27	−.02	.29	.23	−.06
Parent married	.57	.74	.18*	.71	.79	.08
Parents Catholic	.39	.58	.19*	.39	.55	.16
Geography:						
Rural	.00	.00	.00	.00	.00	.00
Suburban	.00	.00	.00	.00	.00	.00
Urban	1.00	1.00	.00	1.00	1.00	.00
Distance to closest Catholic high school (miles)	6.04	1.90	−4.14***	1.90	2.01	.11

Expectations:[a]						
Schooling expectations (years)	15.27	16.10	.83***	16.48	16.05	−.43
Very sure to graduate from high school	.80	.94	.14	.88	.94	.06
Parents expect at least some college	.90	.99	.09	.95	.99	.04
Parents expect at least college graduation	.78	.86	.08	.84	.85	.01
Student expects white-collar job	.53	.72	.19**	.50	.70	.20
8th grade variables:						
Delinquency index (range 0–4)	.88	.63	−.25	1.22	.65	−.57
Student got into fight	.34	.19	−.15*	.05	.19	.15*
Student rarely completes homework	.25	.13	−.12**	.23	.14	−.09
Student frequently disruptive	.19	.17	−.02	.14	.17	.03
Student repeated grade 4–8	.11	.05	−.06*	.10	.05	−.05
Risk index (range 0-4)	1.30	.90	−.40***	1.05	.91	−.14
Grades composite	2.78	2.88	.09	3.01	2.88	−.13
Unpreparedness index (0–25)	10.99	11.28	.29	11.10	11.27	.17
8th grade reading score	46.76	53.25	6.49***	49.99	52.88	2.89
8th grade mathematics score	45.43	48.71	3.28***	48.88	48.61	−.27
Outcomes:						
12th grade reading standardized score	47.29	50.78	3.49**	52.74	50.17	−2.57
12th grade math standardized score	46.40	51.71	5.31***	51.46	50.92	−.54
Enrolled in four-year college in 1994	.23	.52	.28***	.28	.56	.28*
High school graduate	.78	.99	.21***	.89	1.00	.11

* Difference is statistically significant at the .1 level.
** Difference is statistically significant at the .05 level.
*** Difference is statistically significant at the .01 level.
[a] The expectations variables are not included in our empirical models.

C. *Probit and OLS Estimates of the Effect of Catholic High Schools*

Panel A of table 3 reports the coefficient on CH in univariate probit models for HS. The difference in means for the C8 sample is 0.105 when no controls are included, as shown by the marginal effect in the probit with no controls (col. 5). When we add the first set of controls, the average marginal effect falls to 0.084, which is suggestive that the family background and geographic controls explain only a fairly small amount of the raw difference in the graduation rate. This effect is still very large considering that the graduation rate is 0.947 among students from the C8 sample. The point estimate of the marginal effect of CH declines slightly to 0.081 when we add eighth grade test scores in column 7 and increases to 0.088 when we add a large set of eighth grade measures of attendance, attitudes toward school, academic track in eighth grade, achievement, and behavioral problems. The stability of the CH effect is remarkable, especially given the fact that the control variables in column 8 are powerful. The pseudo R^2 of the regression model rises from 0.11 to 0.35 when we add the first set of controls and to 0.58 when we add the full set of controls. These covariates are powerful predictors of dropout behavior but lead to only a small change in the estimated CH effect.[12]

In panel B of table 3 we also report estimates of the effect of CH on the probability that a student is enrolled in a four-year college at the time of the third follow-up survey in 1994, two years after most students graduate from high school. The raw difference of 0.236 (col. 5) declines to 0.154 when basic family background and geographic controls are included in the probit model and to 0.149 when we add detailed controls. Once again the pseudo R^2 rises substantially as we add more control variables.

In panels C and D of table 3 we report estimates of the effect of CH on twelfth grade reading and math scores. For the C8 sample with the full set of controls, we obtain small positive effects of 1.14 (0.46) on the math score and 0.33 (0.62) on the reading score. As Grogger and Neal (2000) emphasize, a positive effect of CH on the high school graduation rate might lead to a downward bias in the CH coefficient in the twelfth grade test equations given that dropouts have lower test scores and that

[12] We obtain similar results when we estimate linear probability models and linear probability models with fixed effects for each eighth grade (see Altonji et al. 2002*b*). These results show that factors that vary across Catholic elementary schools (such as public high school quality) do not drive the large positive estimates of the Catholic high school effect. Bias from individual heterogeneity could well be more severe in the within-school analysis.

dropouts have a lower probability of taking the twelfth grade test. However, the issue appears to be of only minor importance.[13]

To facilitate comparison with other studies, we also present estimates for the combined sample of students from Catholic and public eighth grades. For this sample the effect of CH on HS is reduced from 0.123 to 0.052 after we add the full set of controls (cols. 1–4 of table 3). The college attendance results largely mirror the HS results. The probit estimate of the effect of CH is 0.074 once the full set of controls are included, which is substantial relative to the mean college attendance probability of 0.28.

Note that the choice of controls makes a much larger difference in the full sample than in the C8 sample. We do not fully understand this pattern. However, conditioning on eighth grade variables is problematic in the full sample because a substantial number of variables are supplied by schools and teachers and may reflect school-specific standards. For example, the standards for being a "troublemaker" may differ substantially between Catholic and public eighth grades. As a result, in order to draw inferences for the full sample, ideally one would want to control for type of eighth grade and interact the covariates with this variable. However, since virtually all the Catholic high school students come from Catholic eighth grades, this essentially amounts to using the C8 sample to identify the CH effect. It is thus hard to justify why one should be interested in the full sample when the C8 sample is available.

Once detailed controls for eighth grade outcomes are included, the estimates of the effects on twelfth grade reading and math are only 1.14 and 0.92, respectively, which point to a small but statistically significant positive effect. Given the high degree of selection into Catholic high school in the full sample on the basis of observable traits, these estimates may reflect unobserved differences between public and Catholic high school students rather than actual effects on test scores and should be interpreted with caution.

D. A Sensitivity Analysis

Although the evidence suggests only a small amount of selection on observables in the C8 subsample, it is possible that a small amount of selection on unobservables could explain the whole CH effect. We now explore this possibility by examining the sensitivity of the estimates to

[13] We address the issue by imputing missing data for both high school graduates and dropouts using predicted values from a regression of the twelfth grade score on the full set of controls in the outcome regression, plus the Catholic high school dummy and the tenth grade test scores and a dummy variable for whether the individual graduated from high school. High school graduation has a small and statistically insignificant coefficient. See Altonji et al. (2002*b*) for details.

TABLE 3
OLS and Probit Estimates of Catholic High School Effects in Subsamples of NELS:88 (Weighted)

	Full Sample: Controls				Catholic 8th Grade Attendees: Controls			
	None (1)	Family Background, City Size, and Region[a] (2)	Col. 2 Plus 8th Grade Tests (3)	Col. 3 Plus Other 8th Grade Measures[b] (4)	None (5)	Family Background, City Size, and Region[a] (6)	Col. 2 Plus 8th Grade Tests (7)	Col. 3 Plus Other 8th Grade Measures[b] (8)
	A. High School Graduation							
Probit	.97	.57	.48	.41	.99	.88	.95	1.27
	(.17)	(.19)	(.22)	(.21)	(.24)	(.25)	(.27)	(.29)
	[.123]	[.081]	[.068]	[.052]	[.105]	[.084]	[.081]	[.088]
Pseudo R^{2}[c]	.01	.16	.21	.34	.11	.35	.44	.58
	B. College in 1994							
Probit	.73	.37	.33	.32	.60	.48	.56	.60
	(.08)	(.09)	(.09)	(.09)	(.13)	(.15)	(.15)	(.15)
	[.283]	[.106]	[.084]	[.074]	[.236]	[.154]	[.154]	[.149]
Pseudo R^2	.02	.19	.29	.34	.04	.18	.29	.36

