View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

AN ALTERNATIVE INSTRUMENT FOR PRIVATE SCHOOL COMPETITION

D. Cohen Zada

Discussion Paper No. 07-05

June 2007

Monaster Center for Economic Research
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva, Israel

Fax: 972-8-6472941
Tel: 972-8-6472286

An Alternative Instrument for Private School Competition*

D. Cohen-Zada**

May 2006

Abstract

Empirical studies estimating the effect of private school competition on student outcomes commonly use the share of Catholics in the local population as an instrument for private school competition. However, it has recently been argued that since this instrument is likely to be correlated with unmeasured student characteristics that vary across localities, it cannot be a valid instrument for private school competition. I suggest using instead the local share of Catholics in the population in 1890 and its squared term. I show that these instruments are very strong and are also exogenous to both student achievements and private school competition. These instruments can also be applied to estimate the treatment effect of Catholic schools.

Keywords: School choice, demand for schooling, Catholic share, competition.

JEL classification: I20

* I thank Moshe Justman for many helpful discussions. Comments from the editor and the referee significantly improved the paper. I also thank Naomi Feldman, Christopher Jepsen, and William Sander for their comments and suggestions, and to Roger Finke for his kind guidance regarding the data of the American Religion Data Archive.

** Department of Economics, Ben-Gurion University, P.O. Box 653, Beer Sheva 84105, Israel; *Tel.*: 972-8-6472301; *Fax*: 972-8-6472941; *e-mail*: danoran@bgu.ac.il

1. Introduction

One of the main arguments in support of school choice reforms is that more competition from private schools would increase the quality of public education (Friedman, 1962). In the last decade numerous empirical studies have tried to quantify the effect of competition on student outcomes. Many of these studies used private school enrollment rates to measure competition by estimating an education production function to which they added the private enrollment rate in the student's county of residence as an additional determinant of student outcomes. Since the local private enrollment rate is endogenous to public school quality, most of these studies used the share of Catholics in the population in the student's county of residence as an instrument for the local private enrollment rate (Hoxby 1994, Dee 1998, Sander 1999, Jepsen 2002, among others). Hoxby (1994) was the first study to use this instrument claiming that this is a valid instrument since "religious composition of an area is largely a matter of historical accident" (p. 2). Most later studies followed Hoxby's argument in justifying this instrument.

However, although Hoxby's argument can justify using the historic Catholic share as an instrument for private school competition, it is less valid for the current Catholic share for several reasons. For one thing, Catholics have become quite mobile. As Table 1 shows, the correlation between the Catholic share in 2000 and at earlier points decreases as we go back in time. Thus, if Catholics choose where to live and their current location decisions are correlated with unobserved student characteristics, the current Catholic share in the local population cannot be excluded from the outcome equation. Consequently, the Catholic share in the population cannot serve as a valid instrument for private school competition (Grogger and Neal, 2000; Altonji et.al, 2005a). In addition, if the location decisions of Catholics are correlated with unobserved variables that affect the demand for private schooling, the share of Catholics in the local population is correlated with the error term of the first stage school-choice estimation. This would render the Catholic share in the population endogenous to private school enrollment, thus implying that the first-stage estimation of the local private enrollment rate would yield biased estimates, which, in turn, would generate a bias in the estimated effect of private school competition on student outcomes. Hence, in order to serve as a valid instrument for private school competition, the Catholic share in the population must be exogenous to private school competition.

In this paper we suggest the Catholic share in 1890 and its squared term as alternative instruments for private school competition. Our reasoning is that as one goes back in time one can be more certain that the historical Catholic share is orthogonal to unmeasured student characteristics, which are known to affect individual student achievements. We chose 1890 because it is the earliest year for which data are available on the Catholic share in the population, and because we find the Catholic share in 1890 to be a very strong instrument for private school competition. Moreover, as the distribution of the Catholic population has changed drastically since 1890, the Catholic share in 1890 is substantially less likely to correlate with current unobserved characteristics that influence student achievement and vary across localities.[1] Thus, if there are unmeasured characteristics that affect student achievements they are likely to have changed over more than one hundred years, during which large migration flows substantially altered the composition of local communities.

