Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools[1]

Joseph G. Altonji[2]
Todd E. Elder[3]
Christopher R. Taber[4]

October 23, 2003

[1]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 anonomous referees, as well as participants in seminars at the American Economic Association Winter meetings (January 2000), Boston College, Cornell University, CREST-INSEE, Duke University, IZA, Johns Hopkins University, Harvard University, MIT, 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 at 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.

[2]Department of Economics and Economic Growth Center, Yale University and NBER

[3]Department of Economics and Institute of Labor and Industrial Relations, University of Illinois at Urbana-Champaign

[4]Department of Economics and Institute for Policy Research, Northwestern University and NBER

Abstract

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 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, college attendance. We find little evidence of an effect on test scores.

1 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 post secondary 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 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.

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

[2] Several recent studies, including Evans and Schwab (1995), Neal (1997), Grogger and Neal (2000), Figlio and Stone (2000) and Altonji, Elder and Taber (2002b) 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, Elder and Taber (2002b) raise doubt about the validity of all of 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.

We then use our strategies to assess the effectiveness of Catholic schools. The formal econometric theory is presented in Altonji, Elder, and Taber (2002a, hereafter AET).

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 importantly, we provide a way to quantitatively assess the degree of omitted variables bias.[4]

In section 2 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 12th 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 "control" groups by focusing on the Catholic eighth grade sample, in contrast to most of the literature. 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 12th 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 4 year college 2 years after high school and 0.08 on the probability of graduating 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 3 and 4 we present and apply our methods for using the degree of selection

[3]See for example, Currie and Thomas (1995), Engen et al (1996), Poterba et al (1994), Angrist and Evans (1988), Jacobsen et al. (1999), Bronars and Grogger (1994), Udry (1996),Cameron and Taber (2004), or Angrist and Krueger (1999). 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 especially Murphy and Topel's (1990) study of the importance of selection on unobserved ability as an explanation for industry wage differentials.

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 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. OLS or probit models assume 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 5 we obtain larger univariate effects for urban minorities but also stronger evidence of selection. We conclude the paper in section 6.

2 Preliminary Analysis of the Catholic School Effect

2.1 Data

Our data set is NELS:88, a National Center for Education Statistics (NCES) survey which began in the Spring of 1988. A total of 1032 schools contributed up to 26 eighth graders to the base year survey, resulting in 24,599 eighth graders participating. 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 due to 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 post-secondary 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-up. We use the 8th grade test scores as control variables and 12th grade reading and math tests as outcome measures.

The high school graduation variable HS is equal to one if the respondent graduated 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

[5] Our major findings are robust to whether or not college attendance is limited to 4-year universities, full-time versus part-time, or enrolled 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 AET, footnote 11 for details.

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 AET, Appendix A and B.

2.2 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 versus 0.15), and are almost twice as likely to be enrolled in a four year college in 1994 (0.59 versus 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 versus 0.12) as is the gap in the college attendance rate (0.61 versus 0.38). In contrast, the gap in the 12th 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.

Tables 1 and 2 also show that the means of favorable family background measures, 8th 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

[7] See AET, footnote 12, for details on the sampling scheme.

[8] In Table 1 and Table 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.

[9] See the footnotes 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

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 0 in the C8 sample; both of these variables are powerful predictors of HS. Finally, the gap in the 8th 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 8th 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% 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]

2.3 Probit and OLS Estimates of the Effect of Catholic High Schools

The top panel 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 (column 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

influence expectations.

[10] In unreported results, the pattern of positive selection on the 8th 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%. The percentage is 0.3% in the sample of 16,070 individuals for whom information on sector of eighth-grade and sector of 10th grade is available. It is only 0.7% among children whose parents are Catholic.

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 as 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 Table 3 we also report estimates of the effect of CH on the probability that a student is enrolled in a 4 year college at the time of the 3rd follow-up survey in 1994, 2 years after most students graduate from high school. The raw difference of 0.236 (column 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 the bottom half of Table 3 we report estimates of the effect of CH on 12th 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 12th grade test equations given that dropouts have lower test scores and that dropouts have a lower probability of taking the 12th 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 (Table 3, columns 1-4). 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 make 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,

[12] We obtain similar results when we estimate linear probability models and linear probability models with fixed effects for each eighth grade. See AET. 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.

[13] We address the issue by imputing missing data for both high school graduates and dropouts using predicted values from a regression of the 12th grade score on the full set of controls in the outcome regression, plus the Catholic high school dummy and the 10th 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 AET for details.

the standards for being a "trouble maker" 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 of 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 12th 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.

2.4 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 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

$$CH = 1(X'\beta + u > 0) \tag{2.1}$$

$$Y = 1(X'\gamma + \alpha CH + \varepsilon > 0) \tag{2.2}$$

$$\begin{bmatrix} u \\ \varepsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right). \tag{2.3}$$

While this 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 section is to treat this model as if it were underidentified by one parameter. In particular, we act as if ρ is not identified.

In Table 5 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]

[14] See Rosenbaum (1995) or Rosenbaum and Rubin (1983) for examples of this type of sensitivity analysis.

We report results for HS in the top panel and college attendance in the bottom panel, and 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 0 (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 if $\rho = 0.2$. 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 for a strong CH effect is considerably weaker than suggested by the estimates that take CH as exogenous.

For our preferred sample of Catholic 8th 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% level when $\rho = 0.3$ and is still positive when $\rho = 0.5$, although it is not significant. Thus, ignoring sampling error for the moment, 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 the bottom panel of Table 5 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 ρ is 0.2 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

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 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" from the set of controls.

the analysis after generalizing this assumption in the semiparametric specification

$$u = \theta + u^* \tag{2.4}$$

$$\varepsilon = \theta + \varepsilon^*, \tag{2.5}$$

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 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 3 points of support for θ. The (unreported) results are similar to those in Table 5. For example, in the full sample the effect of CH on high school graduation is 0.034 when $\rho = .2$ and -.0.055 when $\rho = 0.5$. In the C8 sample the estimates are .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 $\rho =.2$ and -0.049 when $\rho = .5$.