	C. 12th Grade Reading Score							
OLS	4.28	2.08	1.18	1.14	1.92	.17	.37	.33
	(.47)	(.54)	(.38)	(.38)	(.82)	(.98)	(.63)	(.62)
R^2	.01	.19	.60	.60	.01	.19	.59	.62
	D. 12th Grade Math Score							
OLS	4.86	1.98	1.07	.92	2.79	1.10	1.46	1.14
	(.44)	(.54)	(.34)	(.32)	(.77)	(1.00)	(.53)	(.46)
R^2	.01	.26	.72	.74	.02	.26	.73	.77

Note.—NELS:88 third follow-up and second follow-up panel weights are used for the educational attainment and twelfth grade models, respectively. Sample sizes for full sample: $N=8{,}560$ (high school graduation), $N=8{,}315$ (college attendance), $N=8{,}116$ (twelfth grade reading), and $N=8{,}119$ (twelfth grade math). For Catholic eighth grade sample: $N=859$ (high school graduation), $N=834$ (college attendance), $N=739$ (twelfth grade reading), and $N=739$ (twelfth grade math). Huber-White standard errors are in parentheses. Marginal effects are in brackets. Marginal effects of probit models are computed as average derivatives of the probability of an outcome with respect to Catholic high school attendance.

[a] Control sets 2–4 include race (white/nonwhite), Hispanic origin, gender, urbanicity (three categories), region (eight categories), and distance to the nearest Catholic high school (five categories). Family background variables used as controls include log family income, mother's and father's education, five dummy variables for marital status of the parents, and eight dummy variables for household composition.

[b] Other eighth grade measures include measures of attendance, attitudes toward school, academic track, achievement, and behavioral problems (from teacher, parent, and student surveys). The NELS:88 variables used are bys55a, bys55e, bys55f, byt1_2, bys56e, byp50, byp57e, bylep, bys55b, bys55d, byrisk, bygrads, byp51, and bys78a–c and also teacher survey variables regarding whether a student performs below ability, completes homework, is attentive or disruptive in class, or is frequently absent or tardy. See app. A of Altonji et al. (2002*b*) for more details.

[c] The pseudo R^2 for probit models is defined as $\mathrm{Var}(\boldsymbol{X}'\boldsymbol{\gamma})/[1+\mathrm{Var}(\boldsymbol{X}'\boldsymbol{\gamma})]$.

the correlation between the unobserved factors that determine CH and the various outcomes Y. We display estimates of the CH effect for a range of values of the correlation between the unobserved determinants of school choice and the outcome.

Consider the bivariate probit model

$$\mathrm{CH} = 1(\boldsymbol{X}'\boldsymbol{\beta} + u > 0), \tag{1}$$

$$Y = 1(\boldsymbol{X}'\boldsymbol{\gamma} + \alpha\mathrm{CH} + \epsilon > 0), \tag{2}$$

and

$$\begin{bmatrix} u \\ \epsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \tag{3}$$

While the above model is formally identified without an exclusion restriction, semiparametric identification requires such an excluded variable. Furthermore, empirical researchers are highly skeptical of results from this model in the absence of an exclusion restriction. Accordingly, our thought exercise in this subsection is to treat (1)–(3) as though the model is underidentified by one parameter. In particular, we act as though ρ is not identified.

In table 4 we display estimates of CH effects that correspond to various assumptions about ρ, the correlation between the error components in the equations for CH and Y.[14] We report results for HS in panel A and college attendance in panel B, and we include both probit coefficients and average marginal effects on the outcome probabilities (in brackets). We include family background, eighth grade tests, and other eighth grade measures as controls.[15] We vary ρ from zero (the univariate probit case that we have already considered above) to 0.5 by estimating bivariate probit models constraining ρ to the specified value. For the full sample, the raw difference in the high school graduation probability is 0.13. When $\rho = 0$, the estimated effect is 0.058, and the figure declines to 0.037 when $\rho = 0.1$ and to 0.011 when $\rho = 0.2$. The latter value is not statistically significant. Given the strong relationship between the observables that determine HS and CH in the full sample, the evidence

[14] See Rosenbaum and Rubin (1983) or Rosenbaum (1995) for examples of this type of sensitivity analysis. In these approaches, the authors essentially restrict the correlation between the error terms in a selection equation and an outcome equation and ask what values are plausible in a particular model.

[15] Because of convergence problems in estimating the bivariate probit models, we eliminated from the set of controls the dummy variables for household composition (but not marital status of parents), urbanicity, region, and indicators for "student rarely completes homework," "student performs below ability," "student inattentive in class," "a limited English proficiency index," and "parents contacted about behavior."

TABLE 4
SENSITIVITY ANALYSIS: ESTIMATES OF CATHOLIC HIGH SCHOOL EFFECTS GIVEN DIFFERENT ASSUMPTIONS ON THE CORRELATION OF DISTURBANCES IN BIVARIATE PROBIT MODELS IN SUBSAMPLES OF NELS:88: MODIFIED CONTROL SET

	Correlation of Disturbances[a]					
	$\rho = 0$	$\rho = .1$	$\rho = .2$	$\rho = .3$	$\rho = .4$	$\rho = .5$
	A. High School Graduation					
Full sample (raw difference = .12)	.459	.271	.074	−.132	−.349	−.581
	(.150)	(.150)	(.150)	(.148)	(.145)	(.140)
	[.058]	[.037]	[.011]	[−.021]	[−.060]	[−.109]
Catholic 8th graders (raw difference = .08)	1.036	.869	.697	.520	.335	.142
	(.314)	(.313)	(.310)	(.306)	(.299)	(.290)
	[.078]	[.064]	[.050]	[.038]	[.025]	[.011]
Urban minorities (raw difference = .22)	1.095	.905	.706	.499	.282	.053
	(.526)	(.538)	(.549)	(.560)	(.570)	(.578)
	[.176]	[.157]	[.132]	[.101]	[.062]	[.013]
	B. College Attendance					
Full sample (raw difference = .31)	.331	.157	−.019	−.196	−.376	−.558
	(.070)	(.070)	(.070)	(.068)	(.067)	(.064)
	[.084]	[.039]	[−.005]	[−.047]	[−.087]	[−.125]
Catholic 8th graders (raw difference = .23)	.505	.336	.165	−.008	−.184	−.362
	(.121)	(.120)	(.119)	(.117)	(.114)	(.110)
	[.140]	[.093]	[.045]	[−.002]	[−.050]	[−.099]
Urban minorities (raw difference = .30)	.447	.269	.090	−.091	−.272	−.455
	(.282)	(.282)	(.280)	(.276)	(.269)	(.259)
	[.116]	[.062]	[.020]	[−.020]	[−.057]	[−.091]

NOTE.—NELS:88 third follow-up sampling weights are used in the computations. Owing to computational difficulties, several variables were excluded from the control sets in the bivariate probit models: all dummy variables for household composition; urbanicity and region; indicators for "student rarely completes homework," "student performs below ability," "student inattentive in class," and "parents contacted about behavior"; and a limited English proficiency index. Other than these exclusions, the controls are identical to those described in notes a and b of table 3. Huber-White standard errors are in parentheses. Marginal effects are in brackets.

[a] Models estimated as bivariate probits with the correlation ρ between u and ϵ set to the values in column headings.

for a strong CH effect is considerably weaker than suggested by the estimates that take CH as exogenous.

For our preferred sample of Catholic eighth graders, the results are less sensitive to ρ. Presumably this arises because the pseudo R^2 is higher, which would generally imply that the same correlation in unobservables would lead to less selection bias. In this circumstance, the type of sensitivity analysis that we conduct in the table is more informative. The effect on HS is 0.078 when $\rho = 0$, which is slightly below the estimate we obtain with the full set of controls in table 3. It declines to 0.038 and is significant at the 10 percent level when $\rho = 0.3$ and is still positive when $\rho = 0.5$, although it is not significant. Thus, when sampling error is ignored, the correlation between the unobservable determinants of CH and HS would have to be greater than 0.5 to explain the estimated effect under the null of no "true" CH effect.

