An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling[1]

Joseph Altonji
Yale University and NBER

Todd Elder
University of Illinois

Christopher Taber
Northwestern University and NBER

January 20, 2005

[1]This research was supported by the Institute for Policy Research, Northwestern University, the National Science Foundation under grant SBR 9512009, and the NIH-NICHD under grant R01 HD36480-03. We are grateful to Timothy Donohue for excellent research assistance and to Jeffrey Grogger, Derek Neal, and participants in seminars at Boston College, Michigan State University, Northwestern University, Purdue University, University College London, University of Missouri, University of Virginia, and University of Wisconsin for helpful comments. We are responsible for the remaining shortcomings of the paper.

Abstract

Several previous studies have relied on religious affiliation and the proximity to Catholic schools as exogenous sources of variation for identifying the effect of Catholic schooling on a wide variety of outcomes. Using three separate approaches, we examine the validity of these instrumental variables. We find that none of the candidate instruments is a useful source of identification of the Catholic school effect, at least in currently available data sets. We also investigate the role of exclusion restrictions versus nonlinearity as the source of identification in bivariate probit models. The analysis in the paper may be useful as a template for the assessment of instrumental variables strategies in other applications.

1 Introduction

The question of whether private schools, including Catholic schools, provide better education than public schools is at the center of the current national debate over the role of vouchers, charter schools, and other reforms that increase choice in education. All serious studies comparing Catholic and public schools acknowledge the problem of nonrandom selection into Catholic schools and most wrestle with it in one way or another.[1] In the absence of experimental data, the main option is to find a nonexperimental source of variation Z_i in Catholic school attendance that is exogenous with respect to the outcome under study. The problem is that most student background characteristics that influence schooling decisions, such as income, attitudes, and education of the parents, are likely to directly influence outcomes. Characteristics of private and public schools such as tuition levels, student body characteristics and school policies are also poor candidates for excluded instruments because they are likely to be related to the effectiveness of the schools.

Two influential papers provide potential instrumental variables. Evans and Schwab (1995) treat Catholic schooling as exogenous in much of their analysis, but also present estimates that rely in part on the assumption that religious affiliation affects whether a person attends a Catholic school but has no independent effect on the outcome under study. Specifically, they use a dummy variable for affiliation with the Catholic church (C_i) as their excluded variable. Some support for this assumption is evidenced by the fact that Catholics are significantly more likely than non-Catholics to attend Catholic school, while Catholics are not far from national averages on many socio-economic indicators. Evans and Schwab find a strong positive effect of Catholic school attendance on high school graduation and on the probability of starting college. However, as Murnane (1985), Tyler (1994), and Neal (1997) note and as Evans and Schwab acknowledge, being Catholic could well be correlated with characteristics of the neighborhood and family that influence the effectiveness of schools.[2]

[1] A few examples of early studies of Catholic schools and other private schools are Coleman et al (1982), Noell (1982), Goldberger and Cain (1982), Alexander and Pallas (1985), and Coleman and Hoffer (1987). Recent studies include Evans and Schwab (1993,1995), Tyler (1994), Neal (1997), Figlio and Stone (1998), Grogger and Neal (2000), Sander (2001), and Jepsen (forthcoming). Murnane (1984), Witte (1992), Chubb and Moe (1990), Cookson (1993), and Sander (2001) provide overviews of the discussion and references to the literature.

[2] Neal (1997) points out that one problem with using C_i as an instrumental variable when estimating Catholic school effects (as in Neal, 1997, and Evans and Schwab, 1995) is that religious identification might be influenced by the school type attended. Neither study investigates the issue. In the case of NELS:88 we use the parent's report of religious affiliation while the student is in eighth grade as our religion measure, but our results are not very sensitive to using the child's tenth grade report instead.

Neal (1997) uses proxies for geographic proximity to Catholic schools and subsidies for Catholic schools as exogenous sources of variation in Catholic high school attendance (see also Tyler, 1994). The basic assumption is that the location of Catholics or Catholic schools was determined by historical circumstances unrelated to unobservables that influence performance in schools. Using data from the National Longitudinal Survey of Youth (NLSY), Neal estimates bivariate probit models of Catholic high school attendance and high school graduation in which Catholic school effects are identified by excluding whether the person is Catholic and either the fraction of Catholics in the county population in the case of urban minorities or the number of Catholic schools per square mile in the county in the case of urban whites (Table 6, p.113). His point estimates are not very sensitive to adding Catholic to the outcome equation.

The interaction between whether a person is Catholic and the availability of Catholic schools is a natural alternative to using proximity or religion separately. It is possible that distance to Catholic schools is related to differences in regional and family characteristics that have a direct influence on schooling and labor market outcomes, given that Catholic schools are somewhat concentrated by region.[3] However, since "tastes" for Catholic schooling depend strongly on religious preference, the interaction between distance (D_i) and religious affiliation will have an effect on Catholic school attendance that is independent of the separate effects of religious affiliation and distance. In particular, Catholic school attendance is likely to be much more sensitive to distance for Catholics than for non-Catholics. Consequently, one can control for both religious affiliation and for distance from Catholic schools, as well as for a set of other geographic characteristics, while excluding the interaction $C_i \times D_i$ from outcome models. However, the case that $C_i \times D_i$ may be a valid instrument even if C_i and D_i are not is far from bulletproof. Catholic parents who want their children to attend Catholic schools might choose to live near Catholic schools. This could lead to a positive or negative bias depending on the relationship between preferences for Catholic school and the error component in the outcome equation. Also, past immigration patterns and internal migration from city to suburb and across regions may have led to differences between Catholics and non-Catholics in the correlation between proximity to Catholic schools and observed and unobserved components of family background.

In this paper we explore the validity of Catholic religion, proximity to Catholic schools,

[3] Hoxby (1995) discusses geographical concentration by region, much of which is associated with the geographic concentration of the Catholic population in the past.

and the interaction between religion and proximity as exogenous sources of variation for identifying the effects of Catholic schooling on educational attainment and achievement. Our analysis may be of broader interest as a prototype for evaluations of instrumental variables strategies that could be conducted in other domains. Given that our conclusions are negative on the utility of the instruments, it is important to stress at the outset that the authors cited above recognize the potential problems with the exogeneity and power of the instruments that we investigate here. In particular, Grogger and Neal (2000, p. 191) reach the conclusion that bivariate probit strategies are not very informative in the context of one of our main data sets, the National Educational Longitudinal Survey of 1988 (NELS:88).

We use multiple data sets and methods in our evaluation. In addition to the National Educational Longitudinal Survey of 1988 (NELS:88), we also report results based on the National Longitudinal Study of the High School Class of 1972 (NLS-72). For each instrument, we present 2SLS and bivariate probit estimates that rely on the particular instrument as the source of identification and compare the results to OLS and univariate probit estimates.

In addition to examining the a priori case for the instruments and the face plausibility and precision of the IV estimates, we assess the quality of the instruments in two other ways. The first takes advantage of the fact that few students who attend public 8th grades attend Catholic high school regardless of religion. This provides some justification for using the coefficient on the instrument in a reduced-form outcome equation from a sample of public eighth grade attendees in NELS:88 as an estimate of the direct link between Catholic religion and the outcome. The second approach uses a methodology introduced in Altonji, Elder and Taber (2002; hereafter, AET) to assess the instrumental variable results. AET's approach is based on the practice of using the degree of selection on observables as a guide to how much selection there is on unobservables.[4]

Finally, we investigate differences in the point estimates and standard errors obtained using 2SLS versus bivariate probits and show that, at least in our data, religion and non-linearities in the effects of religion and family background rather than the location variable instruments are the main source of identification when we use Neal's measures of proximity to Catholic schools.

[4]Researchers often informally argue for the exogeneity of membership in a "treatment group" or of an instrumental variable by examining the relationship between group membership or the instrumental variable and a set of observed characteristics, or by assessing whether point estimates are sensitive to the inclusion of additional control variables. See for example, Currie and Duncan (1995), Engen et al (1996), Angrist and Evans (1998), Jacobsen et al. (1999), Bronars and Grogger (1994), and Udry (1998).

In Section 2 we discuss the data from NLS-72 and NELS:88 that are used in the study. In Section 3 we present results using religion as the source of identification and provide some initial evidence on the direct effect of being Catholic on educational attainment. We also introduce and apply AET's method of using the observables to assess the potential for bias from an association between the instrument and the unobservables. In Sections 4 and 5 we present results using distance and the interaction between distance and religion as the excluded instruments. In Section 6 we investigate the role of exclusion restrictions versus nonlinearity as the source of identification in bivariate probit models. Section 7 concludes with a research agenda.

2 Data

2.1 NELS:88

NELS:88 is a National Center for Education Statistics (NCES) survey which began in the Spring of 1988. A total of 1032 schools contributed as many as 26 eighth grade students to the base year survey, resulting in 24,599 eighth graders participating. Subsamples of these individuals were reinterviewed in 1990, 1992, and 1994. The NCES only attempted to contact 20,062 base-year respondents in the first and second follow-ups, and only 14,041 in the 1994 survey. Additional observations are lost due to attrition.

Parent, student, and teacher surveys in the base year provide information on family and individual background and on pre-high school achievement and behavior. Each student was also administered a series of cognitive tests in the 1988, 1990, and 1992 surveys to ascertain aptitude and achievement in math, science, reading, and history. We use standardized item response theory (IRT) test scores that account for the fact that the difficulty of the 10th and 12th grade tests taken by a student depends on the 8th grade scores. We use the 8th grade test scores as control variables and the 10th and 12th grade reading and math tests as outcome measures.