Summarizing the results 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 CH 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 HS case. 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. In order to solve this problem, in the next sections we use the degree of selection on the observables as a guide.

3 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. Using our Catholic school application, let the outcome of interest Y be a function of a latent variable Y^*, which is determined as

$$
\begin{aligned}
Y^* &= \alpha CH + W'\Gamma \\
&= \alpha CH + X'\Gamma_X + \xi, \qquad (3.1)
\end{aligned}
$$

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^*, W is the full set variables (observed and unobserved) that determine Y^* and Γ is the causal effect of W on Y^*. In the second line of (3.1), X is a vector of the observable components of W, Γ_X is the corresponding subvector of Γ, and the error component ξ is an index of the unobserved variables. Because it is unlikely that the control variables X are all unrelated to ξ, we work with

$$
Y^* = \alpha CH + X'\gamma + \varepsilon, \qquad (3.2)
$$

where γ and ε are *defined* so that $\text{cov}(\varepsilon, X) = 0$. Consequently, γ captures both the direct effect of X on Y^*, Γ_X, as well as the relationship between X and ξ. Let CH^* be the latent variable that determines CH such that $CH = 1(CH^* > 0)$, where the indicator function $1(\cdot)$ is 1 when $CH^* > 0$ and 0 otherwise. Consider the linear projection of CH^* onto $X'\gamma$ and ε,

$$
\text{Proj}(CH^*|X'\gamma, \varepsilon) = \phi_0 + \phi_{X'\gamma}X'\gamma + \phi_\varepsilon \varepsilon. \qquad (3.3)
$$

We formalize the idea that "selection on the unobservables is the same as selection on the observables" as

Condition 1

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

We contrast this with the OLS condition,

Condition 2

$$
\phi_\varepsilon = 0.
$$

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 AET. The following three types of assumptions suffice:

1. the elements of X are chosen at random from the full set of factors W that determine Y;

2. the numbers of elements in X and 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 X and the unobservables obeys an assumption that is very strong but weaker than the standard assumption $cov(X, \xi) = 0$. Roughly speaking, the assumption is that the regression of CH^* on $Y^* - \alpha CH$ is equal to the regression of the part of CH^* that is orthogonal to X on the corresponding part of $Y^* - \alpha CH$.[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: $Cov(CH, \xi) = 0$ and $Cov(X, \xi) = 0$.[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 NLSY, NELS:88, the PSID, 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. Data set content 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 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 W are left out. This is reflected in the relatively low explanatory power of social science models of

[16] The condition is

$$\frac{\sum_{\ell=-\infty}^{\infty} E\left(W_j W_{j-\ell}\right) E\left(\beta_j \Gamma_{j-\ell}\right)}{\sum_{\ell=-\infty}^{\infty} E\left(W_j W_{j-\ell}\right) E\left(\Gamma_j \Gamma_{j-\ell}\right)} = \frac{\sum_{\ell=-\infty}^{\infty} E(\tilde{W}_j \tilde{W}_{j-\ell}) E\left(\beta_j \Gamma_{j-\ell}\right)}{\sum_{\ell=-\infty}^{\infty} E(\tilde{W}_j \tilde{W}_{j-\ell}) E\left(\Gamma_j \Gamma_{j-\ell}\right)}, \tag{3.4}$$

where $\tilde{W}_j$ is the component of W_j that is orthogonal to X. It is easy to show that this condition holds under the standard assumption $E(\xi \mid X) = 0$. However, $E(\xi \mid X) = 0$ is not likely to hold and fortunately is not necessary for (3.4). AET present an example of a model for which (3.4) holds, but $E(\xi \mid X) \neq 0$.

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

individual behavior. Furthermore, in many applications, including ours, the endogenous variable is correlated with many of the elements of X. Given the constraints that shape the choice of X and the fact that many of the elements of 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 $Cov(CH, \xi) = 0$ requires. Since the X variables are inter-correlated, the assumption that $Cov(X, \xi) = 0$ 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^* and/or X 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 X as a more or less random subset of the elements of 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 suggests 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, 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 education 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 12th 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 X. These will influence high school outcomes but not the probability of starting a Catholic high school. To see this, rewrite ε as $\varepsilon = \varepsilon_1 + \varepsilon_2$, 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

$$\phi_{X'\gamma} \equiv \frac{cov(CH^*, X'\gamma)}{var(X'\gamma)} = \frac{cov(CH^*, \varepsilon_1)}{var(\varepsilon_1)}. \tag{3.5}$$

Assume without loss of generality that $cov(CH^*, X'\gamma) \geq 0$ as is true in our data. Since $var(\varepsilon) > var(\varepsilon_1)$ and $cov(CH^*, \varepsilon) = cov(CH^*, \varepsilon_1)$ then

$$\phi_\varepsilon \equiv \frac{cov(CH^*, \varepsilon)}{var(\varepsilon)} \leq \frac{cov(CH^*, \varepsilon_1)}{var(\varepsilon_1)} = \phi_{X'\gamma}.$$

Since $cov(CH^*, \varepsilon_1) \geq 0$ and $\phi_\varepsilon \geq 0$, Condition 1 is replaced by Condition 3

Condition 3

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

In AET we prove that we can identify the set of values of α that satisfy Condition 3. For our data and empirical specification we find that the upper bound on α occurs when one assumes that $\frac{cov(CH^*,\varepsilon)}{var(\varepsilon)} = 0$ and the lower bound occurs when one assumes that $\frac{cov(CH^*,\varepsilon)}{var(\varepsilon)} = \frac{cov(CH^*,X'\gamma)}{var(X'\gamma)}$. 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.**

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

We now return to the bivariate probit model given by (2.1), (2.2), and (2.3). In the bivariate probit case, Condition 3 may be re-written[18] as

[18] Keep in mind that in the binary probit the variances of ε and u are normalized to 1.

$$0 \leq \rho \leq \frac{Cov(X'\beta, X'\gamma)}{Var(X'\gamma)}. \tag{4.1}$$

In the top panel of Table 6, we present estimates that use the $C8$ sample directly and maximize the likelihood subject to $\rho = \frac{Cov(X'\beta,X'\gamma)}{Var(X'\gamma)}$. The standard errors assume that (4.1) holds for the particular 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 95% confidence interval for the bound includes zero.