[1] Finke and Stark (2005) who describe the immigration of Catholics to the US in the years 1870-1926 mention that "the largest increase in the proportion who were Catholics occurred between 1890, when the census recorded more than seven million Catholics (making up 12 percent of the population), and 1906, when there were more than 14 million Catholics (making up 17 percent of the population). From 1906 through 1926 the Catholic 'market share' remained constant at 16 percent of the population."

For the Catholic share in 1890 and its squared term to serve as strong instruments for private school competition they must be correlated with the current private enrollment rate. To show that this is the case, we use data from a cross-section of 2640 counties in the United States for the years 1990 and 2000 and estimate the private enrollment rate as a function of the Catholic share in 1890 and its squared term, controlling for demographic variables that were found to be significant in previous studies.[2] Then, we use the Stock and Yogo (2005) weak instrument test to ascertain that our instruments are not weak. The results show that the share of Catholics in the population in 1890 has a very significant positive concave effect on the private enrollment rate, and the Stock and Yogo (2005) test indicates that our instruments are very strong. These results are robust to several different specifications and estimation methods.

We then show that the commonly maintained assumption that the current Catholic share in the local population is exogenous to the local private enrollment rate is not valid.[3] To do this, we first estimate the determinants of the private enrollment rate using the Catholic share in 1890 as an instrument for the current Catholic share. Then, using several exogeneity tests we show that the current Catholic share in the population is endogenous to the local private enrollment rate. Moreover, we show that assuming that the current Catholic share in the population is exogenous to the private enrollment rate yields a Catholic share effect on private school competition that is biased down by approximately 33%. This result reinforces our argument that the current Catholic share cannot serve as a valid instrument for the local private enrollment rate.

[2] Data on the Catholic share in the population in 1890 were available for about 86% of the counties.

[3] Past school-choice estimations generally treated the Catholic share in the population as an exogenous determinant of the demand for private education (Clotfelter 1976; James 1987, Hamilton and Macauley 1991, West and Palsson 1988, Cohen-Zada and Justman 2003, among others).

Finally, to show that the Catholic share in 1890 is exogenous to the private enrollment rate, we add as an additional source of identification the Catholic share in 1906 and test the identifying restrictions.[4] The results show that we cannot reject the joint null hypothesis that the Catholic shares in 1890 and 1906 and their squared terms are valid instruments. We conclude the paper with a short discussion on how our instruments may also contribute to identify the treatment effect of Catholic schools.[5]

2. Background and related literature

There are numerous studies on the effect of private school competition on student outcomes, many of which measure private school competition according to local private enrollment rates. Several of these studies have used the current Catholic share in the population as an instrument for private school competition; a brief summary of these works is presented below.[6]

Hoxby (1994) justified the use of the current Catholic share as an instrument for private school competition based on two arguments. First, this instrument strongly correlates with local private enrollment rates since parents who live in areas with a high percentage of Catholics pay lower tuition rates and lower transportation costs in traveling to their school of choice, and have more Catholic school options. Her first-stage regression indicated that the Catholic share has a positive concave effect on the private enrollment rate. Her second argument was that the Catholic share is exogenous to student achievements. Using individual data from the NLSY, she found that greater private school competitiveness significantly increases the quality of public schools, as measured by educational attainment, wages, and high school graduation

---

[4] We are able to use the Catholic shares in 1890 and 1906 as two different sources of identification since the Catholic population changed substantially between these years (see note 1).

[5] Previous efforts to estimate the treatment effect of Catholic schools include, among others, Sander and Krautmann (1995), Neal (1997), Sander (1996), Sander (2000), and recently Altonji et al. (2005b).

[6] For a more comprehensive review of the research in this field see Belfield and Levin (2002).

rates of public school students. Dee (1998) used county- and district- level data from 18 states that contain consistently-defined high school graduation rates. In order to identify the causal effect of private school competition on student outcomes he first ranked school districts, in descending order, by the Catholic share in the population, and then created four dummies each indicating the quintile in this ranking. His 2SLS estimates indicate that competition from private schools have a significant positive effect on high school graduation rates of adjacent public schools. In addition, he pointed out that "OLS consistently and dramatically underestimates the effect of competition from private schools on the level of achievements in public schools" (p. 423).