We calculate distance from the nearest Catholic high school as the distance from the zip code centroid of the respondent's eighth grade school to the zip code centroid of the closest Catholic high school.[5] Our distance measure D_i is a vector of mutually exclusive indicators

[5]Detailed information on zip code characteristics of the eighth grade school (at the zip code level) is available on the NELS:88 Restricted Use files. For the NELS:88 analysis, the zip code of every Catholic high school in the United States in 1988 was obtained from Ganley's *Catholic High Schools in America: 1988*. The distance from a particular zip code centroid to the centroids of all the catholic high schools was

for distance less than 1 mile, 1 to 3 miles, 3 to 6 miles, 6 to 12 miles, and 12 to 20 miles, with greater than 20 miles treated as the omitted category. Our religion indicator C_i is 1 if parents indicated that they are Catholic in response to a question about religious affiliation in the base year survey and is 0 otherwise.

Our main outcome measures are high school graduation (HS_i) and college attendance ($COLL_i$). HS_i is one if the respondent graduated high school by the date of the 1994 survey, and zero otherwise.[6] $COLL_i$ is one if the respondent was enrolled in a four-year university at the date of the 1994 survey and zero otherwise.[7] The indicator variable for Catholic high school attendance, CH_i, equals 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.[8] Unless noted otherwise, the results reported in the paper are weighted.[9]

2.2 NLS-72

The NLS-72 is a Department of Education survey of high school students that contains information on 22,652 persons who were seniors during the 1971-1972 academic year. Additional interviews were conducted in 1973, 1974, 1976, 1979, and 1986. The final sample sizes are 19,489 students from 1192 public high schools and 71 Catholic high schools for the college attendance indicator variable, 14,671 students from 879 public high schools and 57 Catholic high schools for the math and reading score variables, and 16,276 students from 1191 public high schools and 71 Catholic high schools for the years of academic education variable.[10]

---

calculated using an algorithm obtained from the U.S. National Oceanic and Atmospheric Administration.

[6] We obtain similar results using a "drop out" dummy variable which equals one if a student dropped out of high school by 1992, or if the student dropped out of high school by 1990 and was not reinterviewed in 1992 or 1994, zero otherwise. This variable catches dropouts who left the survey by 1990 and were either dropped from the sample or were nonrespondents.

[7] 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.

[8] A student 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 ($CH = 0$). If such transfers are frequently motivated by discipline problems, poor performance, or alienation from school, then misclassification of the transfers as public high school students could lead to upward bias in estimates of the effect of CH on educational attainment. AET present evidence that this issue is of minor importance.

[9] The sampling scheme in the NELS:88 is explained in AET and Grogger and Neal (2000). The results are somewhat sensitive to the use of sample weights, although our main findings are robust to weighting. Given the choice based sampling scheme the weighted estimates are clearly preferred.

[10] The 2236 students who did not report their religious affiliation are excluded from the analysis. We also drop an additional 495 students for whom we could not impute distance from the nearest Catholic high school, reducing the sample size to 19,921. We also exclude 111 cases in which the student attended a non-Catholic private school, and additional observations are lost because data for key control variables and outcomes are missing.

The variable C_i is 1 for students who indicated they were Catholic in response to a base year question about religious affiliation and is 0 otherwise. Distance from the nearest Catholic high school was recorded as the distance in air miles between the centroids of the zip code of residence reported in the first follow-up, and the zip code of the nearest Catholic high school, and our distance measure D_i is defined as a set of indicator variables identical to those used in the NELS:88 data.[11] The follow-up survey included an indicator for whether the respondent had moved between their senior year of high school and the survey date, so the 10,530 students who moved were assigned the mean value of distance for all non-movers who attended the same high school.[12]

In the original design, schools with a high percentage of minority students and in low income areas are overrepresented, and sampling weights also vary with whether the school is public or private. The results are not sensitive to weighting procedures, and so we report unweighted estimates.

3 Using Religious Affiliation to Identify the Catholic School Effect

In Table 1 we present univariate probit, OLS, bivariate probit, and 2SLS estimates of the Catholic school effect. The table footnotes provide a list of the family background, city size, student characteristic and eighth-grade behavioral and academic outcome variables that are included in both the equations for CH_i and the outcomes (Y_i). In this section our focus is on the first column in which we use C_i as the excluded instrument and include D_i but not $C_i \times D_i$ in the equations for both CH_i and Y_i. Before proceeding further it worth emphasizing that throughout the paper, we focus on homogenous treatment effects models, although in Section 6 we estimate separate models for urban whites and urban minorities. In reality, the gains from Catholic school attendance probably vary with the quality of the local Catholic school and public school and the match between the student and the school. If one regards the estimates as local average treatment effects in the spirit of Imbens and Angrist (1994),

[11] The zip code of every Catholic high school in existence in the United States is listed in the US Department of Education's "Universe of Private Schools."

[12] The 495 students who were dropped because no distance measures could be created for them either attended one of the 26 high schools for which there are no valid observations on distance, or did not have valid values for the geographic move variable. These schools were part of NLS-72's "backup sample," and the students in this subsample were lost because they were excluded from the first follow-up.

one would expect some variation in estimates to arise because of this.

In NELS:88 the 2SLS estimate of the effect of CH_i on high school graduation is 0.34 (0.08). This estimate is unreasonably large given that the sample mean of HS_i is 0.84. The bivariate probit estimate of the average marginal effect is a more reasonable value of 0.128, but it is still double the univariate probit estimate. The estimates of the effect of CH_i on enrollment in a four-year college in 1994 are also unreasonably large, as the 2SLS coefficient of 0.40 (0.10) is larger than the sample mean of 0.29. The bivariate probit estimate of 0.170 is also well above the univariate probit estimate of 0.094.

We obtain a different pattern in NLS-72 (bottom panel). On one hand, the probit estimate of the effect of CH_i on college attendance is 0.068, which is smaller than but reasonably close to the NELS:88 estimate of 0.094. However, in NLS-72 the use of 2SLS does not lead to a big increase in the estimate. (The point estimate is 0.06, although the apparent similarity should be interpreted cautiously, as the 2SLS standard error is 0.04 and the 2SLS estimate is not significantly different from zero even though it implies a large Catholic schooling effect.) The NELS:88 results change very little when we condition the analysis on making it to 12th grade or on $HS_i = 1$, so we cannot attribute the similarity of the results from 2SLS and single-equation methods in NLS-72 but not NELS:88 to the fact that NLS-72 is limited to those who have made it to 12th grade. We suspect that part of the difference in results for the two data sets is due to improvements over time in the relative social position of the Catholic population with school age children in the U.S.. The larger gap between the observed characteristics of Catholics and non-Catholics in NELS:88 relative to NLS-72 (Tables 3a and 3b) is consistent with this, as we discuss below. However, we do not have a full explanation.

The NLS-72 bivariate probit estimate is only -0.002, but it should be kept in mind that the source of identification in the bivariate probit case is a complicated nonlinear function of the variables in the model for CH_i and not simply C_i, even though only C_i is excluded from the outcome equation. In particular, the analysis in Section 6 below suggest that the interaction between C_i and D_i plays an important role in bivariate probit and leads the bivariate probit point estimate to be smaller than the 2SLS estimate. Our analysis suggests that identification of the bivariate probit comes primarily from the functional form assumptions rather than the exclusion restrictions in some cases. Thus to assess the credibility of the instruments, we focus on the 2SLS results when thinking about the validity of the exclusion restrictions for particular instruments.

Table 2 reports OLS and 2SLS estimates of the effect of Catholic high school on test scores in NELS:88 and a variety of outcomes in NLS-72. Column (1) shows that the 2SLS estimates are larger for both NELS test scores than the single-equation ones, although the 2SLS coefficients are noisy. The standard deviation of these tests is 10, so the 2SLS estimate of 2.64 implies a large impact on 12th grade math scores. However, the fact that the OLS estimates are uniformly smaller indicates that either 2SLS is biased upward or that Catholic high school students are actually negatively selected on the basis of unmeasured factors which are correlated with test scores. The NLS-72 test score results follow the opposite pattern – 2SLS estimates are negative while OLS is large and positive for both reading and math. Although the NLS-72 analysis does not control for eighth grade achievement, this does not account for the differences in patterns between the two data sets, as (unreported) NELS:88 models that do not control for 8th grade achievement generate similar results to the NELS:88 models in Table 2.

To summarize, in NELS:88 the 2SLS estimates using C_i as the excluded instrument imply that the Catholic school effect is very large, particularly for educational attainment. The NLS-72 results are more mixed but are consistent with a substantial positive effect on educational attainment. One might be tempted to conclude that IV estimates, while unreasonably large, bolster the probit and OLS evidence that the true effect is substantial. In the remainder of this section, we argue that this is the wrong interpretation.

3.1 Comparing Catholics and non-Catholics

Column (1) of Table 3a presents sample means of a set of family background characteristics, student characteristics, eighth grade outcomes, and high school outcomes in NELS:88, and Column (2) shows the difference between Catholics and non-Catholics in these means.[13] Catholics are 7 percentage points more likely to graduate high school and 8 percentage points more likely to be enrolled in a four year college in 1994. Differences in tenth and twelfth grade test scores are more modest but all show a significant advantage for Catholic students. If Catholic was as good as randomly assigned, these differences would be entirely attributed to the fact that Catholics are more likely to attend Catholic high school. It would then be

[13] In Table 3a the outcome variables are weighted with the same weights used in the regression analysis, so that the 10th and 12th grade test scores are weighted using first and second follow-up panel weights, respectively, and high school graduation and college attendance are weighted by third follow-up weights. All other variables are weighted using second follow-up panel weights.

troubling if Catholic appeared to be related to a broad set of variables determined prior to high school enrollment that influence high school outcomes. Consequently, we begin our evaluation of Catholic religion as an excluded instrument by following the common practice of simply comparing the characteristics of Catholics and non-Catholics in both NELS:88 and NLS-72.