In panel B of table 4 we present the results for college attendance.

For the full sample, the results are very similar to the HS results. The evidence for a positive effect of CH on college attendance is stronger in the C8 sample than in the full sample. The point estimate is 0.045 (though not significant at conventional levels) when $\rho = 0.02$ and remains positive until ρ is about 0.3. However, in this sample the strongest evidence is for a positive effect of CH on HS.

To see whether these results are driven by the joint normality assumption, we repeated the analysis after generalizing this assumption using the semiparametric specification

$$u = \theta + u^* \tag{4}$$

and

$$\epsilon = \theta + \epsilon^*, \tag{5}$$

where the distribution of θ is unrestricted and u^* and ϵ^* are independent standard normals. As long as the correlation between u and ϵ is nonnegative, the bivariate probit is a special case of this model in which θ is restricted to be normally distributed. We estimate the model using nonparametric maximum likelihood (see, e.g., Heckman and Singer 1984), which involves treating the distribution of θ as discrete; in practice we obtain three points of support for θ. The (unreported) results are similar to those in table 4. For example, in the full sample the effect of CH on high school graduation is 0.034 when $\rho = 0.2$ and -0.055 when $\rho = 0.5$. In the C8 sample the estimates are 0.058 when $\rho = 0.2$ and 0.025 when $\rho = 0$. The estimated effect of CH on college attendance for the C8 sample is 0.055 when $\rho = 0.2$ and -0.049 when $\rho = 0.5$.

To this point, our preferred estimates based on the C8 sample suggest a strong positive effect of CH on HS and COLL. For this subsample, the relationship between Catholic high school attendance and other observables seems weak, and the estimates are not very sensitive to the addition of a powerful set of control variables, especially in the case of high school graduation. Finally, the sensitivity analysis shows that in the C8 sample the degree of selection on unobservables must be quite high to explain the full CH effect. This is where the typical analysis of bias due to selection on unobservables based on patterns in the observables would end, with the conclusion that part of the CH effect on educational attainment is probably real. The problem with this type of analysis is that, without prior knowledge, it is hard to judge the magnitude of ρ. To solve this problem, in the next sections we use the degree of selection on the observables as a guide.

III. Selection Bias and the Link between the Observed and Unobserved Determinants of School Choice and Outcomes

We now discuss a theoretical foundation for the practice of using the relationship between an endogenous variable and the observables to make inferences about the relationship between the variable and the unobservables. With our Catholic school application, let the outcome of interest Y be a function of a latent variable Y^*, which is determined as

$$\begin{aligned} Y^* &= \alpha\mathrm{CH} + \boldsymbol{W}'\boldsymbol{\Gamma} \\ &= \alpha\mathrm{CH} + \boldsymbol{X}'\boldsymbol{\Gamma}_X + \xi, \end{aligned} \tag{6}$$

where CH is an indicator for whether the student attends a Catholic high school, the parameter α is the causal effect of CH on Y^*, $\boldsymbol{W}$ is the full set of variables (observed and unobserved) that determine Y^*, and $\boldsymbol{\Gamma}$ is the causal effect of $\boldsymbol{W}$ on Y^*. In the second line of (6), $\boldsymbol{X}$ is a vector of the observable components of $\boldsymbol{W}$, $\boldsymbol{\Gamma}_X$ is the corresponding subvector of $\boldsymbol{\Gamma}$, and the error component ξ is an index of the unobserved variables. Because it is unlikely that the control variables $\boldsymbol{X}$ are all unrelated to ξ, we work with

$$Y^* = \alpha\mathrm{CH} + \boldsymbol{X}'\boldsymbol{\gamma} + \epsilon, \tag{7}$$

where $\boldsymbol{\gamma}$ and ϵ are *defined* so that $\mathrm{Cov}(\epsilon, \boldsymbol{X}) = 0$. Consequently, $\boldsymbol{\gamma}$ captures both the direct effect of $\boldsymbol{X}$ on Y^*, $\boldsymbol{\Gamma}_X$, and the relationship between $\boldsymbol{X}$ and ξ. Let CH* be the latent variable that determines CH such that $\mathrm{CH} = 1(\mathrm{CH}^* > 0)$, where the indicator function $1(\cdot)$ is one when $\mathrm{CH}^* > 0$ and zero otherwise. Consider the linear projection of CH* onto $\boldsymbol{X}'\boldsymbol{\gamma}$ and ϵ,

$$\mathrm{Proj}(\mathrm{CH}^*|\boldsymbol{X}'\boldsymbol{\gamma}, \epsilon) = \phi_0 + \phi_{X'\gamma}\boldsymbol{X}'\boldsymbol{\gamma} + \phi_\epsilon\epsilon. \tag{8}$$

We formalize the idea that "selection on the unobservables is the same as selection on the observables" in the following condition.

CONDITION 1.

$$\phi_\epsilon = \phi_{X'\gamma}.$$

We contrast this with the OLS condition.

CONDITION 2.

$$\phi_\epsilon = 0.$$

Roughly speaking, condition 1 says that the part of Y^* that is related to the observables and the part related to the unobservables have the *same* relationship with CH*. Condition 2 says that the part of Y related to the unobservables has *no* relationship with CH*.

The precise conditions and formal model leading to condition 1 are given in Altonji et al. (2002*b*). The following three types of assumptions suffice:

1. the elements of $\boldsymbol{X}$ are chosen at random from the full set of factors $\boldsymbol{W}$ that determine Y;
2. the numbers of elements in $\boldsymbol{X}$ and $\boldsymbol{W}$ are large, and none of the elements dominates the distribution of CH or the outcome Y; and
3. the relationship between the observable elements $\boldsymbol{X}$ and the unobservables obeys an assumption that is very strong but weaker than the standard assumption $\text{Cov}(\boldsymbol{X}, \xi) = 0$. Roughly speaking, the assumption is that the regression of CH* on $Y^* - \alpha\text{CH}$ is equal to the regression of the part of CH* that is orthogonal to $\boldsymbol{X}$ on the corresponding part of $Y^* - \alpha\text{CH}$.[16]

While the assumptions that lead to condition 1 are strong and unlikely to hold exactly, they are no more objectionable than the OLS assumptions leading to condition 2: $\text{Cov}(\text{CH}, \xi) = 0$ and $\text{Cov}(\boldsymbol{X}, \xi) = 0$.[17] Assumptions of types 1 and 2, in particular, are likely to be better approximations to reality than the OLS assumptions because of the manner in which most large-scale data sets are designed and collected. Data sets such as the National Longitudinal Study of Youth, NELS:88, the Panel Study of Income Dynamics, and the German Socioeconomic Panel are designed to serve multiple purposes rather than to address one relatively specific question, such as the effectiveness of Catholic schools. Content of a data set is a compromise among the interests of multiple research, policy making, and funding constituencies. Burden on the respondents, budget, and access to administrative data sources serve as constraints. Obviously, content is also shaped by what is known about

[16] Let Γ_j be the coefficient on W_j in (6) and treat CH* symmetrically with $Y^* - \alpha\text{CH}$ so that

$$\text{CH}^* = \sum_{j=1}^{K} W_j\beta_j + \eta_K,$$

where η is defined to be orthogonal to the W_j, K is the number of elements of $\boldsymbol{W}$, and the sequence $\{W_j\beta_j\}$ is stationary. The condition is a limiting result as the number of elements of $\boldsymbol{W}$ goes to infinity. It is

$$\frac{\sum_{l=-\infty}^{\infty} E(W_jW_{j-l})E(\beta_j\Gamma_{j-l})}{\sum_{l=-\infty}^{\infty} E(W_jW_{j-l})E(\Gamma_j\Gamma_{j-l})} = \frac{\sum_{l=-\infty}^{\infty} E(\tilde{W}_j\tilde{W}_{j-l})E(\beta_j\Gamma_{j-l})}{\sum_{l=-\infty}^{\infty} E(\tilde{W}_j\tilde{W}_{j-l})E(\Gamma_j\Gamma_{j-l})},$$

where $\tilde{W}_j$ are the residuals from a regression of W_j onto $\boldsymbol{X}$. It is easy to show that this condition holds under the standard assumption $E(\xi|\boldsymbol{X}) = 0$. However, $E(\xi|\boldsymbol{X}) = 0$ is not likely to hold and fortunately is not necessary for the above equation. Altonji et al. (2002*b*) present an example of a model for which the equation holds but $E(\xi|\boldsymbol{X}) \neq 0$.