The results for $COLL$ follow a similar pattern, but $\rho = \frac{Cov(X'\beta,X'\gamma)}{Var(X'\gamma)}$ leads to a larger reduction in the estimated effect of CH on college attendance probability. The point estimate of 0.03 is still sizeable, 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 8th graders. We partition X and γ into the subvectors $\{X_1, X_2, ..., X_G\}$ and $\{\gamma_1, \gamma_2, ..., \gamma_G\}$ consisting of variables and parameters that fall into similar categories. In practice, G is 6. We estimate γ on the public 8th grade sample on the grounds that very few such students go to Catholic high school, and so selectivity will not influence the estimates of γ even though the mean of the error term may be different for this sample. We assume that the values of γ are the same for students from Catholic and public 8th grades, up to a proportionality factor for each subvector, which slightly relaxes the implicit assumption of the full sample models in Table 3 that γ does not depend on the sector of the 8th grade.[19] The results using the second estimation method are reported in the middle panel of Table 6. In the case of HS, $\hat{\rho}$ is only 0.09, $\hat{\alpha}$ is 0.94 with a p-value of .002, and the effect on the high school graduation probability is 0.09. However, the college effect is only 0.02.

As a further robustness check, in the bottom panel of Table 6 we replace the joint normality assumption implicit in the bivariate probit with the semiparametric specification presented in equations (2.4) and (2.5). The results do not change substantially, with the lower bound estimate of $\hat{\alpha}$ being 0.05 for HS with a p-value of 0.03. The lower bound

[19] The restrictions on γ pass with a p-value of .12 in the HS case, but fail with a p-value of .03 in the $COLL$ case, so perhaps the method 2 results for $COLL$ should be discounted. Details are in Table 6 note 4.

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 the top panel of Table 6.[20]

4.1 The Relative Amount of Selection on Unobservables Required to Explain the CH Effect

In this section 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

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

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. AET 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 α is zero. An advantage of this approach is that we do not have to simultaneously estimate the parameters of the CH and Y equations subject to (4.1). 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 if it were estimated by a regression of the latent variable Y^* on X and CH. Let $X'\beta$ and $\widetilde{CH}$ represent the predicted value and residuals of a regression of CH on X so that $CH = X'\beta + \widetilde{CH}$. Then,

$$Y^* = \alpha\widetilde{CH} + X'[\gamma + \alpha\beta] + \varepsilon.$$

[20] Unrestricted bivariate probit estimates of ρ and α for high school graduation are 0.13 and .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 ρ equal to -0.52 and an implausibly large but very imprecise estimate of α equal to 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.

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

$$\begin{aligned} \text{plim } \widehat{\alpha} &\simeq \alpha + \frac{cov(\widetilde{CH}, \varepsilon)}{var\left(\widetilde{CH}\right)} \\ &= \alpha + \frac{var\,(CH)}{var\left(\widetilde{CH}\right)} \left[E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0)\right]. \end{aligned}$$

Condition 4 allows us to use an estimate of $E(X'\gamma \mid CH = 1) - E(X'\gamma \mid CH = 0)$ to estimate the magnitude of $E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0)$, and therefore this bias.[21] Under the null hypothesis of no CH effect, we can consistently estimate γ, and thus $E(X'\gamma \mid CH)$, from a separate model imposing $\alpha = 0$. The results for HS are reported in the first row of Table 7. The estimate of $\left(E(X'\gamma \mid CH = 1) - E(X'\gamma \mid CH = 0)\right) / Var(X'\gamma)$ is 0.24. That is, the mean/variance of the probit index of X variables that determine HS is 0.24 higher for those who attend CH than for those who do not. Since the variance of ε is 1.00, the implied estimate of $E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0)$ if Condition 4 holds is 0.24 (row 1, column 3). Multiplying by $var\,(CH_i)\,/var\left(\widetilde{CH}_i\right)$ yields a bias of 0.29. The unconstrained estimate of α is 1.03, and the last column of the table reports that the ratio $\widehat{\alpha}/[\frac{var(CH)}{var\left(\widetilde{CH}\right)}(E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0))] = 1.03\ /\ 0.29 = 3.55$. 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, column (6)). Since the ratio of selection on unobservables relative to selection on observables is likely to be less than 1, part of the CH effect on college graduation is probably real.[22]

Rows 3 and 4 of the table present 12th grade test score results. CH has a positive and statistically significant coefficient only in the case of 12th grade math scores, but this small effect of 1.14 (0.46) can be almost completely eliminated assuming 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

[21] Note that when $var(\varepsilon)$ is very large relative to $var(X'\gamma)$,what one can learn is limited, because even a small shift in $\left(E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0)\right) / var(\varepsilon)$ is consistent with a large bias in α.

[22] 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 8th 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 the left hand side of Condition 4 evaluating 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.

univariate models and the likelihood of some positive bias, we conclude that CH probably has little effect on test scores.