Differences by C_i appear in many of the family and student characteristics and eighth grade outcomes in Table 3a. There is a modest positive association between Catholic religion and parental educational expectations, with a gap of 0.04 in the fraction of parents who expect their children to attend some college and 0.03 in the fraction who expect at least a college degree.[14] While the differential in family income is positive, it is negative in mother's and father's education. However, Table 3a also shows that Catholic students are favored across a broad set of measures available in eighth grade, such as test scores, grades, and teacher evaluations of the student's behavior. Among these eighth grade variables, only the "unpreparedness index" variable does not vary favorably with C_i. The discrepancy in the fraction of students who repeated a grade in grades 4-8 is -0.03, and the gap in the fraction of students who are frequently disruptive is -0.02. The existence of gaps in favor of Catholic students across several dimensions suggests that Catholic and non-Catholic students differ in many respects, some of which may be unobservable to empirical researchers. Since these differences also contribute to high school and post-high school outcomes (see AET for evidence), doubts arise regarding the validity of using C_i as an instrumental variable for Catholic high school attendance.

In NLS-72, the differences are less pronounced, although it appears that overall Catholic religion has a weak positive association with favorable family background characteristics. Log family income is 0.07 higher for Catholics, who are also five percentage points less likely to be members of families which meet NLS-72's definition of low socio-economic status. There are also essentially no differences in parental education levels or pre-high school student educational expectations.

Given the overall picture of Tables 3a and 3b, we anticipate that the use of C_i as an instrumental variable will likely result in positively biased estimates of Catholic schooling

[14] Some of the variables used in our multivariate models are excluded from Table 3a to keep them manageable given sample sizes. The expectations variables in Table 3a are excluded from our outcome models because if Catholic school has an effect on outcomes, this may be influence expectations. The association between C_i and the 8th grade varaibles and expectations variables remains substantial even after controlling for demographics, family background and geography.

effects in NELS:88, and perhaps a small positive bias in NLS-72, although it is difficult to gauge the extent of the bias. The richness of the NELS:88 data permits us to use two more formal procedures to gauge its magnitude and direction.

3.2 The Effect of Catholic Religion for Students from Public Eighth Grades

One way to assess the endogeneity of Catholic religion is to identify a sample of persons for whom Catholic high school is not a serious option, and then interpret the coefficient on C_i in a single equation model as an estimate of the direct effect of Catholic religion on the outcome. Public eighth graders provide such a sample, because only 0.3% of public school eighth graders in our effective sample go on to attend Catholic high school. The corresponding percentage among public eighth grade attendees whose parents are Catholic is only 0.7%.

Let the outcome Y_i be determined by

$$Y_i = \alpha CH_i + X_i'\gamma + \varepsilon_i, \tag{1}$$

where γ is defined so that $\text{cov}(\varepsilon_i, X_i) = 0$. CH_i is potentially endogenous and thus correlated with ε_i. We assume that our instrument C_i does not influence Y_i directly but is correlated with CH_i. However, there is concern that C_i is correlated with ε_i.

Define β, π, and λ to be the coefficients of the least squares projections

$$\text{Proj}\,(C_i \mid X_i) = X_i'\pi, \tag{2}$$

$$\text{Proj}\,(CH_i \mid X_i, C_i) = X_i'\beta + \lambda C_i. \tag{3}$$

Define $\widetilde{C}_i$ as the residual of the projection of C_i on X_i so that

$$\widetilde{C}_i \equiv C_i - X_i'\pi. \tag{4}$$

It is well known that the IV estimate of α converges to

$$\widehat{\alpha}_{IV} \xrightarrow{p} \alpha + \frac{cov\left(\widetilde{C}_i, \varepsilon_i\right)}{\lambda var(\widetilde{C}_i)}. \tag{5}$$

Now suppose there is an event p_i for which $\Pr(CH_i = 1 \mid p_i) = 0$. In our application p_i is attendance of a public eighth grade by individual i. We assume for now that the joint distribution of $(X_i, C_i, \varepsilon_i)$ is independent of p_i, but argue at the end of the section that accounting for correlation between C_i and ε_i induced by restricting the analysis to the public eighth-grade sample is likely to strengthen the evidence against C_i as an instrument

Consider a regression of Y_i on X_i and C_i conditional on p_i. Under these conditions, the coefficient on C_i will converge to $cov\left(\widetilde{C}_i, \varepsilon_i\right)/var(\widetilde{C}_i)$. Since we have a consistent estimate of λ from the first stage regression, we can obtain a consistent estimate of the bias $\psi = cov\left(\widetilde{C}_i, \varepsilon_i\right)/\left(\lambda var(\widetilde{C}_i)\right)$ by estimating the parameter ψ in the regression model

$$Y_i = X_i'\Lambda + [C_i\widehat{\lambda}]\psi + \omega_i \quad (6)$$

on the public eighth grade sample.

In column (1) of Table 4 we report estimates of the bias parameter ψ using this approach to evaluate Catholic religion as an instrument.[15] We present separate equations estimated for HS_i, $COLL_i$, and the 12th grade math and reading test scores. The vector X_i includes all of the other controls that were included in our models in Tables 1 and 2. For ease of comparison, the table also presents the corresponding 2SLS estimates from Table 1 and 2.

The results are striking – the implied bias in the 2SLS estimate is 0.34 (0.08) for HS_i, which is identical to the 2SLS coefficient itself.[16] The large potential bias should raise a great deal of concern about using Catholic as an instrument, particularly given the remarkable similarity between the magnitudes of the bias and the 2SLS estimate. In our view, this evidence alone is sufficient to rule out Catholic religion as a useful instrument.

In the college attendance case the (unreported) estimate of $cov\left(\widetilde{C}_i, \varepsilon_i\right)/var(\widetilde{C}_i)$ is 0.038 (0.013). Catholic students are nearly four percentage points more likely to enroll in a four year college than non-Catholics even when Catholic high school is not a serious option. This relationship implies a bias of 0.29 (0.11) in 2SLS estimates, so it seems likely that the large 2SLS estimates in Table 1 result from the endogeneity of C_i with respect to both high

[15] Eliminating the 36 students who attended public 8th grade and went on to Catholic high school makes little difference.

[16] To see how we arrive at this figure, note that the estimate of $cov\left(\widetilde{C}_i, \varepsilon_i\right)/var(\widetilde{C}_i)$ in the HS equation is 0.044 (0.011). That is, the graduation probability among students who go to public eighth grade is estimated to be 0.044 higher for Catholics than non-Catholics, even though hardly any of these students attend Catholic high schools. Since λ is estimated to be 0.130 (0.009), the bias is approximately 0.34 (= 0.044 / 0.130).

school graduation and college attendance. Similar calculations imply that the math test score estimate from Table 2 can largely be explained by potential bias of 1.85 (1.41) for the 12th grade math scores. Part of the college attendance and test score effects may be "real," as these large corrections are still smaller than the 2SLS point estimates, but the substantial evidence of endogeneity of C_i combined with the imprecision of the estimates prevents any firm conclusions about the effect of Catholic high school on these outcomes.

We now address the selection problem induced by focusing only on public eighth graders. The analysis in this section has treated public eighth grade attendance as if it were randomly assigned. We would expect positive selection of Catholics into Catholic grade schools because Catholic schooling requires tuition and parental initiative, and also because we observe positive selection on a broad list of characteristics that improve school outcomes. That is, Catholic students who attend Catholic grade schools are likely to have higher values of ε_i in equation (1) than Catholic public school students. Since non-Catholics are much less likely to attend Catholic schools this effect will lead to a negative bias in $Cov(\widetilde{C}_i, \varepsilon_i)$ when we condition on public school attendance.[17] This would imply that our estimates of ψ are biased downward, which makes the results in this section even more surprising. While we believe that the above scenario is the most likely one, one could write down a Roy (1951) model of comparative advantage in which children who gain less from attending a Catholic 8th grade school conditional on the observables are more likely to attend a public 8th grade school, or a model in which parents substitute (unobserved) parental inputs for Catholic schooling when Catholic school is relatively expensive.[18] In either case, Catholic children may outperform non-Catholic children conditional on public eighth grade attendance even when C_i is a valid instrument.

3.3 Using the Observables to Assess the Bias from Unobservables

In this section we extend the methodology of AET to assess the potential bias in the instrumental variables estimates in equation (1). AET consider the case in which an instrument (such as C_i) is not necessarily valid, and the researcher does not have a strong prior about how it is determined. In particular, rather than assume that the choice of X_i ensures that

[17] To see this in a simple case, abstract from observables so that $\widetilde{C}_i = C_i$, and assume that non-Catholics do not attend Catholic schools, that $E(\varepsilon_i \mid C_i) = 0$ unconditional on p_i, and that there is positive selection into Catholic eighth grades so that $E(\varepsilon_i \mid C_i = 1, p_i^c) > E(\varepsilon_i \mid C_i = 1, p_i)$, where p_i^c is the complement of p_i. This implies that $E(\varepsilon_i \mid C_i = 1, p_i) < 0$ and thus the bias is negative.