[17] Technically, OLS can be unbiased if the conditions $\text{Cov}(\text{CH}, \xi) = 0$ and $\text{Cov}(\boldsymbol{X}, \xi) = 0$ happen to fail in a way that leads to a perfect cancelation of biases, or if $\text{Cov}(\text{CH}, \xi) = 0$, $\text{Cov}(\boldsymbol{X}, \xi) \neq 0$, but $\text{Cov}(\text{CH}, \boldsymbol{X}')\,\text{Var}(\boldsymbol{X})^{-1}\,\text{Cov}(\boldsymbol{X}, \xi) = 0$. Neither of these cases is very interesting.

the factors that really matter for particular outcomes and by variation in the feasibility of collecting useful information on particular topics. Explanatory variables that influence a large set of important outcomes (such as family income, race, education, gender, or geographical information) are more likely to be collected. But as a result of the limits on the number of the factors that we know matter and that we know how to collect and can afford to collect, many elements of $\boldsymbol{W}$ are left out. This is reflected in the relatively low explanatory power of social science models of individual behavior. Furthermore, in many applications, including ours, the endogenous variable is correlated with many of the elements of $\boldsymbol{X}$. Given the constraints that shape the choice of $\boldsymbol{X}$ and the fact that many of the elements of $\boldsymbol{X}$ are systematically related to CH*, it is unlikely that the many unobserved variables that determine ξ are unrelated to CH*, which is basically what Cov (CH, ξ) $= 0$ requires. Since the $\boldsymbol{X}$ variables are intercorrelated, the assumption that Cov ($\boldsymbol{X}$, ξ) $= 0$ is also likely to be a poor approximation to reality even though it is made in virtually all empirical studies in the social sciences.

These considerations suggest that condition 2, which underlies single-equation methods in econometrics, will rarely hold in practice. Many factors that influence Y^* and are correlated with CH* or $\boldsymbol{X}$ or both are left out. The assumptions leading to condition 1 represent the other extreme, which is that the constraints on data collection are sufficiently severe that it is better to think of the elements of $\boldsymbol{X}$ as a more or less random subset of the elements of $\boldsymbol{W}$ rather than a set that has been systematically chosen to eliminate bias.

In our case, we have data on a broad set of family background measures, teacher evaluations, test scores, grades, and behavioral outcomes in eighth grade, as well as measures of proximity to a Catholic high school. These measures cover most of the socioeconomic, academic, and behavioral factors stressed in the literature on educational attainment. They have substantial explanatory power for the outcomes that we examine, and a large number of the variables play a role, particularly in the case of high school graduation and college attendance. Once we restrict the sample to Catholic eighth graders and condition on Catholic religion and distance from a Catholic high school, a broad set of variables make minor contributions to the probability of Catholic high school attendance. The relatively large number and wide variety of observables that enter into our problem suggest that the observables may provide a useful guide to the unobservables.

However, the "random selection of observables" assumption that leads to condition 1 is not to be taken literally. In fact, there are strong reasons to expect the relationship between the unobservables and CH (or, more generally, any potentially endogenous treatment) to be weaker than the relationship between the observables and CH. First, $\boldsymbol{X}$ often has been selected with an eye toward reducing bias in single-equation estimates rather than at random. For example, we control for race and ethnicity, which are strongly related to both CH and educational attainment. We also include detailed eighth grade achievement and behavior measures as well as parental background measures that figure prominently in discussions of selection bias. Second, in the case of the twelfth grade test scores, ϵ will also reflect the substantial variability in test performance on a particular day, which presumably has nothing to do with the decision to start Catholic high school. Finally, and most importantly, shocks that occur after eighth grade are excluded from $\boldsymbol{X}$. They will influence high school outcomes but not the probability of starting a Catholic high school. To see this, rewrite ϵ as $\epsilon = \epsilon_1 + \epsilon_2$, where ϵ_1 includes factors determined prior to high school and ϵ_2 is the independent innovation in the error term that is determined during high school. Since CH* is determined in eighth grade, we can impose our data generation condition on the variables determined prior to high school, in which case

$$\phi_{X'\gamma} \equiv \frac{\text{Cov}\,(\text{CH}^*,\ \boldsymbol{X}'\boldsymbol{\gamma})}{\text{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})} = \frac{\text{Cov}\,(\text{CH}^*,\ \epsilon_1)}{\text{Var}\,(\epsilon_1)}. \tag{9}$$

Assume without loss of generality that $\text{Cov}\,(\text{CH}^*, \boldsymbol{X}'\boldsymbol{\gamma}) \geq 0$, as is true in our data. Since $\text{Var}\,(\epsilon) > \text{Var}\,(\epsilon_1)$ and $\text{Cov}\,(\text{CH}^*, \epsilon) = \text{Cov}\,(\text{CH}^*, \epsilon_1)$, then

$$\phi_\epsilon \equiv \frac{\text{Cov}\,(\text{CH}^*,\ \epsilon)}{\text{Var}\,(\epsilon)} \leq \frac{\text{Cov}\,(\text{CH}^*,\ \epsilon_1)}{\text{Var}\,(\epsilon_1)} = \phi_{X'\gamma}.$$

Since $\text{Cov}\,(\text{CH}^*, \epsilon_1) \geq 0$ and $\phi_\epsilon \geq 0$, condition 1 is replaced by condition 3.

Condition 3.

$$0 \leq \phi_\epsilon \leq \phi_{X'\gamma}.$$

In Altonji et al. (2002*b*), we prove that we can identify the set of values of α that satisfy condition 3. In theory this set can be quite complicated, but in practice in our empirical work we find that the identified set is an interval with an upper bound on α that occurs when one assumes that $\text{Cov}\,(\text{CH}^*, \epsilon)/\,\text{Var}\,(\epsilon) = 0$ and a lower bound that occurs when one assumes that

$$\frac{\text{Cov}\,(\text{CH}^*,\ \epsilon)}{\text{Var}\,(\epsilon)} = \frac{\text{Cov}\,(\text{CH}^*,\ \boldsymbol{X}'\boldsymbol{\gamma})}{\text{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})}.$$

Thus, in the empirical work below, we interpret estimates of α that impose condition 1 as a lower bound for α and single-equation estimates with CH treated as exogenous (which impose condition 2) as an upper bound. This simplifies the analysis substantially. If the lower-bound estimates point to a substantial CH effect, we interpret this as strong evidence in favor of such an effect. *We view analysis based on condition 1 and condition 3 as a complement to the standard analysis based on condition 2, not as a replacement for it.*

IV. Estimates of the CH Effect Using Selection on the Observables to Assess Selection Bias

We now return to the bivariate probit model given by (1), (2), and (3). In the bivariate probit case, condition 3 may be rewritten as[18]

$$0 \leq \rho \leq \frac{\mathrm{Cov}\,(\boldsymbol{X}'\boldsymbol{\beta},\ \boldsymbol{X}'\boldsymbol{\gamma})}{\mathrm{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})}. \tag{10}$$

In panel A of table 5, we present estimates that use the C8 sample directly and maximize the likelihood imposing $\rho = \mathrm{Cov}\,(\boldsymbol{X}'\boldsymbol{\beta}, \boldsymbol{X}'\boldsymbol{\gamma})/\,\mathrm{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})$. The standard errors assume that (10) holds for the particular $\boldsymbol{X}$ variables that we have and ignore variation that would arise because that set is not sufficiently large for such variation to be negligible. For HS, the estimate of ρ is 0.24 (0.13). The estimate of α is 0.59 (0.33), which is significant at the .07 level and implies an effect of 0.05 on the probability of high school graduation. Consequently, even with the extreme assumption of equality of selection on observables and unobservables imposed, the point estimate suggests a large positive effect of attending Catholic high school on HS, although the 95 percent confidence interval for the bound includes zero.