In contrast, some studies indicated either no significant effect or mixed effects of private school competition on student outcomes. Using school-level data, Sander (1999) found that within the state of Illinois the percentage in private schools has no significant effect on public school achievements. Jepsen (2002) used two individual data sets: the National Longitudinal Survey of Youth 1979 and the National Educational Longitudinal Survey of 1988. Similar to Hoxby (1994), his first stage regression indicated that the Catholic share has a positive concave effect on the private enrollment rate.[7] He found that the estimated effect of private school competition on student achievement depends on the choice of the dataset (NLSY or NELS88), the measure of student achievement, and the aggregation level of the competition variables. He concluded that private school competition does not have a consistently positive significant effect on student achievements.

More recently, using a large school district level dataset in upstate New York, Greene and Kang (2004) found that private school competition has a significant

[7] Cohen-Zada (2006) provides a fully fledged model explaining why the Catholic share has a non-linear concave effect on the demand for private education.

positive effect on the total average score of mathematics and science but a negative effect on the percent of students receiving a Regents diploma. Following Hoxby (1994) and Jepsen (2002), they also used as instruments the Catholic share in the population and its squared term, but found that the Catholic share has a *convex* rather than a concave effect on the private enrollment rate. Finally, using data on school districts in Georgia, Geller et al. (2006) found that third- and tenth-grade test scores for both reading and mathematics are not significantly higher in areas with greater private school competition. In summary, as we can see from these findings, no consensus has been reached regarding the effect of private school competition on student outcomes.

3. Data

We combine data from six sources which are matched geographically. County data on K-12 enrollment by school type were created by the National Educational Data Resource Center using school-level data from the Public Elementary/Secondary School Universe Survey and the Private School Survey. We supplemented these data with demographic variables taken from the County and City Data Books 1994 and 2000. County data on the share of population that lives in a rural area were taken from the STF3 files of the 1990 and 2000 census. Data on the number of Catholic members and the share of Catholics in each county population in 1990 and 2000 were taken from the Religious Congregation and Membership in the US (2000). Historical data on the number of Catholic members in each county in 1890 were made available by the American Religion Data Archive and were originally collected by the Census Office as part of the 11th Census. The U.S. Census collected data on the number of members of each denomination from 1890 through 1936. Historical data on the number of Catholic members in each county in 1906 were also made available by the

American Religion Data Archive though the data were originally collected by the United States Census of Religious Bodies.[8]

The reliability of these data has previously been convincingly discussed and verified in a study by Rodney Stark (1992), a well-known sociologist of religion, who has published numerous studies which rely on these data. He outlines several reasons for trusting these data. First, they are in line with the picture drawn by historians regarding the number of church members in the US in the end of the nineteen century. Second, the data are "extremely stable over space and time" (p. 92). Third, the Bureau describes a very careful procedure for collecting the data (see note 8). In addition, the Bureau also reports that they compared their results with other sources whenever possible. Stark concludes that "[i]t is time that we accepted the dedication and sophistication of data collectors long dead. Counting is not a recent invention, and a wealth of good quantitative historical data awaits analysis" (p. 94).

Historical county data on the size of population in 1890 and 1910 were made available by the Geospatial and Statistical Data Center at the University of Virginia and were originally collected by the Decennial Census of the Unites States of 1890 and 1910.

We obtained the share of Catholics in each county population by dividing the number of Catholic members of each county by its total population. As data on the size of population in 1906 were not available, we obtained a proxy for the share of Catholics in the population in 1906 by dividing the number of Catholic members in

[8] The method for collecting the data adopted in 1906 was different from that of 1890. In 1890, the Census Office used the diocese as the ecclesiastical unit for gathering the data rather than contacting the local churches directly. The correspondence would be addressed to the bishops in charge of each diocese and they were requested to furnish the necessary information. For 1906, however, the Bureau of the Census contacted the leaders of each identifiable denomination in the US and asked them to provide lists of churches. Then, these lists were used to contact local church leaders directly. Churches that did not respond were sent several follow-up surveys and as a final step they were visited by a census officer.