[18] See Willis and Rosen (1979) as an example of an empirical model of this type.

$\widetilde{C}_i$ is uncorrelated with ε_i, as is required for consistency of 2SLS, AET develop a model of data collection which implies that the effect on C_i of a unit change in the index of observables $X_i'\gamma$ that determine Y_i is the same as the effect on C_i of a unit change in the index of unobservables ε_i. When the instrument is an indicator variable such as C_i, the condition may be written as

$$(7) \qquad \frac{E(\varepsilon_i \mid C_i = 1) - E(\varepsilon_i \mid C_i = 0)}{Var(\varepsilon_i)} = \frac{E(X_i'\gamma \mid C_i = 1) - E(X_i'\gamma \mid C_i = 0)}{Var(X_i'\gamma)}.$$

The term $\frac{E(X_i'\gamma|C_i=1)-E(X_i'\gamma|C_i=0)}{Var(X_i'\gamma)}$ is the normalized shift in the index of observables in the outcome equation that is associated with C_i, while the term $\frac{E(\varepsilon_i|C_i=1)-E(\varepsilon_i|C_i=0)}{Var(\varepsilon_i)}$ is the corresponding normalized shift in the distribution of unobservables. Using $\frac{E(X_i'\gamma|C_i=1)-E(X_i'\gamma|C_i=0)}{Var(X_i'\gamma)}$ to assess the possibility that $\frac{E(\varepsilon_i|C_i=1)-E(\varepsilon_i|C_i=0)}{Var(\varepsilon_i)}$ is substantially different from 0 is a formalization of the common practice of checking for a systematic relationship between an instrumental variable and each of the elements of X_i, as we performed in Section 3.1 above. Intuitively, if one estimates $\frac{E(X_i'\gamma|C_i=1)-E(X_i'\gamma|C_i=0)}{Var(X_i'\gamma)}$ and finds that it is substantially different from zero, one may be worried that the null hypothesis $E(\varepsilon_i \mid C_i) = 0$ is wrong. The precise conditions that imply equation (7) are given in AET.

We can use (7) to approximate the amount of bias in 2SLS estimates of Catholic schooling effects if selection on unobservables is similar to selection on observables. It is straightforward to show that the asymptotic bias from 2SLS would be

$$(8) \qquad \frac{cov(\lambda\widetilde{C}_i, \varepsilon_i)}{var\left(\lambda\widetilde{C}_i\right)} = \frac{var\,(C_i)}{\lambda var\left(\widetilde{C}_i\right)}\left[E(\varepsilon_i \mid C_i = 1) - E(\varepsilon_i \mid C_i = 0)\right]$$

$$(9) \qquad = \frac{var\,(C_i)}{\lambda var\left(\widetilde{C}_i\right)}\frac{Var(\varepsilon_i)}{Var(X_i'\gamma)}\left[E(X_i'\gamma \mid C_i = 1) - E(X_i'\gamma \mid C_i = 0)\right].$$

where we have used (7) to obtain (9) from (8). The hypothesis of equal selection on observables and unobservables provides a way of identifying $[E(\varepsilon_i \mid C_i = 1) - E(\varepsilon_i \mid C_i = 0)]$, and therefore the asymptotic bias of instrumental variable estimates, since the terms in (9) are readily and consistently estimable. AET develops extensions to the case of latent dependent variables, so both probit and linear 2SLS bias calculations are given where appropriate.

One should not make too much of the specific estimates of bias, which are based on strong assumptions about the symmetry of selection of observables and unobservables. In AET, we argue that the relationship between the indices of unobservables that determine CH_i and Y_i is likely to be weaker than the relationship between the indexes of observables, in part because many of the factors that determine graduation and college attendance are determined after 8th grade and are excluded from X_i by design. We are less clear about the force of this argument in the case of C_i and the other instruments we consider. The variables C_i, D_i, and $C_i \times D_i$ could all be correlated with pre and post 8th grade influences on Y_i that are not correlated with CH_i, but these correlations could be stronger or weaker than the link between factors that determine CH_i and Y_i. However, we suspect that they are considerably weaker, which means that bias estimates will be too large in absolute value.

One may refine the bias calculations to account for the fact that the variation in the instrument may only be over a specific dimension. For example, D_i only varies across zip code, and so must be orthogonal to within zip code variation in $X_i'\gamma$ and ε_i. Consequently, we adjust the bias estimates by using variance in $E(X_i'\gamma)$ across zip codes relative to the variance within zip codes as a guide to the variance in $E(\varepsilon_i \mid D_i)$ relative to the cross area variance in ε_i.

Column (1) of Table 5 presents the results, which are quite striking. In the case of high school graduation, for linear 2SLS we calculate a bias of 0.52 (0.23) in $\widehat{\alpha}$ if we include D_i among the set of variables used to form the index of observables and 0.84 (0.26) if we exclude it. These are both huge potential biases, greater in magnitude than the implausibly large 2SLS point estimate, which is repeated in this table for convenience. In the case of $COLL$ the bias estimate under the assumptions leading to (7) is 0.45 (0.21), which is slightly larger than the 2SLS estimate of 0.40. If selection on unobservables follows the same pattern as selection on observables, there is a huge bias in the IV estimates when C_i is used as an instrument, at least for the cohort of children sampled in NELS:88.[19] The results reinforce our conclusions based on the public 8th grade sample. However, the bias estimates have large standard errors and are best interpreted as a sign of potential trouble rather than a precise estimate of the extent of the bias.

The bottom panels of Table 5 repeat the calculations for 12th grade test scores. These calculations use estimates of the reliability of the NELS:88 tests to provide a rough adjustment for the fact that much of the variance in ε_i is due to noise in the tests and thus is unrelated to C_i.[20] The calculations suggest that there is the potential for substantial bias

[19] This conclusion is also supported by calculations not reported that use a two stage probit procedure. See Elder (2002) for details.

[20] The adjustment is performed by multiplying the estimate of $\text{plim}(\hat{\alpha} - \alpha)$ based on (8) by (*reliability*-R^2)$/(1 - R^2)$, where *reliability* is the estimate of the reliability of the particular test, and R^2 is the R^2 of the model for the particular test. To see the justification, let the composite error term be $\varepsilon^* = \varepsilon + \varsigma$ where ς

when using C_i as an instrument, but the estimates are very imprecise. In the case of math the bias estimates of 2.02 (0.75) and 1.87 (0.74) (depending again on whether D_i is used in the calculations) preclude any firm conclusions. In general, we cannot rule out the possibility of a positive effect of Catholic high school attendance on achievement test scores, but the large potential biases are suggestive that the use of C_i as an instrument is not a reliable way to assess the magnitude of these effects.

3.4 Summary of C_i results

All three approaches that we have used cause one to question the use of C_i as an instrument. It seems closely related to observable covariates, which causes one to be worried that it may lead to bias. We estimate the bias in two very different ways, both of which suggest that the estimates may be substantially positively biased. We conclude from these calculations that IV procedures based on C_i lead to huge point estimates but may also be subject to a great deal of bias. In this circumstance, C_i is not a useful instrumental variable despite its powerful association with CH_i.

We have already noted that we do not have a good understanding of why the gap between the IV estimates of the Catholic school effect and the probit or linear probability estimates are so much larger in NELS:88 than in NLS-72 or in High School and Beyond (See Evans and Schwab, 1995). Unfortunately, we lack the rich set of primary school data required to use the relative degree of selection on observables to explore the discrepancy in IV results across data sets. The variability across data sets, which in part may reflect changes over time in the composition of the Catholic population in the U.S., is an additional reason to be cautious about the use of C_i as an instrument.

---

is the component of test scores due to noise in the test. One minus the reliability of the test is an estimate of $var(\varsigma)/var(Y_i+\varsigma)$ where Y_i is the true test score. The value 1 minus the R^2 of the test score model is an estimate of $[var(\varepsilon)+var(\varsigma)]/var(Y_i+\varsigma)$, and note that since $var(\varepsilon)=[var(\varepsilon)/(var(\varepsilon)+var(\varsigma))]var(\varepsilon^*)$,

$$var(\varepsilon)=\frac{(1-R^2)-(1-reliability)}{1-R^2}var(\varepsilon^*).$$

The R^2 is 0.60 for 12th grade reading and 0.74 for 12th grade math (using the 2SLS estimate of the model and ignoring the correlation between CH_i and ε_i), and the reliability is 0.85 for 12th grade reading and 0.94 for 12th grade math. Consequently, the correction scales down the bias estimates by 0.625 for reading and 0.770 for math.

4 Instrumental Variables Estimates using Proximity to Catholic Schools

In this section we evaluate proximity (D_i) as a source of identifying variation. In Column (2) of Table 1 we report estimates with D_i as the excluded instrument, and again we focus on linear 2SLS because of concerns that functional form assumptions are driving identification in bivariate probit models. The 2SLS estimate of -0.04 (0.10) for high school graduation is too imprecise for us to draw any inferences from it. The 2SLS estimates for $COLL_i$ 0.31 (0.11) are in NELS:88 and 0.44 (0.20) in NLS-72. Both estimates are much larger than the estimated marginal effect of 0.085 from the univariate probit in NELS:88 and 0.070 from NLS-72. Column (2) of Table 2 presents the results for test scores in NELS:88 and NLS-72. These coefficients vary across specifications, but for the NLS-72 test scores they imply very large effects. On their face, the findings for $COLL$ and the NLS-72 test score results appear implausible. In the remainder of this section we look for evidence of bias.