The results for COLL follow a similar pattern, but $\rho = \mathrm{Cov}\,(\boldsymbol{X}'\boldsymbol{\beta}, \boldsymbol{X}'\boldsymbol{\gamma})/\,\mathrm{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})$ leads to a larger reduction in the estimated effect of CH on the probability of attending college. The point estimate of 0.03 is still sizable, although it is not statistically significant.

To improve precision of the estimates of α and as a check on the robustness of the results, we also try an alternative method that uses information contained in the sample of public eighth graders. We partition $\boldsymbol{X}$ and $\boldsymbol{\gamma}$ into the subvectors $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_G\}$ and $\{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_G\}$ consisting of variables and parameters that fall into similar categories. In practice, $G = 6$. We estimate $\boldsymbol{\gamma}$ on the public eighth grade sample on the grounds that very few such students go to Catholic high school, and so selectivity will not influence the estimates of $\boldsymbol{\gamma}$ even though the mean of the error term may be different for this sample. We assume that the values of $\boldsymbol{\gamma}$ are the same for students from Catholic and public eighth grades, up to a proportionality factor for each subvector, which

[18] Keep in mind that in the binary probit the variances of ϵ and u are normalized to one.

TABLE 5
Sensitivity of Estimates of Catholic Schooling Effects on College Attendance and High School Graduation to Assumptions about Selection Bias in NELS:88, Catholic Eighth Grade Subsample: Modified Control Set

Model	Constraint on ρ (1)	High School Graduation Coefficients $\hat{\rho}$ (2)	High School Graduation Coefficients $\hat{\alpha}$ (3)	College Attendance Coefficients $\hat{\rho}$ (4)	College Attendance Coefficients $\hat{\alpha}$ (5)
	A. Estimation Method 1[a]				
1	$\rho = \text{Cov}(X'\beta, X'\gamma)/\text{Var}(X'\gamma)$	.24	.59	.24	.11
		(.13)	(.33)	(.06)	(.16)
			[.05]		[.03]
2	$\rho = 0$	0	1.04	0	.51
			(.31)		(.12)
			[.08]		[.14]
	B. Estimation Method 2[b]				
3	$\rho = \text{Cov}(X'\beta, X'\gamma)/\text{Var}(X'\gamma)$	.09	.94	.27	.06
		(.08)	(.30)	(.05)	(.10)
			[.07]		[.02]
	C. Estimation Method 3[c]				
4	$\rho = \text{Cov}(X'\beta, X'\gamma)/\text{Var}(X'\gamma)$	.25	.80	.25	.15
		(.16)	(.37)	(.09)	(.22)
			[.05]		[.04]

Note.—The estimation is performed on a sample of Catholic eighth grade attendees from NELS:88. $N=859$ for the high school graduation sample, and $N=834$ for the college attendance sample. The NELS:88 third follow-up sampling weights are used in the computations. Owing to computational difficulties, several variables were excluded from the control sets in the bivariate probit models. See the note to table 4. Huber-White standard errors are in parentheses. Marginal effects are in brackets.

[a] In panel A, the model is $\text{CH} = 1(X'\beta + u > 0)$ and $Y = 1(X'\gamma + \alpha\text{CH} + \epsilon > 0)$; β, γ, and α are estimated simultaneously as a constrained bivariate probit model.

[b] In panel B, the model is the same as in panel A. It is two-step, with β obtained from a univariate probit and γ from a univariate probit using the public eighth grade subsample. Next, α is computed from a bivariate probit with β fixed at this initial value and γ fixed up to six proportionality factors. The categories of proportionality factors are demographics/family background, test scores, behavioral problems, school attendance and attitudes toward school, grades and achievement, and distance measures. The coefficients (and standard errors) of the proportionality factors for these categories are .82 (.19), .87 (.22), .92 (.03), 1.07 (.04), .59 (.08), and .90 (6.08), respectively, in the high school graduation case. For college attendance, the coefficients (and standard errors) are .80 (.01), 1.01 (.04), .95 (.15), .43 (.17), 1.44 (.03), and 1.04 (1.59).

[c] In panel C, the model is $\text{CH} = 1(X'\beta + \theta + u > 0)$ and $Y = 1(X'\gamma + \alpha\text{CH} + \theta + \epsilon > 0)$; β, γ, and α are estimated simultaneously as a constrained semiparametric model. Models are estimated as univariate probits conditional on θ, the distribution of which is estimated nonparametrically. In col. 1 of this panel, $\rho = \text{Var}(\theta)/[1 + \text{Var}(\theta)]$.

slightly relaxes the implicit assumption of the full-sample models in table 3 that γ does not depend on the sector of the eighth grade.[19] The results using the second estimation method are reported in panel B of table 5. In the case of HS, $\hat{\rho} = 0.09$, $\hat{\alpha} = 0.94$ with a p-value of .002, and the effect on the probability of graduating from high school is 0.09. However, the college effect is only 0.02.

[19] The restrictions on γ pass a likelihood ratio test with a p-value of .12 in the HS case, but fail with a p-value of .03 in the COLL case, so perhaps these alternative results for COLL should be discounted. Details are in note b of table 5.

As a further robustness check, in panel C of table 5 we replace the joint normality assumption implicit in the bivariate probit with the semiparametric specification presented in equations (4) and (5). The results do not change substantially, with the lower-bound estimate of $\hat{\alpha}$ being 0.05 for HS with a *p*-value of .03. The lower-bound estimate for COLL is 0.04 but is not statistically significant. The estimates of ρ also change little relative to the bivariate probit case, which we view as evidence that condition 1, rather than joint normality of the unobservables, drives identification of the models in panel A of table 5.[20]

A. *The Relative Amount of Selection on Unobservables Required to Explain the CH Effect*

In this subsection we provide a different, more informal way to use information about selection on the observables as a guide to selection on the unobservables. Consider the following restriction, which uses the CH indicator directly.

CONDITION 4.

$$\frac{E(\epsilon|\text{CH} = 1) - E(\epsilon|\text{CH} = 0)}{\text{Var}(\epsilon)} = \frac{E(\boldsymbol{X}'\boldsymbol{\gamma}|\text{CH} = 1) - E(\boldsymbol{X}'\boldsymbol{\gamma}|\text{CH} = 0)}{\text{Var}(\boldsymbol{X}'\boldsymbol{\gamma})}.$$

This condition states that the relationship between CH and the mean of the distribution of the index of unobservables that determine outcomes is the same as the relationship between CH and the mean of the observable index, after adjusting for differences in the variance of these distributions. Altonji et al. (2002*b*) show that this condition is equivalent to condition 1 and holds under the same assumptions.

One way to gauge the strength of the evidence for a CH effect is to ask how large the ratio on the left side of condition 4 would have to be relative to the ratio on the right to account for the entire estimate of α under the null hypothesis that $\alpha = 0$. An advantage of this approach is that we do not have to simultaneously estimate the parameters of the CH and Y equations subject to (10). Consequently, we are able to use the full control set used in columns 4 and 8 of table 3.