5 Results by minority status and urban residence

A number of studies using NELS:88, 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 CH came from Catholic eighth grades. Only 15 of the 700 urban minority students in public 10th 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 years 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 4 we report univariate results from the urban sample of white students as well as the urban sample of minorities. All of 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 CH, 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 of results in Table 4, 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 4 also presents test score results for the urban minority sample. We obtain a coefficient of -0.19 (1.39) for the 12th grade reading score and a coefficient of 1.25 (1.09) for the 12th 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 8th grade outcomes. This result reflects the large gap in the means of the controls in favor of minorities attending CH. 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 (2.1)-(2.3) for the urban minority sample. Turning again to Table 5, 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 0. This value seems unreasonable.

We also estimated the restricted bivariate probit model as in Table 6 for urban minorities. We experienced computational difficulties in estimating the model for HS that we suspect are related to the fact that only 1 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. Due to the computational problems, we focus on an analysis involving the differences in indices of observable variables based on Condition 4. In the bottom panel of Table 7, under Condition 4 and the null that α is 0 the implied shift in $(E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0))$ 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 to us; the evidence suggests 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 CH improves the reading scores of minorities. Table 7 shows that in the case of 12th grade reading scores $(E(X'\gamma \mid CH = 1) - E(X'\gamma \mid CH = 0))\ /Var(X'\gamma)$ is 0.090. Under Condition 4 this amount of favorable selection on the observables implies an estimate of $(E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH_i = 0))$ equal to 3.28. Since the point estimate of α is already negative, there is certainly no evidence that Catholic schools boost 12th grade reading scores.

In the case of 12th grade math, the point estimate of α is 1.25 and in row 4 of the table we report that the implied estimate of $(E(\varepsilon \mid CH = 1) - E(\varepsilon \mid CH = 0))$ 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.[23] Consequently, we would not rule out a small positive effect on math but there is little evidence that CH substantially boost the test scores of urban minorities.[24]

6 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 CH 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 ranges between 0.02 and 0.03 depending on estimation details. We conclude that part of the effect of CH on college attendance is probably 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

[23] The estimates of the reliability of the math test reported in the NELS:88 documentation, while probably downward biased, are in the 0.87 to 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.

[24] These test score findings are robust to the imputation procedures for dropouts described in Section 2.3. 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.

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.[25] 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.[26] The theoretical analysis in AET 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.

[25] 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, Dave, Kaestner, and Markowitz (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.

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

7 References

Altonji, Joseph G., "The Effects of Family Background and School Characteristics on Education and Labor Market Outcomes," unpublished manuscript, Northwestern University, 1988.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," NBER Working Paper 7831, 2002a.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "IV Strategies for Evaluating the Catholic School Effect," NBER Working Paper 9358, Northwestern University, 2002b.

Angrist and Evans, "Children and their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 88 (1988), 450-477.

Angrist, Joshua D., and Alan B. Krueger, "Empirical Strategies in Labor Economics," *Handbook of Labor Economics Vol. 3A,* Ashenfelter and Card (eds.), North Holland, 1999.

Bronars, Stephen G. , and Jeff Grogger, "The Economic Consequences of Unwed Motherhood: Using Twins as a Natural Experiment," *American Economic Review* 80(1994), 1141-56.

Cameron, Stephen V. and Christopher R. Taber, "Estimating Borrowing Constraints Using the Returns to Schooling," forthcoming, *Journal of Political Economy*, 2004.

Chatterji, Pinka, Dave Dhaval, Robert Kaestner, and Sara Markowitz, "Alcohol Abuse and Suicide Attempts Among Youth - Correlation or Causation?" NBER Working Paper #9638, 2003.

Chubb, John E., and Terry M. Moe, *Politics, Markets, and America's Schools* (Washington, D.C.: The Brookings Institution, 1990).

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

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

Cookson, Peter W., Jr., "Assessing Private School Effects: Implications for School Choice," in Edith Rasell and Richard Rothstein, eds., *School Choice: Examining the Evidence* (Washington, D.C.: Economic Policy Institute, 1993).

Currie, Janet, and Thomas Duncan, "Does Head Start Make a Difference?" *American Economic Review*, 85 (1990), 341-64.

Engen, Eric, William Gale, and John Karl Sholz, "The Illusory Effects of Saving Incentives on Saving," *Journal of Economic Perspectives,* 10 (1996), 113-138.

Evans, William N., and Robert M. Schwab, "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics,* 110 (1995), 947-974.

Figlio, David N., and Joe A. Stone, "Are Private Schools Really Better?," forthcoming *Research in Labour Economics*, (2000).

Goldberger, Arthur S., and Glen C. Cain, "The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report," *Sociology of Education*, LV (1982), 103-122.

Grogger, Jeff, and Derek Neal, “Further Evidence on the Benefits of Catholic Secondary Schooling,” *Brookings-Wharton Papers on Urban Affairs* (2000), 151-193.

Heckman, J., and B. Singer, “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data,” *Econometrica,* 52 (1984), 271-320.

Jacobsen, Joyce P., James W. Pearce III, and Joshua L. Rosenbloom, “The Effect of Childbearing on Married Women’s Labor Supply and Earnings,” *Journal of Human Resources* 34(3), Summer 1999, pp. 449-474.

Jepsen, Christopher, “The Effectiveness of Catholic Primary Schools,” unpublished manuscript, Northwestern University, 2000.

Krauth, Brian, "Peer Effects and Selection Effects in Youth Smoking," unpublished manuscript, Simon Frasier University, 2003.

Ludwig, Jens, “Educational Achievement in Public, Private, and Catholic Schools: New Evidence on What We Know (and Don’t Know),” working paper, Georgetown University, 1997.

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

Murphy, Kevin M., and Robert H. Topel, “Efficiency Wages Reconsidered: Theory and Evidence,” in Y. Weiss and R. Topel eds., *Advances in the Theory and Measurement of Unemployment.* New York, St. Martin’s Press, 1990, 204-40.