In Column (3) of Table 3a we report the relationship between a wide set of observables in NELS:88 and a student's distance from the nearest Catholic high school. For simplicity we collapsed the vector D_i into a dummy variable $D6_i$, which is equal to 1 for person i if she lives less than 6 miles from the nearest Catholic high school and zero otherwise, and present the difference in these means by $D6_i$. Among the eighth grade measures, such as teacher evaluations of the student's behavior, there is little difference between those who live close to Catholic high schools and those who do not. However, there is a positive relationship between $D6_i$ and most of the family background measures. There is also a positive association between proximity and both student and parental educational expectations. Similar differences by $D6_i$ appear in NLS-72 (Table 3b). These differences in family motivation and students' home environment introduce the possibility that there might also be unmeasured differences which could affect outcomes and lead to bias in models using D_i as an instrumental variable in both NLS-72 and NELS:88.

In column (2) of Table 4 we report estimates of the bias coefficient ψ based on the equation

$$Y_i = X_i'\gamma + [D_i'\widehat{\lambda}]\psi + \omega_i \tag{10}$$

for public eighth graders from NELS:88. In (10), $D_i'\widehat{\lambda}$ is the index of distance dummies weighted by their coefficients $\widehat{\lambda}$ in the first stage equation for CH_i. The estimate of ψ is

-0.05 (0.12) in the equation for HS_i and 0.37 (0.12) in the equation for $COLL_i$. There is not much evidence for bias in the HS_i equation given the large standard error, but this is not surprising given that the 2SLS estimate is also noisy and does not indicate a positive effect. For $COLL_i$, the implied bias is slightly larger than the 2SLS estimate, reaffirming the notion that one should not put too much stock in inferences using D_i as an instrument for college attendance, at least in NELS:88. In the case of reading scores the bias check is uninformative given the large standard error on ψ. For 12th grade math scores, the evidence in favor of a positive effect of CH_i is dampened by the fact that implied bias estimates are large in this case as well.

Finally, we apply the AET methodology for assessing the potential bias due to selection on unobservables. The extension of the methods to account for fact that D_i is a vector is straightforward, with the relevant condition analogous to (7) being $\frac{cov(D_i'\lambda,\varepsilon_i)}{var(\varepsilon_i)} = \frac{cov(D_i'\lambda,X_i'\gamma)}{var\left(X_i'\gamma\right)}$. The results are in Column (2) of Table 5. The estimates computed under the assumption of equal selection on observables and unobservables show the potential for large positive biases for both HS_i and $COLL_i$. The fact that the bias estimates for the two different outcomes have the same sign is not surprising, since it reflects the similarity in the effects of X_i on the two education outcomes. While the specific bias estimates are noisy and are probably overstated for reasons discussed above, the large estimate for $COLL_i$ suggests that the 2SLS coefficients are not informative. Finally, for 12th grade math scores, the estimates of 1.72-1.76 (depending on whether C_i is included in the calculations involving $X_i'\gamma$) again do not preclude a small Catholic schooling effect, but given both the evidence of endogeneity and the large standard errors of the 2SLS estimates, we conclude that the 2SLS estimates using D_i are not useful in drawing conclusions regarding test scores.[21]

5 Instrumental Variables Estimates Using Catholic $\times$ Distance

Finally, we turn to the interaction between C_i and D_i as the source of identifying variation. All of the models include both C_i and D_i among the controls. In Column (3) of Table 1 we report probit, bivariate probit, linear probability and 2SLS estimates of the effect of CH_i

[21] The public 8th grade analysis is probably less informative for D_i than for C_i because of the likelihood that distance from Catholic elementary school and distance from Catholic high school are closely related. Consequently, selection issues may have a bigger effect on the coefficient on the index when the distance variables are involved than when only religion is involved.

on high school graduation and college attendance. The bivariate probit and 2SLS point estimates are negative in two of the three cases. Column (3) of Table 2 presents results for test scores. The 2SLS estimates lie below the OLS ones in three of the four cases, with 12th grade math score coefficients being fairly large and negative in both data sets. However, in all cases in NELS:88 the standard errors are too large in relation to the difference between the OLS and 2SLS estimates for the 2SLS estimates to help much in modifying conclusions about α. This is less true in the NLS-72.

We have investigated the properties of the instrument using the same set of procedures that we used for C_i and D_i with the same bottom line. Given the imprecision in some of the estimates and space considerations, we will skip the details.[22] However, the weight of the evidence in Tables 1-5 leads us to be very skeptical of the interaction as an exclusion restriction. In particular, there is evidence in both data sets that the difference between Catholics and non-Catholics in favorable family background characteristics rises with distance from the nearest Catholic high school. If the link between $C_i \times D_i$ and ε_i followed the same pattern, the 2SLS estimates would be biased downward. We suspect that this underlies the negative coefficients for some outcomes in both data sets, particularly NLS-72. We conclude that $C_i \times D_i$ is not a very useful source of variation for the purpose of estimating the Catholic school effect, at least not in the context of NELS:88 or NLS-72.

[22] In Column (4) of Table 3a we report the coefficient on $C \times D6_i$ from regressions of the various background and outcome variables indicated in the rows on C_i, D_i, and $C_i \times D6_i$. The results for the eighth grade measures are mixed, with $C_i \times D6_i$ being positively associated with indicators for whether the student got into a fight at school, but negatively correlated with the "repeated grade" indicator. There are also slight comparative advantages in eighth grade GPA and reading scores. In contrast, family background, student expectations, and parental expectations are generally negatively correlated with $C_i \times D6_i$, with striking differences in parental education levels and expectations.

For NLS-72, the estimates in Table 3b imply that the difference in mother's and father's education between Catholics and non- Catholic students who live within 6 miles of a Catholic high school is 0.33 and 0.32 years lower, respectively, than the difference among Catholic and non Catholic student who live more than 6 miles from a Catholic high school. The incomes of Catholics relative to non-Catholics also rise with distance, and all of these figures are nearly identical to the corresponding ones in NELS:88. Additionally, student educational expectations are strongly correlated with $C_i \times D6_i$, with a coefficient of -0.06 (0.016). We have not investigated why low SES Catholics are disproportionately located near Catholic high schools, but if the unobservable parental traits that influence the outcomes we study follow a similar pattern, then our 2SLS estimates of the effect of Catholic schools are likely to be negatively biased for both the NLS-72 and NELS:88 cohorts.

6 Exclusion Restrictions or Nonlinearity as the Source of Identification? A Comparison of Bivariate Probit and 2SLS

Thus far we have focused on whether the instruments are uncorrelated with the error components. In this section we focus on the power of the instruments for identification in nonlinear models. Evans and Schwab (1995) and Neal (1997) apply bivariate probits of Catholic schooling and educational attainment using data from High School and Beyond and NLSY, respectively, and both papers emphasize the importance of an exclusion restriction in the model for identification. As we have already noted, Evans and Schwab (1995) primarily rely on Catholic religion, excluding it from the outcome equation but including it in the Catholic schooling decision. Neal (1997) uses an indicator for Catholic religion along with county level measures of the density of Catholics (Catholic church adherents as a fraction of county population) in the case of minorities and Catholic religion and Catholic secondary schools per square mile in the county in the case of whites. Both of these papers report positive effects of CH_i on educational attainment that are estimated reasonably precisely, although Evans and Schwab (1995) experiment with 2SLS and obtain implausible results in some specifications.[23] Our bivariate probit results generally follow the same pattern, with estimates being much more precise and reasonable than linear specifications. It is therefore worth investigating the reasons why our instrumental variables results are so noisy and in many cases seem unreasonable, while the bivariate probits tend to generate plausible, precise estimates.

It is useful to start by reviewing identification in the bivariate probit model. The specification used in Neal (1997), Evans and Schwab (1995), and here is

$$CH_i = 1(X_i'\beta + Z_i'\lambda + u_i > 0)$$
$$Y_i = 1(\alpha CH_i + X_i'\gamma + \varepsilon_i > 0),$$

[23] Neal (1997) does not report results based on linear 2SLS, and we were not able to produce them for NLSY79 because we do not have access to the data on Catholic high school attendance. He kindly agreed to run several specifications of the model. The 2SLS results are quite similar to those in Table 6 in that they show very large standard errors in the urban minority sample regardless of which instruments are excluded, as well as in the urban white sample when C_i is not excluded from the model. In the urban white sample when C_i is one of the instruments, the point estimates are reasonable with standard errors around 0.10. In both samples, two stage models in which the first stage is a probit produce results similar to the bivariate probit models that he reports, suggesting that the functional form of the selection equation (and the resulting predicted probability) is primarily what drives the differences between linear two stage least squares and bivariate probit models.

where $1(\cdot)$ is the indicator function taking the value one if its argument is true and zero otherwise, and (u_i, ε_i) are jointly normal each with unit variance but with an unknown correlation. Identification of the α coefficient is the primary focus of these studies. It is well known that exclusion restrictions are useful for semiparametric identification in limited dependent variable models (see, e.g., Heckman, 1990, Cameron and Heckman, 1998, or Taber, 2000), but in this parametric case the linearity and normality assumptions are sufficient, so an exclusion restriction is not necessary. When one uses both exclusion restrictions and functional form restrictions, both contribute to parameter identification in practice. In this subsection we explore whether the source of identification is primarily coming from the exclusion restrictions or primarily coming from the functional form restrictions in the Catholic schools case.