To gauge the role of selection bias in a simple way, we ignore the fact that Y is estimated by a probit and treat α as though it were estimat-

[20] Unrestricted bivariate probit estimates of ρ and α for high school graduation are 0.13 (0.16) and 0.77 (1.12), which are quite close to the restricted estimates; but this is a matter of luck because the standard errors are very large. In the college attendance case we obtain a large and implausibly negative value of $\rho = -0.52$ (0.09) and an implausibly large but very imprecise estimate of $\alpha = 1.18$ (0.50). Without exclusion restrictions or a restriction such as condition 1, identification of α and ρ is strictly based on functional form and is very tenuous. The results are not informative about the Catholic school effect and the nature of selection bias, and this is reflected in part in the very large standard errors.

ed by a regression of the latent variable Y^* on $\boldsymbol{X}$ and CH. Let $\boldsymbol{X}'\boldsymbol{\beta}$ and $\widetilde{\mathrm{CH}}$ represent the predicted value and residuals of a regression of CH on $\boldsymbol{X}$ so that $\mathrm{CH} = \boldsymbol{X}'\boldsymbol{\beta} + \widetilde{\mathrm{CH}}$. Then

$$Y^* = \alpha\widetilde{\mathrm{CH}} + \boldsymbol{X}'(\boldsymbol{\gamma} + \alpha\boldsymbol{\beta}) + \epsilon.$$

If the bias in a probit is close to the bias in OLS applied to the above model, then the fact that $\widetilde{\mathrm{CH}}$ is orthogonal to $\boldsymbol{X}$ leads to the familiar formula

$$\begin{aligned}\operatorname{plim}\hat{\alpha} &\simeq \alpha + \frac{\operatorname{Cov}(\widetilde{\mathrm{CH}},\ \epsilon)}{\operatorname{Var}(\widetilde{\mathrm{CH}})} \\ &= \alpha + \frac{\operatorname{Var}(\mathrm{CH})}{\operatorname{Var}(\widetilde{\mathrm{CH}})}[E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)].\end{aligned}$$

Condition 4 allows us to use an estimate of $E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 1) - E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 0)$ to estimate the magnitude of $E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)$ and therefore this bias. (Note that when $\operatorname{Var}(\epsilon)$ is very large relative to $\operatorname{Var}(\boldsymbol{X}'\boldsymbol{\gamma})$, what one can learn is limited, because even a small shift in $[E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)]/\operatorname{Var}(\epsilon)$ is consistent with a large bias in α.) Under the null hypothesis of no CH effect, we can consistently estimate $\boldsymbol{\gamma}$, and thus $E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH})$, from a separate model imposing $\alpha = 0$. The results for HS are reported in the first row of table 6. The estimate of $[E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 1) - E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 0)]/\operatorname{Var}(\boldsymbol{X}'\boldsymbol{\gamma})$ is 0.24. That is, the mean/variance of the probit index of $\boldsymbol{X}$ variables that determine HS is 0.24 higher for those who attend Catholic high school than for those who do not. Since the variance of ϵ is 1.00, the implied estimate of $E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)$ if condition 4 holds is 0.24 (row 1, col. 3). Multiplying by $\operatorname{Var}(\mathrm{CH}_i)/\operatorname{Var}(\widetilde{\mathrm{CH}}_i)$ yields a bias of 0.29. The unconstrained estimate of α is 1.03, and column 6 of the table reports that the ratio

$$\frac{\hat{\alpha}}{[\operatorname{Var}(\mathrm{CH})/\operatorname{Var}(\widetilde{\mathrm{CH}})][E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)]} = \frac{1.03}{0.29} = 3.55.$$

That is, the normalized shift in the distribution of the unobservables would have to be 3.55 times as large as the shift in the observables to explain away the entire CH effect. This seems highly unlikely.

For college attendance the estimated ratio is 1.43 (row 2, col. 6). Since the ratio of selection on unobservables relative to selection on

TABLE 6

Amount of Selection on Unobservables Relative to Selection on Observables Required to Attribute the Entire Catholic School Effect to Selection Bias

Outcome	$[\hat{E}(X'\hat{\gamma}\mid CH = 1) - \hat{E}(X'\hat{\gamma}\mid CH = 0)] \div \widehat{Var}(X'\hat{\gamma})$ (1)	$\widehat{Var}(\hat{\epsilon})$ (2)	$E(\epsilon\mid CH = 1) - E(\epsilon\mid CH = 0)$[a] (3)	$Cov(\epsilon, \widetilde{CH}) \div Var(\widetilde{CH})$ (4)	$\hat{\alpha}$ (5)	Implied Ratio[b] (6)
	A. $\hat{\alpha}$ Estimated from the Catholic Eighth Grade Subsample, Full Set of Controls[c]					
High school graduation (N=859)	.24	1.00	.24	.29	1.03 (.31)	3.55
College attendance (N=834)	.39	1.00	.39	.47	.67 (.16)	1.43
12th grade reading (N=739)	.091	36.00	3.28	3.94	.33 (.62)	.08
12th grade math (N=739)	.038	24.01	.91	1.09	1.14 (.46)	1.04
	B. $\hat{\alpha}$ Estimated from the Urban Minority Subsample					
High school graduation (N=698)	.73	1.00	.73	.88	1.59 (.67)	1.81
College attendance (N=698)	.58	1.00	.58	.69	.68 (.30)	.99
12th grade reading (N=561)	.090	30.58	2.76	3.31	−.19 (1.39)	−.06
12th grade math (N=561)	.058	20.25	1.17	1.40	1.25 (1.09)	.89

Note.—Huber-White standard errors are in parentheses. The model is $Y = 1(X'\gamma + \alpha CH + \epsilon > 0)$ for high school graduation and college attendance, estimated as a probit, and $Y = X'\gamma + \alpha CH + \epsilon$ for twelfth grade test scores, estimated by OLS. The $\hat{\gamma}$ used to evaluate

$$\frac{\hat{E}(X'\hat{\gamma}\mid CH = 1) - \hat{E}(X'\hat{\gamma}\mid CH = 0)}{\widehat{Var}(X'\hat{\gamma})}$$

is estimated under the restriction $\alpha = 0$, using the Catholic eighth grade sample for panel A and the urban minority sample for panel B. The NELS:88 third follow-up and second follow-up panel weights are used for the educational attainment and twelfth grade models, respectively.

[a] If condition 4 holds. Condition 4 states that the standardized selection on unobservables is equal to the standardized selection on observables, i.e.,

$$\frac{E(\epsilon\mid CH = 1) - E(\epsilon\mid CH = 0)}{Var(\epsilon)} = \frac{E(X'\gamma\mid CH = 1) - E(X'\gamma\mid CH = 0)}{Var(X'\gamma)}.$$

[b] The implied ratio in col. 6 is the ratio of standardized selection on unobservables to observables under the hypothesis that there is no Catholic school effect.

[c] See notes a and b of table 3 for a description of the controls. In the urban minority sample, the indicator "black" is excluded.

observables is likely to be less than one, part of the CH effect on college graduation is probably real.[21]

[21] As a robustness check, we also used two separate methods for estimating γ in order to evaluate $E(X'\gamma\mid CH = 1) - E(X'\gamma\mid CH = 0)$, since bias in α will lead to bias in γ. The first method uses the γ from the public eighth grade sample to form the index $X'\gamma$ for each Catholic eighth grade student. In the case of high school graduation, the normalized shift in the distribution of the unobservables would have to be 2.78 times as large as the shift in the observables to explain away the entire Catholic school effect. When we evaluate

Rows 3 and 4 of the table present twelfth grade test score results. The variable CH has a positive and statistically significant coefficient only in the case of twelfth grade math scores, but this small effect of 1.14 (0.46) can be almost completely eliminated under the assumption that condition 4 holds. Even if selection on unobservables is only one-half as strong as that on observables, the effect of CH would be negligible and statistically insignificant. Given the weak evidence from the univariate models and the likelihood of some positive bias, we conclude that CH probably has little effect on test scores.