1906 by the total county population in 1910. Data on the Catholic share in 1890 and 1906 were available for 86% and 94% of the counties, respectively. This left us with 5280 observations for which data were available on the Catholic share in 1890, and 5254 observations for which data were available for both 1890 and 1906. Table 2 presents the number of observations in each 10% range in the distribution of the Catholic share in the population for 1890, 1906 and 1990-2000. It shows that for all the years in more than 90% of the counties the Catholic share in the population was lower than 40%. On the other hand, there is one observation for which the Catholic share in 1990-2000 is higher than 100%, and two such observations in 1906 and four in 1890. According to *Religious Congregation and Membership* 2000, which reports the data for 1990-2000, this discrepancy can be explained by US undercount, church membership overcount, and county of residence differing from county of membership. That is, as the membership data for 1990-2000 are gathered from the Congregations, the reported share of Catholics in the county population may be biased up or down if Churches from one county draw in Catholics from surrounding counties. This bias is more likely to occur in the few counties in which the reported share of Catholics in the population is very high. In these counties, the actual share of Catholics may be substantially lower if these are urban areas that attract a large number of Catholics from surrounding areas. In this case, if Catholics from surrounding counties are less likely than local Catholics to attend the local Catholics schools, the correlation between the share of Catholics in the population and the private enrollment rate may drop off. We deal with this concern in the empirical estimation.

Detailed data sources are provided in the Appendix, and Table 3 presents descriptive statistics. The table shows that the average Catholic share among the counties was 5.50% in 1890, 6.7% in 1906, 12.88% in 1990, and 13.41% in 2000. The correlation between the share of Catholics in 1890 and 1906 is 0.74. Finally, we obtained the private enrollment rate by dividing enrollment in private schools by total K-12 enrollment. Table 2 shows that the average private enrollment rate was 5.33% in 1990, and slightly increased to 5.54% in 2000. In about 30% of the counties the local private enrollment rate was zero.

4. Empirical estimation

4.1 Instrument relevance

A concern that may arise about our instruments is that as we went quite far back in time, the correlation between the Catholic share in 1890 and the current private enrollment rate is not high enough and thus our instruments may be considered weak. In a seminal paper, Bound, Jaeger and Baker (1995) discuss how instrumental variables can perform poorly if the instruments are weak and conclude that "the use of instruments that jointly explain little variation in the endogenous variable can do more harm than good" (p. 449). Specifically, they focused on two problems that may arise owing to weak instruments. First, if the correlation between the instruments and the endogenous variable is low, even a weak correlation between the instruments and the error in the structural equation can lead to large inconsistencies in the IV estimates. Second, IV estimates are biased in the same direction as OLS estimates, with the magnitude of the bias increasing as the F statistic on the excluded instruments in the first stage regression of IV approaches zero. Moreover, even enormous sample sizes do not guarantee that finite-sample biases will be eliminated from IV estimates. In regard to detecting "weak instruments", they pointed out that "F statistics close to 1 should be cause for concern" (p. 446). Staiger and Stock (1997) suggested that instruments be considered weak if the F statistic on the excluded instruments in the first stage is less than 10. More recently, Stock and Yogo (2005) developed two alternative quantitative definitions of weak instruments. First, a set of instruments is weak if the bias of the IV estimator, relative to the bias of the OLS estimator, exceeds a certain threshold. The second definition is that a set of instruments is weak if the α level Wald test of the endogenous variable, based on the IV statistics, has a size that exceeds a certain limit. For each definition, they also provided a table of critical values that enable using the first stage *F*-statistic to test whether given instruments are weak. [9] In the next section we implement the Stock and Yogo (2005) test in two stages. First, we estimate the private enrollment rate as a function of the Catholic share in 1890 and its squared term, controlling for a set of demographic variables that were found to be significant in previous studies. Then, we test whether the *F*-statistic on the excluded instruments exceeds the relevant critical value proposed by Stock and Yogo (2005).