In order to better assess what is identifying the bivariate probit models, as well as to facilitate comparison between the results of this paper and the previous literature, we examine the sensitivity of our results from NLS-72 to different specifications using bivariate probit models of educational attainment. We use a sample design based loosely on Neal (1997), in that we look at individuals from urban areas and examine separate effects for minorities and whites.[24] In contrast to Neal (1997), we focus on college attendance instead of high school graduation due to the sample design of NLS-72. The results are reported in Table 6. Our results are similar to Neal's in several respects. First, the univariate probit coefficient of 0.640 (0.198) implies a large positive effect for non-whites. Second, the coefficient of 0.879 (0.523) from a bivariate probit specification which uses Neal's exclusion restrictions for urban minorities – Catholic religion and the county-level ratio of Catholics to the overall population – is larger than the univariate one, although the difference is not significantly different from zero. Third, the estimates appear at first glance to be of a reasonable magnitude. In particular, the probit coefficients are comparable to the ones reported both in Neal (1997) and in Table 1 of this paper. However, the marginal effects of 0.239 and 0.329 for the univariate and bivariate models, respectively, are suspiciously large.

Table 6 also shows that for urban minorities, the estimated bivariate probit coefficients and standard errors are relatively insensitive to exclusion restrictions and appear to be largely driven by the functional form assumptions embedded in these models. To see this, note that

[24] We have not replicated the analysis for NELS:88 for several reasons. Most importantly, we could not accurately match students to counties, as no county-level identifiers are available in these data at present and zip codes frequently cross county lines.

the precision of the estimates does not vary much across specifications, even when only a "weak" instrument such as $C_i \times D_i$ is excluded or when no instruments are excluded (bottom row). The standard error of the coefficient on CH_i is smaller in both of these cases than when the more powerful instrument, C_i, is excluded, which seems at odds with the notion that the exclusions are driving identification. In contrast, 2SLS estimates swing wildly across specifications, with the results being similar to Evans and Schwab (1995) and our own earlier results; we typically find huge effects with standard errors that are sufficiently large that any plausible estimate would not be significantly different from zero at conventional levels. In the most precisely estimated specification involving all three exclusion restrictions, the 2SLS coefficient of 0.331 (0.254) implies a large effect yet is not statistically significant. In the case of the weakest instrument, $C_i \times D_i$, the 2SLS coefficient of 2.572 (2.442) is so large that it does not make sense within the linear probability framework, yet it is still not significantly different from zero.

The bivariate probit results for whites are again fairly similar across specifications, although the precision of the estimates now varies with the choice of instrument. In the 2SLS case, both precision and the coefficients themselves are relatively constant except when $C_i \times D_i$ is used as an exclusion restriction. It appears that in the urban white subsample, the exclusion restrictions are driving a larger share of identification than they are for urban minorities, but that the linear index assumption in conjunction with normality is still playing a large role. Neal (p. 113 in the notes to Table 6) reports that in high school graduation models the standard error of the bivariate probit estimate of α rises from 0.476 when only Catholic schools/square mile is excluded from the high school graduation equation to 0.589 when there are no exclusion restrictions, which suggests that functional form is playing a substantial role in identification in the NLSY as well.

One problem in interpreting the 2SLS results in Table 6 and in Evans and Schwab's and Neal's data is that both Catholic school and the educational attainment outcomes are binary events, so the imprecision in 2SLS may arise because the linear probability model provides a poor approximation for these decisions relative to the bivariate probit. With this in mind, in Table 6 we take an alternative approach to examining the extent to which nonlinearities are contributing to identification in the nonlinear models. Columns (3) and (6) present results from two stage probit models in which the first stage models the probability of Catholic high school attendance as

$$\Pr(CH_i = 1 \mid X_i, Z_i) = \Phi(X_i'\beta + Z_i'\lambda),$$

where $\Phi(\cdot)$ represents the standard normal cdf and Z_i is the vector of instruments. The second stage includes the X_i variables as controls, but rather than including $\Phi(X_i'\widehat{\beta}+Z_i'\widehat{\lambda})$ as the key variable as is commonly done, we include separate predicted probabilities holding X_i and Z_i constant at their sample means, respectively.[25] The second stage models for college attendance are then

$$(11) \quad \Pr(COLL_i = 1 \mid X_i, Z_i, \overline{X}_i, \overline{Z}_i) = \Phi\left[X_i'\gamma + \alpha_1\Phi(\overline{X}_i'\widehat{\beta} + Z_i'\widehat{\lambda}) + \alpha_2\Phi(X_i'\widehat{\beta} + \overline{Z}_i'\widehat{\lambda})\right].$$

The idea behind this exercise is to isolate the effects of variation in Z_i, given by the second term on the right of (11), from the effects of variation in X_i, which is captured by the third term in brackets in (11). The estimated coefficients α_1, which are reported in the table, measure the extent to which variation in the excluded instruments are influencing college attendance, rather than just nonlinearities in X_i in the score function $\Phi(X_i'\widehat{\beta}+\overline{Z}_i'\widehat{\lambda})$. It is important to point out that this is an informal exercise to explore the extent of the identifying power of Z_i. We do not know of a set of conditions under which we could justify this procedure as a consistent estimate of α.

The most useful information presented in columns (3) and (6) lie in the degree to which the point estimates and standard errors differ from the corresponding estimates given in columns (1) and (4). Consider two extreme cases, one in which C_i, $\%CCH_i$, and CH/P_i are all used as excluded instruments for urban whites, and one in which only $\%CCH_i$, and CH/P_i are used to identify the model for urban minorities. In the first case, the estimate of α_1 of -0.069 (0.125) is similar to the corresponding bivariate probit coefficient of -0.085 (0.118) (column (4)) in terms of both magnitude and precision. In the second case, the estimate of α_1 is 5.541 (2.082), which is markedly different from the bivariate probit estimate of 1.471 (0.442) with respect to both magnitude and precision. We view these contrasting results to mean that variation in the instruments contributes substantially to identification in the first case, but not in the second.

[25] In the specifications we use, two stage probit models in which the outcome models include the estimated predicted probability $\Phi(X_i'\widehat{\beta} + Z_i'\widehat{\lambda})$ of Catholic high school attendance yield estimates which are similar to those obtained from the bivariate probit models reported in Columns (1) and (4) in all cases reported in the table. Neal also finds this to be the case for his preferrred specifications for urban whites and urban minorities (private communication).

Comparing the reported estimates in columns (1) and (3), in every case the point estimates and standard errors differ dramatically, implying that no combination of instruments drives all (or nearly all) of the identification of these models for urban minorities. For urban whites, the exclusion restrictions show substantially more power, but only when Catholic religion (C_i) is used as an instrument - the models using $\%CCH_i$ and CH/P_i or $C_i \times D_i$ still exhibit large discrepancies between columns (4) and (6). The implication is that Catholic religion drives identification in models for urban whites, but none of the other candidate instruments are effective for this sample, and no combination of instruments appears to be powerful for urban minorities. This provides supporting evidence that in many of the estimated models, functional form assumptions are mainly responsible for identification of the bivariate probit estimates.

Although the specifications of Table 6 do not involve exact replications of the analyses of either Evans and Schwab (1995) or Neal (1997), we believe that they do shed some light on the sources of the apparent discrepancies in the results. Table 6 suggests that the proximity measures in both of these studies do not play a key role in identification in NLS-72, as standard errors in the 2SLS models are prohibitively large in cases in which Catholic religion is not an excluded instrument. Joint normality by itself will not always generate reasonable and precise estimates of the Catholic schooling effect, as Grogger and Neal's (2000) analysis shows. (They conclude that bivariate probit models using county level instruments are not very informative in NELS:88 and rely primarily on univariate results in drawing conclusions.) We show that bivariate probits can *sometimes* produce misleading results which are consistent with a powerful instrumental variable, when in fact identification is stemming from a weak instrument in combination with functional form assumptions. To isolate the role of each of these factors, one should experiment with specifications that rely solely on exclusion restrictions for identification.

7 Conclusions and Research Directions

We present evidence on the validity of using three sources of variation in Catholic school attendance – religious affiliation, proximity to Catholic schools, and the interaction between religion and proximity – as a way to identify the effect of attending Catholic high school. We conclude that none of the candidate instruments is useful as a source of identification of the Catholic school effect, at least in the NELS:88 data set. While we will not attempt to restate all the results, a key concern lies in the fact that we find a strong relationship between Catholic religion and educational achievement in the sample of public eighth graders, who almost never attend Catholic high school. Similarly, we find a strong relationship between distance from the nearest Catholic high school and college attendance among public eighth grade students. We also find a fairly strong relationship between the instruments and the index of observed variables that determine the outcomes. Finally, we show that the nonlinearity inherent in bivariate probit is the main source of identification when measures of proximity to Catholic school and/or the density of the Catholic population are the only excluded instruments. Users of the bivariate probit model in other settings should investigate the relative contribution of nonlinearity and exclusion restrictions to identification.