V. Results by Minority Status and Urban Residence

A number of studies, including Evans and Schwab (1995), Neal (1997), and Grogger and Neal (2000), have found much stronger effects of CH for minority students in urban areas than for other students. Table 2 reports differences in the means of outcomes and control variables, by high school type, for all urban minority students and for urban minority students who attended Catholic eighth grades. Note that 54 of the 56 minority students who attended Catholic high schools came from Catholic eighth grades. Only 15 of the 700 urban minority students in public tenth grades came from Catholic eighth grades, which is too few observations to support an analysis on the Catholic eighth grade subsample. In the full urban minority sample the control variables provide evidence of strong positive selection into Catholic high schools. The gaps in mother's education and father's education are 0.66 year and 1.69 years, respectively. The gap in the log of family income is 0.83. There are also very large discrepancies in the base year measures of parental expectations for schooling, student expectations for schooling and white-collar work, and the eighth grade behavioral measures and gaps of 6.49 and 3.28 in the eighth grade reading and math tests, respectively.

In table 7 we report univariate results from the urban sample of white students as well as the urban sample of minorities. All the regression models include our full set of controls. For the minority sample, the probit estimate implies that the average marginal effect of CH on HS is 0.191. One important caveat in interpreting these results is that of the 110 urban minority students who attend a Catholic high school, only one subsequently drops out. There clearly appears to be a strong CH effect on graduation, but one should be wary of small-sample bias in calculating the asymptotic standard errors. Turning to the second set

---

the left-hand side of condition 4 using the estimate of γ obtained from the single-equation probit estimate of the high school graduation equation on the Catholic school sample, the implied ratio is 4.29. For college attendance the corresponding ratios are 1.30 and 2.03.

TABLE 7
OLS, Fixed-Effect, and Probit Estimates of Catholic High School Effects by Race and Urban Residence: Full Set of Controls

	Sample			
	Urban and Suburban White Only (1)	Urban and Suburban Minorities Only (2)	Urban White Only (3)	Urban Minorities Only (4)
	A. High School Graduate			
Sample mean	.88	.80	.88	.80
	(N=3,799)	(N=1,308)	(N=1,002)	(N=697)
Probit	.443	.524	1.176	1.592
	(.279)	(.338)	(.417)	(.673)
	[.046]	[.085]	[.091]	[.191]
	B. College in 1994			
Sample mean	.37	.26	.32	.26
	(N=3,695)	(N=1,258)	(N=981)	(N=666)
Probit	.354	.697	.506	.677
	(.107)	(.201)	(.167)	(.303)
	[.087]	[.158]	[.110]	[.144]
	C. 12th Grade Reading Score			
Sample mean	52.94	47.72	53.33	47.61
	(N=3,638)	(N=1,051)	(N=978)	(N=561)
OLS	1.30	−.72	1.59	−.19
	(.44)	(.98)	(.67)	(1.39)
	D. 12th Grade Math Score			
Sample mean	53.09	47.33	53.90	48.88
	(N=3,638)	(N=1,053)	(N=979)	(N=563)
OLS	1.07	1.17	1.69	1.25
	(.35)	(.76)	(.52)	(1.09)

Note.—All models include controls for Hispanic origin, gender, region, city size, distance to the nearest Catholic school (five categories), family background, eighth grade tests, and other eighth grade measures. See notes a and b of table 3. NELS:88 third follow-up and second follow-up panel weights are used for the educational attainment and twelfth grade models, respectively. Huber-White standard errors are in parentheses. Marginal effects are in brackets. Marginal effects of probit models are computed as average derivatives of the probability of an outcome with respect to Catholic school attendance.

of results in table 7, we find a substantial effect of CH on college attendance, with estimates for the urban minority sample varying from 0.144 to 0.182 depending on the estimation methods. Consistent with previous work, the effects are generally larger for minorities than for the samples of whites. However, since there is more selection on observable variables for this subsample, it seems quite plausible that there could be more selection on unobservables as well and that this could explain the large measured CH effects.

Table 7 also presents test score results for the urban minority sample. We obtain a coefficient of −0.19 (1.39) for the twelfth grade reading score and a coefficient of 1.25 (1.09) for the twelfth grade math score.

Evidently, most or all of the substantial CH advantage for urban minorities in test scores disappears once we control for family background and eighth grade outcomes. This result reflects the large gap in the means of the controls in favor of minorities attending Catholic high schools. As one can see in the table, we obtain similar results when we add suburbanites and extend our analysis to a pooled urban/suburban minority subsample.

We also perform a sensitivity analysis based on the bivariate probit model (1)–(3) for the urban minority sample. Turning again to table 4, note that the raw differential in the high school graduation probability is 0.22 and the estimate of the CH effect under the assumption $\rho = 0$ is 0.176. The estimate is 0.132 when $\rho = 0.2$ and 0.013 when $\rho = 0.5$. Thus the correlation between the unobservables would have to be in the neighborhood of 0.5, a very large correlation, for one to conclude that the true effect of CH on the graduation rates of urban minorities is zero. This value seems unreasonable.

We also estimated the restricted bivariate probit model as in table 5 for urban minorities. We experienced computational difficulties in estimating the model for HS that we suspect are related to the fact that only one Catholic school attendee failed to graduate. For college attendance, we obtained an estimate of ρ of 0.5 and a negative but insignificant estimate of α. As a result of the computational problems, we focus on an analysis involving the differences in indices of observable variables based on condition 4. In panel B of table 6, under condition 4 and the null that $\alpha = 0$, the implied shift in $E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)$ is 0.73 in the case of HS and 0.58 in the case of COLL, which reflects strong selection on the observables that influence these outcomes. Still, selection on the unobservables would have to be 1.81 times as strong as selection on the observables to explain away the entire high school graduation effect. This seems very unlikely, suggesting that for urban minorities a substantial part of the estimated effect of CH on HS is real. On the other hand, we cannot rule out the possibility that much of the effect of CH on COLL is due to selection bias.

As we have already noted, there is little evidence that attending a Catholic high school improves the reading scores of minorities. Table 6 shows that in the case of twelfth grade reading scores,

$$\frac{E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 1) - E(\boldsymbol{X}'\boldsymbol{\gamma}|\mathrm{CH} = 0)}{\mathrm{Var}\,(\boldsymbol{X}'\boldsymbol{\gamma})} = 0.090.$$

Under condition 4 this amount of favorable selection on the observables implies an estimate of $E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0) = 3.28$. Since the point estimate of α is already negative, there is certainly no evidence that Catholic schools boost twelfth grade reading scores.

In the case of twelfth grade math, the point estimate of α is 1.25, and in row 4 of the table we report that the implied estimate of $E(\epsilon|\mathrm{CH} = 1) - E(\epsilon|\mathrm{CH} = 0)$ under condition 4 is 1.17. The implied ratio of selection on unobservables to selection on observables required to explain away the entire estimate of α is 0.89, which seems large given that a substantial part of the unexplained variance is due to unreliability in the tests.[22] Consequently, we would not rule out a small positive effect on math, but there is little evidence that CH substantially boosts the test scores of urban minorities.[23]

VI. Conclusion

Our analysis of the Catholic high school effect is based on the premise that for this problem the degree of selection on the observables in the rich NELS:88 data is informative about selection on unobserved characteristics. Our methodological contribution is to show how one can use such information to quantitatively assess the degree of selection bias. We have three main substantive findings. First, attending a Catholic high school substantially raises high school graduation rates. In the C8 sample, the standard multivariate analysis indicates that only 0.02 of the 0.105 Catholic high school advantage in graduation rates is explained by eighth grade outcomes and family background. We obtain a lower-bound estimate of 0.05 when we impose equality of selection of observables and unobservables and an upper-bound estimate of 0.08 when we assume that there is no selection on unobservables. While estimates that treat CH as exogenous almost certainly overstate the effect of Catholic high schools, the degree of selection on the unobservables would have to be much stronger than the degree of selection on the observables to explain away the entire effect. We also find that the estimate of the effect of CH on the probability of college attendance is very large (0.15) when CH is treated as exogenous, but the lower-bound estimates range between 0.02 and 0.03 depending on estimation details. We conclude that part of the effect of CH on college attendance is probably

[22] The estimates of the reliability of the math test reported in the NELS:88 documentation, while probably downward biased, are in the 0.87–0.90 range. Consequently, a substantial part of the test score residual probably reflects random variation in test performance and is unrelated to achievement levels.