Neal, Derek, “The Effects of Catholic Secondary Schooling on Educational Attainment,” *Journal of Labor Economics* 15 (1997), 98-123.

Neal, Derek, “What Have We Learned About the Benefits of Private Schooling?” *Economic Policy Review,* Vol. 4, (1998), 79-86.

Poterba, James, Steven Venti, and David Wise, “Targeted Retirement Saving and the Net Worth of Elderly Americans,” *American Economic Review,* 84 (1994), 180-185.

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

Rosenbaum, Paul, and Donald Rubin, “Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome,” *Journal of the Royal Statistical Society B* 45, no. 2, (1083), 212-218.

Sander, William, “Parochial Schools and Students Achievement: Findings for Older Adults,” *Education Economics* 8 (2000), 259-68.

Tyler, Brian S., “An Analysis of Public and Catholic Secondary Education and the Earnings of Men,” University of Chicago Ph.D. Thesis (1994).

Udry, Christopher, “Gender, Agricultural Production, and the Theory of the Household,” *Journal of Political Economy*, 104 (1996), 1010-1046.

Witte, John F., “Private School Versus Public School Achievement: Are There Findings That Should Affect the Educational Choice Debate,” *Economics of Education Review*, XI (1992), 371-394.

Wooldridge, Jeffrey, *Introductory Econometrics*, South-Western College Publishing, 2000.

Table 1
Comparison of Means of Key Variables by Sector

Variable	Full Sample			Catholic 8th Grade		
	Public 10th (N=11,167)	Cath 10th (N=672)	Difference	Public 10th (N=366)	Cath 10th (N=640)	Difference
Demographics						
FEMALE	0.52	0.45	-0.07	0.61	0.50	-0.11
ASIAN	0.03	0.04	0.01	0.05	0.05	0.00
HISPANIC	0.09	0.09	0.00	0.08	0.09	0.01
BLACK	0.10	0.09	-0.01	0.07	0.11	0.04
WHITE	0.78	0.78	0.00	0.80	0.74	-0.06
Family Background						
MOTHER'S EDUCATION IN YEARS	13.21	13.96	0.75	13.34	13.88	0.54
FATHER'S EDUCATION IN YEARS	13.49	14.51	1.01	13.39	14.38	0.99
LOG OF FAMILY INCOME	10.23	10.72	0.49	10.47	10.66	0.19
MOTHER ONLY IN HOUSE	0.14	0.09	-0.05	0.07	0.09	0.02
PARENT MARRIED	0.79	0.89	0.10	0.90	0.88	-0.02
PARENTS CATHOLIC	0.28	0.82	0.54	0.84	0.84	0.00
Geography						
RURAL	0.36	0.03	-0.33	0.13	0.01	-0.12
SUBURBAN	0.45	0.51	0.06	0.40	0.48	0.08
URBAN	0.19	0.46	0.27	0.47	0.51	0.04
DISTANCE TO CLOSEST CATHOLIC HS, MILES	22.16	2.97	-19.19	6.91	2.37	-4.53
Expectations [1]						
SCHOOLING EXPECTATIONS IN YEARS	15.25	15.97	0.72	15.52	15.92	0.40
VERY SURE TO GRADUATE HS	0.84	0.89	0.05	0.84	0.90	0.06
PARENTS EXPECT AT LEAST SOME COLLEGE	0.89	0.98	0.09	0.94	0.98	0.04
PARENTS EXPECT AT LEAST COLLEGE GRAD	0.79	0.92	0.13	0.88	0.91	0.03
STUDENT EXPECTS WHITE-COLLAR JOB	0.47	0.61	0.14	0.55	0.59	0.04
8th Grade Variables						
DELINQUENCY INDEX, RANGE FROM 0 TO 4	0.64	0.53	-0.11	0.54	0.46	-0.08
STUDENT GOT INTO FIGHT	0.24	0.23	-0.02	0.20	0.19	-0.01
STUDENT RARELY COMPLETES HOMEWORK	0.19	0.08	-0.11	0.08	0.06	-0.01
STUDENT FREQUENTLY DISRUPTIVE	0.12	0.08	-0.05	0.08	0.08	0.00
STUDENT REPEATED GRADE 4-8	0.06	0.02	-0.05	0.03	0.02	-0.01
RISK INDEX, RANGE FROM 0 TO 4	0.69	0.35	-0.34	0.39	0.39	0.00
GRADES COMPOSITE	2.94	3.16	0.22	3.09	3.20	0.11
UNPREPAREDNESS INDEX, FROM 0 TO 25	10.77	11.08	0.31	10.84	11.02	0.17
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						
10TH GRADE READING STANDARDIZED SCORE	51.02	54.69	3.66	54.63	54.62	-0.01
10TH GRADE MATH STANDARDIZED SCORE	51.12	55.03	3.91	53.40	54.52	1.12
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 4 YEAR COLLEGE IN 1994	0.31	0.59	0.28	0.38	0.61	0.23
HS GRADUATE	0.85	0.98	0.13	0.88	0.98	0.10