In the absence of good instruments, what can one do? In AET we develop a new approach to estimation based on the use of the degree of selection on observables as a guide to selection on unobservables. Using this approach in AET (2002 and 2005) we obtain lower bound estimates of the Catholic school effect and conclude that there is a substantial positive effect on high school graduation but not test scores. A second strategy is to search for new instruments. For example, in work in progress, Susan Dynarski and Jonathan Gruber are obtaining data on tuition levels and tuition discounts based on the number of children in a family, with the goal of using the dependence of tuition on family size as a source of identifying variation. A third, complementary research avenue would involve isolating additional family background control variables that can eliminate the direct association between Catholic religion and outcomes using other data sets, such as Children of NLSY79, which have richer data on parents and the home environment. Finally, no one has yet attempted to specify and estimate a structural model of location choice, school choice and education outcomes. Ultimately this will require assumptions about human capital production technologies and household preferences over various education outputs, travel and monetary costs of educational alternatives, and instruction in religion. The construction of such a model strikes us as a very difficult undertaking, but the process of building it might lead to sharper thinking about what the key omitted variables are and about the mechanisms through which Catholic schooling might affect outcomes. Ultimately, progress on this problem, like many others, will be made by examining it using a variety of different empirical approaches.

References

[1] Alexander, Karl L., and Aaron M. Pallas, "School Sector and Cognitive Performance: When is a Little a Little," *Sociology of Education*, LVIII (1985), 115-128.

[2] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," unpublished paper, Northwestern University, (2002).

[3] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, forthcoming (2005).

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

[5] 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.

[6] Bryk, Anthony S., Valerie E. Lee, and Peter B. Holland, *Catholic Schools and the Common Good*, Cambridge, Mass. : Harvard University Press, 1993.

[7] Cameron, Stephen V. and Heckman, James J., "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy,* 106(1998).

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

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

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

[11] 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).

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

[13] Elder, Todd, "Essays in the Economics of Aging and Education," Northwestern University Ph.D. Thesis (2002).

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

[15] 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.

[16] Evans, William N., and Robert M. Schwab, "Who Benefits from Private Education: Evidence from Quantile Regressions," Department of Economics Working Paper, University of Maryland, August 1993.

[17] Figlio, David N., and Joe A. Stone, "Are Private Schools Really Better?," *Research in Labor Economics*, (2000).

[18] 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.

[19] Grogger, Jeff and Derek A. Neal, "Further Evidence on the Effects of Catholic Secondary Schooling," *Brookings-Wharton Papers on Urban Affairs,* 2000, 151-201.

[20] Heckman, James J., "Varieties of Selection Bias," *American Economic Review,* 80(1990).

[21] Hoxby, Caroline Minter., "Do Private Schools Provide Competition for Public Schools," NBER Working paper, 1995.

[22] Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, Vol. 62, No. 2. (Mar., 1994), pp. 467-475.

[23] Jepsen, Christopher, "The Effectiveness of Catholic Primary Schools, " forthcoming in *Journal of Human Resources.*

[24] 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.

[25] 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.

[26] Murnane, Richard J., "A Review Essay - Comparisons of Public and Private Schools: Lessons from the Uproar." *Journal of Human Resources* 19 (1984), 263-77.

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

[28] Roy, A. "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (1951), 135-146.

[29] Sander, William H. *Catholic Schools: Private and Social Effects.* Boston: Kluwer Academic Publishers, 2001.

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

[31] Taber, Christopher, "Semiparametric Identification and Heterogeneity in Discrete Choice Dynamic Programming Models," *Journal of Econometrics* 96(2000), 201-229.

[32] Udry, Christopher, "Gender, Agricultural Production, and the Theory of the Household", *Journal of Political Economy*, 104 (1996), 1010-1046.

[33] Willis, Robert J. and Sherwin Rosen, "Education and Self-Selection," Journal of Political Economy,University of Chicago Press, vol. 87 (October 1979), pp. S7-36

[34] Udry, Christopher, "Gender, Agricultural Production, and the Theory of the Household", *Journal of Political Economy*, 104 (1996), 1010-1046.

[35] 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.

Table 1

Probit, Bivariate Probit, OLS, and 2SLS Estimates of Catholic Schooling Effects
NELS:88 and NLS-72
Weighted, Marginal Effects of Nonlinear Models Reported, (Huber-White Standard Errors in Parentheses)

	Excluded Instruments		
	(1)	(2)	(3)
	Catholic (C_i)	*Distance* (D_i)	*Catholic*×*Distance* ($C_i \times D_i$)
HS Graduation (NELS:88)			
Probit (controls exclude "instrument")	0.065	0.047	0.052
	(0.025)	(0.025)	(0.026)
Bivariate Probit	0.128	-0.007	-0.022
	(0.032)	(0.085)	(0.119)
OLS	0.041	0.021	0.023
	(0.014)	(0.014)	(0.015)
2SLS	0.34	-0.04	0.09
	(0.08)	(0.10)	(0.11)
College in 1994 (NELS:88)			
Probit (controls exclude "instrument")	0.094	0.085	0.077
	(0.022)	(0.022)	(0.022)
Bivariate Probit	0.170	0.103	-0.043
	(0.055)	(0.062)	(0.070)
OLS	0.128	0.119	0.111
	(0.026)	(0.026)	(0.026)
2SLS	0.40	0.31	-0.11
	(0.10)	(0.11)	(0.12)
College in 1976 (NLS-72)			
Probit (controls exclude "instrument")	0.068	0.070	0.067
	(0.016)	(0.016)	(0.016)
Bivariate Probit	-0.002	-0.052	-0.080
	(0.028)	(0.035)	(0.035)
OLS	0.071	0.075	0.072
	(0.015)	(0.016)	(0.016)
2SLS	0.06	0.44	-0.25
	(0.04)	(0.20)	(0.11)

Notes:

(1) All models other than univariate probits instrument for Catholic High School attendance (CH_i).

(2) Controls for all NELS:88 models include the demographic, family background, geography, and 8th grade variables listed in Table 3a. Controls for all NLS-72 models include the demographic, family background, and geography variables listed in Table 3b. When D_i is used as an instrument, C_i is included as a control; when C_i is an instrument, D_i is included; and when $D_i \times C_i$ is an instrument, both D_i and C_i are included.

(3) Sample sizes: N=8560 (HS Graduation), N=8313 (College Attendance in NELS), N=19,489 (College Attendance in NLS-72)

Table 2

OLS and 2SLS estimates of Catholic Schooling Effects
NELS:88 and NLS-72
Weighted, (Huber-White Standard Errors in Parentheses)

	Excluded Instruments		
	(1)	(2)	(3)
	Catholic (C_i)	*Distance* (D_i)	*Catholic*×*Distance* ($C_i \times D_i$)
12th Grade Reading Score (NELS:88)			
OLS	1.16 (0.37)	1.03 (0.37)	1.14 (0.38)
2SLS	1.40 (1.54)	-1.09 (1.84)	1.24 (1.82)
12th Grade Math Score (NELS:88)			
OLS	1.03 (0.31)	1.00 (0.31)	0.92 (0.32)
2SLS	2.64 (1.21)	2.43 (1.45)	-2.63 (1.57)
12th Grade Reading Score (NLS-72)			
OLS	2.06 (0.34)	2.54 (0.37)	2.50 (0.36)
2SLS	-1.34 (0.99)	8.69 (4.53)	0.50 (2.32)
12th Grade Math Score (NLS-72)			
OLS	1.52 (0.33)	1.77 (0.35)	1.71 (0.36)
2SLS	-0.07 (0.96)	11.05 (4.47)	-3.94 (2.27)

Notes:
(1) All 2SLS models instrument for Catholic High School attendance (CH_i).
(2) Controls for all models include those described in notes to Table 1. When D_i is used as an instrument, C_i is included as a control; when C_i is an instrument, D_i is included; and when $D_i \times C_i$ is an instrument, both D_i and C_i are included as controls.
(3) Sample sizes: N=8,166 (NELS 12th Reading), N=8,119 (NELS 12th Math)
N=16,276 (NLS Academic Years of School), N=14,671 (NLS Reading and Math scores),

Table 3a

Comparison of Means of Key Variables
by Value of Distance, Catholic, and their Interaction
NELS:88

	(1)	(2)	(3)	(4)
	Overall Mean	Difference by C_i	Difference by D_i	Difference by $C_i \times D_i$
Demographics				
Female	0.50	0.01	0.00	0.00
Asian	0.04	0.01	0.04	-0.02
Hispanic	0.10	0.19	0.08	0.03
Black	0.13	-0.15	0.08	-0.13
White	0.73	-0.05	-0.20	0.12
Family Background				
Mother's education	13.14	-0.26	0.17	-0.36
Father's education	13.42	-0.07	0.17	-0.31
Log of family income	10.20	0.11	0.12	-0.02
Mother only in house	0.15	-0.04	0.02	-0.03
Parent married	0.78	0.06	-0.02	0.03
Geography				
Rural	0.32	-0.15	-0.44	0.05
Suburban	0.44	0.06	0.08	0.00
Urban	0.24	0.09	0.36	-0.05
Expectations				
Schooling expectation	15.17	0.15	0.31	-0.06
Very sure to graduate high school	0.83	-0.01	0.00	-0.01
Parents expect some college	0.88	0.04	0.05	-0.02
Parents expect college grad	0.78	0.03	0.06	-0.04
Expect white collar job	0.46	0.03	0.06	-0.01
8th Grade Variables				
Delinquency Index	0.69	-0.05	0.03	-0.04
Got into fight	0.27	-0.01	0.01	0.05
Rarely completes homework	0.21	-0.05	0.00	0.00
Frequently disruptive	0.13	-0.02	-0.01	0.00
Repeated grade 4-8	0.08	-0.03	0.01	-0.03
Risk Index	0.72	-0.07	-0.01	0.01
Grades Composite	2.89	0.04	0.00	0.07
Unpreparedness Index	10.82	0.00	0.08	-0.09
8th Grade reading score	50.32	0.40	0.03	1.15
8th Grade math score	50.33	0.55	0.45	0.06
Outcomes				
10th Grade reading score	50.16	0.65	0.58	0.60
10th Grade math score	50.21	0.93	0.75	-0.50
12th Grade reading score	50.40	0.52	0.88	-0.17
12th Grade math score	50.38	1.18	1.03	-0.70
Enrolled in 4 year college in 1994	0.29	0.08	0.08	-0.05
HS Graduate	0.84	0.07	0.01	0.01
Attended Catholic HS	0.06	0.13	0.12	0.15