[23] These test score findings are robust to the imputation procedures for dropouts described in Sec. II.*C*. In contrast, Grogger and Neal (2000) find some evidence for a Catholic school effect on minority test scores using median regression, particularly when they restore high school dropouts with missing test score data to the sample by simply assigning them 0. We have not fully investigated the source of the discrepancy but suspect that our use of a more extensive set of control variables, our imputation process, differences in the samples used, and differences between mean and median regression all play a role.

real, but the evidence is less clear-cut than in the high school graduation case.

Second, CH substantially raises the probability of high school graduation for urban minorities. Single-equation estimates of the impact on college attendance are also very large for this group, but the degree of positive selection on the observables that determine college attendance is sufficiently large that one could not rule out selection bias as the full explanation for the CH effect on college attendance. Third, we do not find much evidence that CH boosts test scores for the C8 sample or for urban minorities.

In closing, we caution against the potential for misuse of the idea of using observables to draw inferences about selection bias.[24] The assumptions required for condition 1 and condition 4 imply that it is dangerous to infer too much about selection on the unobservables from selection on the observables if the observables are small in number and explanatory power, or if they are unlikely to be representative of the full range of factors that determine an outcome.[25] The theoretical analysis in Altonji et al. (2002*b*) that we summarize here is only the start of the methodological work that is needed. Priorities include a Monte Carlo analysis of how the methods perform in the context of real-world examples and a systematic look at how the performance of our methods varies with the content of major data sets.

References

Altonji, Joseph G. 1988. "The Effects of Family Background and School Characteristics on Education and Labor Market Outcomes." Manuscript, Dept. Econ., Northwestern Univ.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2002*a*. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schools." Working Paper no. 9358 (December), NBER, Cambridge, MA.

———. 2002*b*. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." Manuscript (April), Dept. Econ., Northwestern Univ.

Angrist, Joshua D., and William N. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *A.E.R.* 88 (June): 450–77.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical Strategies in Labor

[24] Examples of questions that strike us as candidates for application of our methods include the effect of drugs and alcohol on future socioeconomic outcomes, the effect of criminal activity on future labor market success, and the effects of peer characteristics on school outcomes. Chatterji et al. (2003) have recently applied our methods to study the link from drinking to suicide. Krauth (2003) is applying them to study peer effects on youth smoking.

[25] Administrative data sets often cover some domains very well but lack the broad set of observables that our methods require.

Economics." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: North-Holland.

Bronars, Stephen G., and Jeff Grogger. 1994. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *A.E.R.* 84 (December): 1141–56.

Cameron, Stephen V., and Christopher R. Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *J.P.E.* 112 (February): 132–82.

Chatterji, Pinka, Dhaval Dave, Robert Kaestner, and Sara Markowitz. 2003. "Alcohol Abuse and Suicide Attempts among Youth—Correlation or Causation?" Working Paper no. 9638 (April), NBER, Cambridge, MA.

Chubb, John E., and Terry M. Moe. 1990. *Politics, Markets, and America's Schools.* Washington, DC: Brookings Inst.

Coleman, James S., and Thomas Hoffer. 1987. *Public and Private High Schools: The Impact of Communities.* New York: Basic Books.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. *High School Achievement: Public, Catholic, and Private Schools Compared.* New York: Basic Books.

Cookson, Peter W., Jr. 1993. "Assessing Private School Effects: Implications for School Choice." In *School Choice: Examining the Evidence*, edited by M. Edith Rasell and Richard Rothstein. Washington, DC: Econ. Policy Inst.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *A.E.R.* 85 (June): 341–64.

Engen, Eric M., William G. Gale, and John Karl Scholz. 1996. "The Illusory Effects of Saving Incentives on Saving." *J. Econ. Perspectives* 10 (Fall): 113–38.

Evans, William N., and Robert M. Schwab. 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Q.J.E.* 110 (November): 941–74.

Figlio, David N., and Joe A. Stone. 1999. "Are Private Schools Really Better?" In *Research in Labor Economics*, vol. 18, edited by Solomon W. Polachek. Stamford, CT: JAI.

Goldberger, Arthur S., and Glen C. Cain. 1982. "The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report." *Sociology of Educ.* 55 (April–July): 103–22.

Grogger, Jeffrey, and Derek Neal. 2000. "Further Evidence on the Benefits of Catholic Secondary Schooling." *Brookings-Wharton Papers Urban Affairs*, 151–93.

Heckman, James J., and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52 (March): 271–320.

Jacobsen, Joyce P., James W. Pearce III, and Joshua L. Rosenbloom. 1999. "The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment." *J. Human Resources* 34 (Summer): 449–74.

Jepsen, Christopher. 2003. "The Effectiveness of Catholic Primary Schooling." *J. Human Resources* 38 (Fall): 928–41.

Krauth, Brian. 2003. "Peer Effects and Selection Effects in Youth Smoking." Manuscript, Dept. Econ., Simon Frasier Univ., Burnaby, BC.

Ludwig, Jens. 1997. "Educational Achievement in Public, Private, and Catholic Schools: New Evidence on What We Know (and Don't Know)." Manuscript, School Public Policy, Georgetown Univ.

Murnane, Richard J. 1984. "A Review Essay—Comparisons of Public and Private Schools: Lessons from the Uproar." *J. Human Resources* 19 (Spring): 263–77.

Murphy, Kevin M., and Robert H. Topel. 1990. "Efficiency Wages Reconsidered: Theory and Evidence." In *Advances in the Theory and Measurement of Unemployment*, edited by Yoram Weiss and Robert H. Topel. New York: St. Martin's Press.

Neal, Derek. 1997. "The Effects of Catholic Secondary Schooling on Educational Attainment." *J. Labor Econ.* 15, no. 1, pt. 1 (January): 98–123.

———. 1998. "What Have We Learned about the Benefits of Private Schooling?" *Fed. Reserve Bank New York Econ. Policy Rev.* 4 (March): 79–86.

Poterba, James M., Steven F. Venti, and David A. Wise. 1994. "Targeted Retirement Saving and the Net Worth of Elderly Americans." *A.E.R. Papers and Proc.* 84 (May): 180–85.

Rosenbaum, Paul R. 1995. *Observational Studies.* New York: Springer-Verlag.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome." *J. Royal Statis. Soc. B* 45 (2): 212–18.

Sander, William. 2000. "Parochial Schools and Student Achievement: Findings for Older Adults." *Educ. Econ.* 8 (December): 259–68.

Tyler, Brian S. 1994. "An Analysis of Public and Catholic Secondary Education and the Earnings of Men." PhD diss., Dept. Econ., Univ. Chicago.

Udry, Christopher. 1996. "Gender, Agricultural Production, and the Theory of the Household." *J.P.E.* 104 (October): 1010–46.

Witte, John F. 1992. "Private School versus Public School Achievement: Are There Findings That Should Affect the Educational Choice Debate?" *Econ. Educ. Rev.* 11 (December): 371–94.

Wooldridge, Jeffrey M. 2000. *Introductory Econometrics: A Modern Approach.* Cincinnati: South-Western College Pub.