4.2 The variables

In estimating the private enrollment rate we include the following demographic variables, which were found to be significant in previous studies:

Catholic share in 1890. Previous studies of school choice estimated the demand for private school enrollment as a function of the current Catholic share in the population, assuming that it is an exogenous variable. Most of these studies assumed a linear relationship between enrollment in private schooling and the share of Catholics in the population and found significant positive effects (Clotfelter 1976, James 1987, Long and Toma 1988, Hamilton and Macauley 1991, among many others). More recent studies assumed a quadratic relationship between the two variables, and found that the

[9] Hann and Hausman (2002) present an alternative test for detecting weak instruments that is based on reverse regressions. However, this test has been shown by Hausman et al. (2005) to have low asymptotic power, and Andrews et al. (2005) recommend not using it for detecting weak instruments.

share of Catholics in the population has a positive concave effect on the demand for private schooling (Hoxby 1994, Jepsen 2002, Cohen-Zada and Justman 2003, Cohen-Zada 2006). Therefore, we expect the historical Catholic share in 1890 to have a positive concave effect on the private enrollment rate.

Mean income reflects parents' ability to pay for differentiated private education (private education is costly while public education is free). Therefore, we expect this variable to have a positive effect on the private enrollment rate (Sonstelie, 1982; West and Palsson, 1988; Cohen-Zada and Justman, 2003).

Density of population affects the cost of education in general, but more so in private schooling, where scale effects and transportation costs are generally more pronounced than in public schooling. Therefore, we expect this variable to have a positive effect on the private enrollment rate.

Share of population that lives in a rural area. Living in a rural area has been shown to have a strong negative effect on the probability to attend a private school. For example, Chiswick and Koutroumanes (1996) showed that a typical household is four times more likely to attend private schooling when it lives in a central city than when it lives in a rural area.

Share of African-Americans in the local population, according to previous empirical studies, is expected to have a positive impact on the private enrollment rate (Coltfelter, 1976; James, 1987; Hamilton and Macauley, 1991; and McCormick et al., 1994).

Share of Hispanics in the population. Sonstelie (1979) found that the proportion of Hispanics, after controlling for Catholic share, has a negative effect on the share of children who attend private schooling. On the other hand, using micro-level data, Chiswick and Koutroumanes (1996) showed that being Hispanic had no significant effect on the demand for private schooling, while Buddin et al. (1998) showed that it had a positive effect.

Following Cohen-Zada and Justman (2003) and Cohen-Zada (2006) who found all these variables to have a concave effect on the private enrollment rate, we allow for non-linearity by including squared terms for each variable.

***Percent of school-age population* (5-17)** is associated with the number of school-age children per household, which affects the cost of education quality relative to other spending categories. In theory, the price of education has two conflicting effects on the demand for private schooling. On the one hand, as the number of school-age children in the household increases, households have less money to send their children to private schools which decreases the demand for private schooling. On the other hand, for a given education budget, a larger number of school-age children per household implies lower quality public schools which increases the demand for private schooling as a substitute for public schooling. Thus, the direction of the effect of this variable cannot be determined a priori.

State and year fixed effects. State fixed effects are included in the regression in order to control for state-specific factors that may influence local private enrollment rates. For example, different states apply different state aid formulas that affect households' choice between public and private schools. Also, high average wages in a state may raise the cost of hiring teachers (Poterba, 1997), but this would have a stronger effect on public schools and other private schools than in Catholic schools, where nuns and priests, who are usually part of the teaching staff, are willing to work for low wages.[10]

---

[10] Our database does not include any measure of public school quality, and therefore we are unable to control for it in our regressions. We could include instead state measures of public school quality, such as pubic spending per student or teacher student ratio, but state factors are already captured in our state fixed effects. Furthermore, previous studies which used such state measures of public schools quality mostly obtained insignificant results or even opposite results than those expected (Gemmello and Osman 1984, James 1987, Long and Toma 1988, and Chiswick and Koutroumanes 1996).

A year fixed effect is included in order to capture all omitted variables that vary over time, and influence the demand for private schooling.

5. Results

5.1 Instrument relevance

The results of a simple pooled OLS regression of the private enrollment rate on the Catholic share in 1890 and its squared term are presented in Table 4, Column 1. As the observations vary greatly in size, we report *t*-statistics corrected for heteroscedasticity according to White (1980). We can see that the Catholic share in 1890 has a very significant positive concave effect on the private enrollment rate, explaining almost 13% of the variance in the private enrollment rate. In addition, the *F*-statistic on the excluded instruments is 248.8, which is well above the critical value required by Stock and Yogo (2005) for rejecting the null hypothesis that the instruments are weak according to the maximum Wald test size distortion definition.[11] This implies that in our regression weak instruments do not appear to be a concern.