Notes:
(1) Difference by $C_i \times D_i$ is obtained from the coefficient on $C_i \times D_i$ in a regression including C_i and D_i as controls
(2)Sample Size: N=16,070

Table 3b

Comparison of Means of Key Variables by Value of Distance, Catholic, and their Interaction NLS-72

	(1)	(2)	(3)	(4)
	Overall Mean	Difference by C_i	Difference by D_i	Difference by $C_i \times D_i$
Demographics				
Female	0.50	-0.01	0.03	0.03
Hispanic	0.04	0.11	0.01	-0.07
Black	0.15	-0.15	0.04	-0.08
Family Background				
Mother's education	12.19	-0.13	0.16	-0.33
Father's education	12.43	0.06	0.40	-0.32
Log of family income	8.93	0.07	0.11	-0.03
Father Blue Collar	0.24	0.01	-0.03	-0.01
Low SES Indicator	0.29	-0.05	-0.06	0.00
English Primary Language	0.92	-0.06	-0.02	0.03
Family Receives Daily Newspaper	0.88	0.04	0.06	0.01
Mother Works	0.50	-0.06	0.03	0.01
Geography				
Rural	0.23	-0.14	-0.30	0.05
Suburban	0.48	0.06	0.02	-0.04
Urban	0.29	0.08	0.28	-0.01
Expectations				
Decided to go to college pre-HS	0.41	-0.01	0.04	-0.06
Outcomes				
Enrolled in college by 1976	0.38	0.01	0.05	-0.06
Reading Score	50.01	0.30	0.46	0.55
Math Score	49.98	0.58	0.40	-0.10
Years of Academic PSE, 1979	1.61	0.03	0.22	-0.23
Attended Catholic HS	0.06	0.19	0.07	0.15

Notes:

(1) Difference by $C_i \times D_i$ is obtained from the coefficient on $C_i \times D_i$ in a regression including C_i and D_i as controls

(2)Sample Size: N=19,921

Table 4

Comparison of 2SLS Estimates[1] and Bias Implied by OLS Estimation of $Y_i = X_i'\gamma + [Z_i'\widehat{\lambda}]\psi + \omega_i$ on the Public Eighth Grade Subsample[2]; Various Outcomes and instruments; NELS:88 Sample Weighted, (Huber-White Standard Errors in Parentheses)

OUTCOME (Y)	**INSTRUMENTS** (Z_i)		
	(1)	(2)	(3)
	Catholic	*Distance*	*Catholic×Distance*
High School Graduation			
Implied Bias in 2SLS (ψ)	0.34 (0.08)	-0.05 (0.12)	0.15 (0.12)
2SLS Coefficient	0.34 (0.08)	-0.04 (0.10)	0.09 (0.11)
College Attendance			
Implied Bias in 2SLS (ψ)	0.29 (0.11)	0.37 (0.12)	-0.23 (0.13)
2SLS Coefficient	0.40 (0.10)	0.31 (0.11)	-0.11 (0.12)
12th Grade Reading Score			
Implied Bias in 2SLS (ψ)	0.54 (1.68)	-0.51 (2.08)	-0.50 (1.99)
2SLS Coefficient	1.40 (1.54)	-1.09 (1.84)	1.24 (1.82)
12th Grade Math Score			
Implied Bias in 2SLS (ψ)	1.85 (1.41)	1.83 (1.69)	-4.37 (2.06)
2SLS Coefficient	2.64 (1.21)	2.43 (1.45)	-2.63 (1.57)

Notes:
(1) Controls for all models include those described in notes to Table 1. In Column 1, D is included as a control; in Column 2, C_i is included as a control; and in Column 3, both D_i and C_i are included as controls.
(2) The model $Y_i = X_i'\gamma + [Z_i'\widehat{\lambda}]\psi + \omega_i$ is estimated by OLS using the NELS:88 sample of those who attended public eighth grade schools. Sample sizes: N=7,701 (HS Graduation), N=7,481 (College Attendance), N=7377 (12th reading), N=7380 (12th math). $\widehat{\lambda}$ is the coefficient on Z_i in the first stage equation for CH_i. The sample sizes for the first stage equations are listed in Tables 1 and 2 for the various outcomes. The 2SLS coefficients are from Tables 1 and 2.
(3) Reported standard errors of ψ account for the fact that $\widehat{\lambda}$ is previously estimated from a model of CH_i attendance.

Table 5

Estimates of Catholic Schooling Effects and Estimates of Potential Bias Using *AET* Methodology, NELS:88 Weighted, (Huber-White Standard Errors in Parentheses)

	Excluded Instruments		
	(1)	(2)	(3)
	Catholic	*Distance*	*Catholic×Distance*
HS Graduation			
2SLS Coefficient	0.34 (0.08)	-0.04 (0.10)	0.09 (0.11)
Bias 1	0.52 (0.23)	0.15 (0.16)	0.14 (0.24)
Bias 2	0.84 (0.26)	0.06 (0.14)	...
College in 1994			
2SLS Coefficient	0.40 (0.10)	0.31 (0.11)	-0.11 (0.12)
Bias 1	0.45 (0.21)	0.46 (0.22)	0.15 (0.26)
Bias 2	0.45 (0.21)	0.40 (0.20)	...
12th Reading Score			
2SLS Coefficient	1.40 (1.54)	-1.09 (1.84)	1.24 (1.82)
Bias 1	1.18 (1.06)	2.49 (1.59)	2.59 (1.14)
Bias 2	1.42 (1.07)	2.11 (1.40)	...
12th Math Score			
2SLS Coefficient	2.64 (1.21)	2.43 (1.45)	-2.63 (1.57)
Bias 1	2.02 (0.75)	1.76 (1.03)	1.42 (0.88)
Bias 2	1.87 (0.74)	1.72 (0.98)	...

Notes:

(1) Controls included are described in Table 1 notes.

(2) Sample sizes: N=8560 (HS Graduation), N=8313 (College Attendance in NELS), N=8,166 (12th Reading), N=8,199 (12th Math).

(3) “Bias 1” calculations use all variables, while ”Bias 2” excludes D_i and C_i in the bias calculations.

(4) Standard Errors of the bias calculations obtained from a 100-replication bootstrap

Table 6
Comparison of Linear and Non-Linear Models of College Attendance in NLS-72
(Standard Errors in Parentheses)
[Marginal Effects of Non-Linear Models in Brackets]

	Sample					
	Non-whites in cities (N=1532)			Whites in cities (N=5326)		
	Nonlinear Models (Probits) (1)	Linear Models (OLS/2SLS) (2)	Nonlinear Models Holding X_i Constant[4] (3)	Nonlinear Models (Probits) (4)	Linear Models (OLS/2SLS) (5)	Nonlinear Models Holding X_i Constant[4] (6)
Single Equation Model	0.640	0.239		0.253	0.093	
(OLS/Probit)	(0.198)	(0.070)		(0.062)	(0.022)	
	[0.239]			[0.093]		
Two Equation Models:						
Excluded Instruments:						
$\%CCH_i$ and CH/P_i	1.471	1.375	5.541	0.048	0.115	0.084
	(0.442)	(0.583)	(2.082)	(0.250)	(0.158)	(0.783)
	[0.517]		[0.706]	[0.018]		[0.031]
C_i and $\%CCH_i$	0.879	0.054	0.012	-0.090	-0.036	-0.084
	(0.523)	(0.309)	(1.443)	(0.121)	(0.050)	(0.148)
	[0.329]		[0.004]	[-0.033]		[-0.031]
C_i, $\%CCH_i$, and CH/P_i	1.106	0.331	1.302	-0.085	-0.034	-0.069
	(0.460)	(0.254)	(0.706)	(0.118)	(0.048)	(0.125)
	[0.409]		[0.471]	[-0.031]		[-0.025]
C_i only	0.761	-0.093	-0.505	-0.133	-0.056	-0.149
	(0.543)	(0.324)	(1.638)	(0.130)	(0.054)	(0.151)
	[0.285]		[-0.148]	[-0.049]		[-0.054]
$C_i \times D_i$	1.333	2.572	1.409	-0.121	-0.395	2.624
	(0.516)	(2.442)	(1.276)	(0.262)	(0.169)	(5.173)
	[0.478]		[0.497]	[-0.044]		[0.559]
None	1.224	...		-0.094	...	
	(0.542)	...		(0.301)	...	
	[0.446]			[-0.034]		

Notes:
(1) Sample is taken from counties in the NLS-72 which had a population of greater than 250,000 in 1980.
(2) All equations control for parents' education and income levels and SES, whether father is a blue-collar worker, county population, gender and race.
(3) The Instrument "$\%CCH_i$" refers to the percent of the county which reports they are Catholic church members, and "CH/P_i" to Catholic schools per person in the county.
(4) "Nonlinear Models Holding X_i constant" refer to two stage probit models in which the second stage includes the full set of controls apart from the excluded instruments, the predicted probability of Catholic high school attendance holding Z_i constant, and the predicted probability of Catholic high school attendance holding X_i constant. The coefficient reported is that associated with the latter term.