Notes:
(1) 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	Cath 10th	Difference	Public 10th	Cath 10th	Difference
Demographics	(N=700)	(N=56)		(N=15)	(N=54)	
FEMALE	0.57	0.57	0.00	0.60	0.61	0.01
ASIAN	0.00	0.00	0.00	0.00	0.00	0.00
HISPANIC	0.44	0.49	0.05	0.34	0.45	0.11
BLACK	0.56	0.51	-0.05	0.66	0.55	-0.11
WHITE	0.00	0.00	0.00	0.00	0.00	0.00
Family Background						
MOTHER'S EDUCATION, IN YEARS	12.61	13.27	0.66	13.58	13.21	-0.37
FATHER'S EDUCATION, IN YEARS	12.64	14.33	1.69	12.66	14.36	1.70
LOG OF FAMILY INCOME	9.62	10.45	0.83	10.16	10.38	0.22
MOTHER ONLY IN HOUSE	0.29	0.27	-0.02	0.29	0.23	-0.06
PARENT MARRIED	0.57	0.74	0.18	0.71	0.79	0.08
PARENTS CATHOLIC	0.39	0.58	0.19	0.39	0.55	0.16
Geography						
RURAL	0.00	0.00	0.00	0.00	0.00	0.00
SUBURBAN	0.00	0.00	0.00	0.00	0.00	0.00
URBAN	1.00	1.00	0.00	1.00	1.00	0.00
DISTANCE TO CLOSEST CATHOLIC HS, MILES	6.04	1.90	-4.14	1.90	2.01	0.11
Expectations [1]						
SCHOOLING EXPECTATIONS, IN YEARS	15.27	16.10	0.83	16.48	16.05	-0.43
VERY SURE TO GRADUATE HS	0.80	0.94	0.14	0.88	0.94	0.06
PARENTS EXPECT AT LEAST SOME COLLEGE	0.90	0.99	0.09	0.95	0.99	0.04
PARENTS EXPECT AT LEAST COLLEGE GRAD	0.78	0.86	0.08	0.84	0.85	0.01
STUDENT EXPECTS WHITE-COLLAR JOB	0.53	0.72	0.19	0.50	0.70	0.20
8th Grade Variables						
DELINQUENCY INDEX, RANGE FROM 0 TO 4	0.88	0.63	-0.25	1.22	0.65	-0.57
STUDENT GOT INTO FIGHT	0.34	0.19	-0.15	0.05	0.19	0.15
STUDENT RARELY COMPLETES HOMEWORK	0.25	0.13	-0.12	0.23	0.14	-0.09
STUDENT FREQUENTLY DISRUPTIVE	0.19	0.17	-0.02	0.14	0.17	0.03
STUDENT REPEATED GRADE 4-8	0.11	0.05	-0.06	0.10	0.05	-0.05
RISK INDEX, RANGE FROM 0 TO 4	1.30	0.90	-0.40	1.05	0.91	-0.14
GRADES COMPOSITE	2.78	2.88	0.09	3.01	2.88	-0.13
UNPREPAREDNESS INDEX, FROM 0 TO 25	10.99	11.28	0.29	11.10	11.27	0.17
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	-0.27
Outcomes						
10TH GRADE READING STANDARDIZED SCORE	47.14	51.46	4.32	48.62	50.75	2.13
10TH GRADE MATH STANDARDIZED SCORE	45.80	48.92	3.12	48.16	48.09	-0.07
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	-0.54
ENROLLED IN 4 YEAR COLLEGE IN 1994	0.23	0.52	0.28	0.28	0.56	0.28
HS GRADUATE	0.78	0.99	0.21	0.89	1.00	0.11

Notes:
(1) The Expectations variables are not included in our empirical models

Table 3

OLS and Probit Estimates of Catholic High School Effects[1,2]
in Subsamples of NELS:88
Weighted, (Huber-White Standard Errors in Parentheses)
[Marginal Effects in Brackets[3]]

	Full Sample				*Catholic 8th Grade Attendees*			
	Controls							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	None	Fam. BG, city size, and region.[4]	(2) plus 8th grade tests	(3) plus other 8th grade measures[5]	None	Fam. BG, city size, and region.[4]	(2) plus 8th grade tests	(3) plus other 8th grade measures[5]
HS Graduation								
Probit	0.97	0.57	0.48	0.41	0.99	0.88	0.95	1.27
	(0.17)	(0.19)	(0.22)	(0.21)	(0.24)	(0.25)	(0.27)	(0.29)
	[0.123]	[0.081]	[0.068]	[0.052]	[0.105]	[0.084]	[0.081]	[0.088]
Pseudo R^2	0.01	0.16	0.21	0.34	0.11	0.35	0.44	0.58
College in 1994								
Probit	0.73	0.37	0.33	0.32	0.60	0.48	0.56	0.60
	(0.08)	(0.09)	(0.09)	(0.09)	(0.13)	(0.15)	(0.15)	(0.15)
	[0.283]	[0.106]	[0.084]	[0.074]	[0.236]	[0.154]	[0.154]	[0.149]
Pseudo R^2	0.02	0.19	0.29	0.34	0.04	0.18	0.29	0.36
12th Grade Reading Score								
OLS	4.28	2.08	1.18	1.14	1.92	0.17	0.37	0.33
	(0.47)	(0.54)	(0.38)	(0.38)	(0.82)	(0.98)	(0.63)	(0.62)
R^2	0.01	0.19	0.60	0.60	0.01	0.19	0.59	0.62
12th Grade Math Score								
OLS	4.86	1.98	1.07	0.92	2.79	1.10	1.46	1.14
	(0.44)	(0.54)	(0.34)	(0.32)	(0.77)	(1.00)	(0.53)	(0.46)
R^2	0.01	0.26	0.72	0.74	0.02	0.26	0.73	0.77

Notes:
(1) NELS:88 third follow-up and 2nd follow-up panel weights used for the educational attainment and 12th grade models, respectively.
(2) Sample sizes for Full sample: N=8560 (HS Graduation), N=8315 (College Attendance), N=8116 (12th Reading), N=8119 (12th Math).
For Catholic 8th Grade sample, N=859 (HS Graduation), N=834 (College Attendance), N=739 (12th Reading), N=739 (12th Math).
(3) Marginal effects of probit models are computed as average derivatives of the probability of an outcome with respect to catholic high school attendance.
(4) Control sets (2)-(4) include race (white/nonwhite), hispanic origin, gender, urbanicity (3 categories), region (8 categories), and distance to the nearest Catholic high school (5 categories). Family background variables used as controls include log family income, mother's and father's education, 5 dummy variables for marital status of the parents, and 8 dummy variables for household composition.
(5) "Other 8th 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 Appendix A for more details.
(6) The pseudo-R^2 for probit models is defined as $\frac{var(X'\gamma)}{1+var(X'\gamma)}$.