In Column 2 we report results from a similar regression but now control for the set of demographic variables mentioned above. The results show that except for percent Hispanics squared, which is not significant, all the variables are very significant and with the predicted signs. The Catholic share in 1890 still has a very significant positive concave effect on the private enrollment rate, and the regression as a whole explains about 35% of the variance in the private enrollment rate. The *F*-statistic on the excluded instruments is 96, which again indicates that weak

[11] Stock and Yogo (2005) provide two definitions of weak instruments. However, as our model includes only one degree of over-identification, the only definition that can be applied is the one that is based on the maximum Wald test size distortion. Stock and Yogo (2005) calculate critical values for testing the hypothesis that the quality of the instruments is below one of four levels. In our case of two instruments (the Catholic share in 1890 and its squared term) and a single endogenous regressor (the private enrollment rate), the critical value required for rejecting the hypothesis that the quality of the instrument is below the highest level is 19.93.

instruments are not a concern. Mean income has an increasing concave effect on the share of private enrollment. The share of African-Americans in the population also has a positive concave effect on the private enrollment rate, suggesting that a larger African-American population increases the proportion of whites who choose private schooling. The density of population, which offers a greater advantage for private rather than public schooling, has an increasing concave effect on the private enrollment rate, and the share of population that lives in a rural area decrease the demand for private schooling. Finally, the share of Hispanics in the population has a negative effect on the demand for private schooling.

Stock and Yogo (2005) also define instruments as weak if the bias of 2SLS is greater than 5% of the bias of OLS. To test whether our instruments are not weak according to this definition, our model would need to have at least two degrees of over-identification. Thus, we run an additional regression, in which we add as additional instruments the Catholic share in 1906 and its squared term. According to this test, we reject the assumption of weak instruments if the joint *F*-statistic on the four excluded instruments is higher than (Stock and Yogo 2005, Table 1) 16.85. The results, presented in Column 3 of Table 4, indicate that both the Catholic share in 1890 and 1906 have significant positive concave effects on the private enrollment rate. In addition, the Stock and Yogo (2005) test indicates that the hypothesis of weak instruments is strongly rejected. As we provide more than one instrument, one is able to test whether in a particular dataset our instruments are exogenous to student outcomes.

To check the strength of the instruments in a linear specification, we run two additional regressions. In the first regression we use the Catholic share in 1890 as the only instrument and in the second we include both the Catholic shares in 1890 and in

1906. The results are reported in Columns 4 and 5 of Table 4. In both regressions the historical Catholic shares have very significant positive effects on the private enrollment rate, and the *F*-statistic on the excluded instruments exceeds the critical value that avoids weak instrument concern.

Next, we run a regression without the instruments but with the other explanatory variables, which allows us to evaluate how much extra explanatory power the instruments have over the other explanatory variables. The results, which are presented in Column 6 of Table 4, show that the regression explains 30.7% of the variance in the dependent variable, which implies that the Catholic share in 1890 and its squared term explain an extra 4% of the variance in the dependent variable (see Column 2).

Since private enrollment is zero in about 29% of the counties, we also estimated each specification using a pooled Tobit regression (see Table 5). The results are generally very similar to the OLS results. The variables are generally very significant and with the predicted signs, and the historical Catholic shares in 1890 and 1906 still have a very significant positive concave effect on the private enrollment rate. This result strengthens our argument that the Catholic share in 1890 and its squared term can serve as strong instruments of private school competition.

As mentioned in the data section, the reported share of Catholics may be biased up or down if churches from one county draw in Catholics from surrounding counties. This is more likely to occur in the few counties in which the reported share of Catholics is very high. In this case, if Catholics from surrounding counties are less likely than local Catholics to attend the local Catholics schools, the correlation between the share of Catholics in the population and the private enrollment rate may drop off and we would obtain a biased Catholic share effect. Moreover, because some of our specifications include the squared Catholic share, these few observations could have a large impact on the results. To exclude this possibility, we run two more regressions for each OLS and Tobit specification. In the first regression we exclude from the regression only the observations with Catholic shares higher than 100% (four observations in 1890 and two observations in 1906), and in the second regression we exclude also observations with Catholic shares higher than 60%. Table 6 reports for each regression the coefficients of the Catholic share variables and the *F*-statistic on the excluded instruments. We can see that the historical Catholic shares are still very strong instruments of private school competition, and that our results are not driven by the few observations with very high Catholic shares. Taken together, all our specifications strongly indicate that the Catholic share in 1890 and its squared term are strong instruments for the private enrollment rate.