Table 4

OLS, Fixed Effect, and Probit Estimates of Catholic High School Effects by Race and Urban Residence. Full Set of Controls[1,2] (Huber-White Standard Errors in Parentheses) [Marginal Effects in Brackets[3]]

	Sample			
	(1)	(2)	(3)	(4)
	Urban and Suburban White Only	Urban and Suburban Minorities Only	Urban White Only	Urban Minorities Only
HS Graduate	(N=3799)	(N=1308)	(N=1002)	(N=697)
Sample Mean	0.88	0.80	0.88	0.80
Probit	0.443	0.524	1.176	1.592
	(0.279)	(0.338)	(0.417)	(0.673)
	[0.046]	[0.085]	[0.091]	[0.191]
College in 1994	(N=3695)	(N=1258)	(N=981)	(N=666)
Sample Mean	0.37	0.26	0.32	0.26
Probit	0.354	0.697	0.506	0.677
	(0.107)	(0.201)	(0.167)	(0.303)
	[0.087]	[0.158]	[0.110]	[0.144]
12th Grade Reading Score	(N=3638)	(N=1051)	(N=978)	(N=561)
Sample Mean	52.94	47.72	53.33	47.61
OLS	1.30	-0.72	1.59	-0.19
	(0.44)	(0.98)	(0.67)	(1.39)
12th Grade Math Score	(N=3638)	(N=1053)	(N=979)	(N=563)
Sample Mean	53.09	47.33	53.90	48.88
OLS	1.07	1.17	1.69	1.25
	(0.35)	(0.76)	(0.52)	(1.09)

Notes:

(1) All models include controls for hispanic origin, gender, region, citysize, distance to the nearest Catholic school (5 categories), family background, 8th grade tests, and other 8th grade measures. (from teacher, parent, and student surveys). See Table 3 notes 1 and 2.

(2) NELS:88 third follow-up and 2nd follow-up panel weights used for the educational attainment and 12th grade models, respectively.

(3) Marginal effects of probit models are computed as average derivatives of the probability of an outcome with respect to Catholic school attendance.

Table 5

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[1]. Modified Control Set[2].
(Huber-White Standard Errors in Parentheses) [Marginal Effects in Brackets]

	Correlation of Disturbances[3]					
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0.4$	$\rho = 0.5$
HS Graduation:						
Full Sample	0.459	0.271	0.074	-0.132	-0.349	-0.581
(Raw difference=0.12)	(0.150)	(0.150)	(0.150)	(0.148)	(0.145)	(0.140)
	[0.058]	[0.037]	[0.011]	[-0.021]	[-0.060]	[-0.109]
Catholic 8th Graders	1.036	0.869	0.697	0.520	0.335	0.142
(Raw difference=0.08)	(0.314)	(0.313)	(0.310)	(0.306)	(0.299)	(0.290)
	[0.078]	[0.064]	[0.050]	[0.038]	[0.025]	[0.011]
Urban	1.095	0.905	0.706	0.499	0.282	0.053
Minorities	(0.526)	(0.538)	(0.549)	(0.560)	(0.570)	(0.578)
(Raw difference=0.22)	[0.176]	[0.157]	[0.132]	[0.101]	[0.062]	[0.013]
College Attendance:						
Full Sample	0.331	0.157	-0.019	-0.196	-0.376	-0.558
(Raw difference=0.31)	(0.070)	(0.070)	(0.070)	(0.068)	(0.067)	(0.064)
	[0.084]	[0.039]	[-0.005]	[-0.047]	[-0.087]	[-0.125]
Catholic 8th Graders	0.505	0.336	0.165	-0.008	-0.184	-0.362
(Raw difference=0.23)	(0.121)	(0.120)	(0.119)	(0.117)	(0.114)	(0.110)
	[0.140]	[0.093]	[0.045]	[-0.002]	[-0.050]	[-0.099]
Urban	0.447	0.269	0.090	-0.091	-0.272	-0.455
Minorities	(0.282)	(0.282)	(0.280)	(0.276)	(0.269)	(0.259)
(Raw difference=0.30)	[0.116]	[0.062]	[0.020]	[-0.020]	[-0.057]	[-0.091]

Notes:
(3) Models estimated as bivariate probits with the correlation ρ between u and ε set to the values in column headings.
(1) NELS:88 3rd follow-up sampling weights used in the computations.
(2) Due 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", "parents contacted about behavior", and a limited-English proficiency index. Other than these exclusions, the controls are identical to those described in Table 3 notes 1 and 2.

Table 6

Sensitivity of Estimates of Catholic Schooling Effects on College Attendance and HS Graduation to Assumptions about Selection Bias in NELS:88, Catholic 8th Grade Subsample[1,2], Modified Control Set[3] (Huber-White Standard Errors in Parentheses) [Marginal Effects in Brackets]

Model:

$$CH = 1(X'\beta + u > 0)$$

$$Y = 1(X'\gamma + \alpha CH + \epsilon > 0)$$

Estimation Method 1: β, γ, and α estimated simultaneously as a constrained bivariate probit model:

Model	Constraint on ρ	HS Graduation Coefficients		College Attendance Coefficients	
		$\widehat{\rho}$	$\widehat{\alpha}$	$\widehat{\rho}$	$\widehat{\alpha}$
(1)	$\rho = \frac{cov(X\beta, X\gamma)}{var(X\gamma)}$	0.24	0.59	0.24	0.11
		(0.13)	(0.33)	(0.06)	(0.16)
			[0.05]		[0.03]
(2)	$\rho = 0$	0	1.04	0	0.51
			(0.31)		(0.12)
			[0.08]		[0.14]

Estimation Method 2: 2-step, with β obtained from a univariate probit, γ from a univariate probit using the public 8th grade subsample. Next, α is computed from a bivariate probit with β fixed at this initial value and γ fixed up to 6 proportionality factors.[4]

Model	Constraint on ρ	HS Graduation Coefficients		College Attendance Coefficients	
		$\widehat{\rho}$	$\widehat{\alpha}$	$\widehat{\rho}$	$\widehat{\alpha}$
(3)	$\rho = \frac{cov(X\beta, X\gamma)}{var(X\gamma)}$	0.09	0.94	0.27	0.06
		(0.08)	(0.30)	(0.05)	(0.10)
			[0.07]		[0.02]

Estimation Method 3: β, γ, and α estimated simultaneously as a constrained semiparametric model[5]:

$$CH = 1(X'\beta + \theta + u > 0)$$

$$Y = 1(X'\gamma + \alpha CH + \theta + \epsilon > 0)$$

Model	Constraint on ρ, where $\rho = \frac{var(\theta)}{1+var(\theta)}$	HS Graduation Coefficients		College Attendance Coefficients	
		$\widehat{\rho}$	$\widehat{\alpha}$	$\widehat{\rho}$	$\widehat{\alpha}$
(4)	$\rho = \frac{cov(X\beta, X\gamma)}{var(X\gamma)}$	0.25	0.80	0.25	0.15
		(0.16)	(0.37)	(0.09)	(0.22)
			[0.05]		[0.04]

Notes:

(1) Estimation performed on a sample of Catholic 8th grade attendees from NELS:88. N=859 for the HS graduation sample, and N=834 for the college attendance sample.

(2) NELS:88 3rd follow-up sampling weights used in the computations.

(3) Due to computational difficulties, several variables were excluded from the control sets in the bivariate probit models. See Table 5, note 2.

(4) 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 0.82 (0.19), 0.87 (0.22), 0.92 (0.03), 1.07 (0.04), 0.59 (0.08), and 0.90 (6.08) respectively, in the high school graduation case. For college attendance, the coefficients and (standard errors) are 0.80 (0.01) , 1.01 (0.04), 0.95 (0.15), 0.43 (0.17), 1.44 (0.03), and 1.04 (1.59).

(5) Models estimated as univariate probits conditional on θ, the distribution of which is estimated nonparametrically.

Table 7

The Amount of Selection on Unobservables Relative to Selection on Observables Required to Attribute the Entire Catholic School Effect to Selection Bias (Huber-White Standard Errors in Parentheses)

Model: $Y_i = 1(X_i'\gamma + \alpha CH_i + \epsilon_i > 0)$ for HS Graduation and College Attendance, estimated as a probit
$Y_i = X_i'\gamma + \alpha CH_i + \epsilon_i$ for 12th Grade test scores, estimated by OLS

$\widehat{\alpha}$ estimated from the Catholic 8th Grade Subsample, Full Set of Controls[2]

Outcome:	$\frac{\widehat{E}(X_i'\widehat{\gamma}\mid CH_i=1)-\widehat{E}(X_i'\widehat{\gamma}\mid CH_i=0)}{\widehat{Var}(X_i'\widehat{\gamma})}$ (1)	$\widehat{Var}\ (\widehat{\epsilon})$ (2)	$E(\epsilon_i \mid CH_i = 1)$ $-E(\epsilon_i \mid CH_i = 0)$ if Cond. 4 Holds (3)	$\frac{Cov(\epsilon_i,\widetilde{CH}_i)}{Var(\widetilde{CH}_i)}$ (4)	$\widehat{\alpha}$ (5)	Implied Ratio (6)
HS Graduation (N=859)	0.24	1.00	0.24	0.29	1.03 (0.31)	3.55
College Attendance (N=834)	0.39	1.00	0.39	0.47	0.67 (0.16)	1.43
12th Grade Reading (N=739)	0.091	36.00	3.28	3.94	0.33 (0.62)	0.08
12th Grade Math (N=739)	0.038	24.01	0.91	1.09	1.14 (0.46)	1.04
$\widehat{\alpha}$ estimated from the Urban Minority Subsample						
HS Graduation (N=698)	0.73	1.00	0.73	0.88	1.59 (0.67)	1.81
College Attendance (N=698)	0.58	1.00	0.58	0.69	0.68 (0.30)	0.99
12th Grade Reading (N=561)	0.090	30.58	2.76	3.31	-0.19 (1.39)	-0.06
12th Grade Math (N=561)	0.058	20.25	1.17	1.40	1.25 (1.09)	0.89

Notes:

(1) The $\widehat{\gamma}$ used to evaluate $\frac{\widehat{E}(X_i'\widehat{\gamma}\mid CH_i=1)-\widehat{E}(X_i'\widehat{\gamma}\mid CH_i=0)}{\widehat{Var}(X_i'\widehat{\gamma})}$ is estimated under the restriction $\alpha = 0$, using the Catholic eighth grade sample for the top panel and the urban minority sample for the bottom panel.

(2) See Table 3 notes 4 and 5 for a description of the controls. In the urban minority sample, the indicator "Black" is excluded.

(3) Condition 4 states that the standardized selection on unobservables is equal to the standardized selection on observables. i.e. $\frac{E(\epsilon_i\mid CH_i=1)-E(\epsilon_i\mid CH_i=0)}{Var(\epsilon_i)} = \frac{E(X_i'\gamma\mid CH_i=1)-E(X_i'\gamma\mid CH_i=0)}{Var(X_i'\gamma)}$.

(4) “Implied Ratio” in column 6 is the ratio of standardized selection on unobservables to observables under the hypothesis that there is no catholic school effect.

(3) NELS:88 third follow-up and 2nd follow-up panel weights used for the educational attainment and 12th grade models, respectively.