As the Catholic share in 1890 is an historical variable dating back over a century before the period of interest, a century during which there have been large waves of Catholic immigration and large changes in local populations, any function of the Catholic share in 1890 is not likely to be correlated with current unmeasured student characteristics that affect individual student achievement, which implies that they can serve as valid instruments for private school competition.

5.2 Endogeneity of the current Catholic share

After we have shown that our instruments are valid, In this section we use them to test whether the maintained assumption that the current Catholic share is exogenous to the private enrollment rate is also valid. As previous studies on the effect of private school competition on student outcomes assumed either a linear (Sander 1996, Geller et al. 2006) or a quadratic (Hoxby 1994, Jepsen 2002, Green and Kang 2004) effect of the current Catholic share on private school competition, we test the exogeneity of the current Catholic share under both specifications. Our results indicate that the current Catholic share is endogenous to private school competition under both specifications. In addition, assuming that the current Catholic share is exogenous to private school competition yields a Catholic share effect that is biased down by more than 30%. This bias in the first-stage estimation of the private enrollment rate on the current Catholic share would generate a bias in the estimated effect of private school competition on student outcomes.

We start with providing simple OLS and Tobit results of the private enrollment rate on the current Catholic share and its squared term, controlling for those demographic variables which were found to be significant in previous studies. Columns 1 and 2 of Table 7 indicate that in both regressions the variables are generally very significant and with the predicted signs, and that the current Catholic share has a significant concave effect on the private enrollment rate.

To test the maintained assumption that the current Catholic share and its squared term are exogenous to the private enrollment rate, we use as their instruments the Catholic share in 1890 and the square of this variable. For the Catholic share in 1890 to be a legitimate instrument for the current Catholic share it must be correlated with the current Catholic share (quality condition) and also not have a direct effect on the private enrollment rate (validity condition). The results of the first stage estimations, presented in Columns 3 and 4 of Table 7, indicate that the Catholic share in 1890 is highly correlated with the current Catholic share. In addition, the calculated Cragg-Donald (1993) statistic in our IV estimation is 475, which easily passes the critical value proposed by Stock and Yogo (2005). [12] Thus, the Catholic share in 1890 and its

[12] Stock and Yogo (2005) tabulate critical values that enable the use of the Cragg-Donald (1993) statistic to test whether a set of instruments are weak in models with more than one endogenous variable. In this case, testing the hypothesis that the instruments are weak according to the *F*-statistic on the excluded instruments of each first stage regression may be misleading. The instruments can be

squared term satisfy the quality condition.

These historical data are unlikely to have a separate direct effect on the present private enrollment rate, unless there is an omitted factor that affects demand for private schooling which is fixed over more than one hundred years. Therefore, it is very likely that the Catholic share in 1890 and its squared term also satisfy the validity condition. Formally, we test for the validity of our instruments in the end of this section.

To determine whether the current Catholic share and its squared term are exogenous to private school enrollment we apply the Hausman test (1978). This test is important because if the maintained assumption that %Catholics is exogenous is valid, 2SLS would then yield less efficient estimates than OLS. The statistic for this test is computed using a two-stage procedure. In the first stage, we estimate reduced-form equations for the current Catholic share and its squared term by regressing them on all the exogenous variables of the model including the instruments.[13] In the second stage, we estimate the private enrollment rate, including the two residual terms estimated in the first stage. A significant *F*-test that the two residual terms together are different from zero rejects the exogeneity of percent Catholics and percent Catholics squared. The P-value of the *F*-test, displayed in Column 5 of Table 7, shows that the commonly-used assumption that percent Catholics and percent Catholics squared can be treated as exogenous is rejected at a high level of significance and thus instrumentation is necessary.

---

weak although they are very significant in each first stage regression. The reason for this is that when the predicted endogenous explanatory variables are close to collinear, it is difficult to separate their effects. For our case of two endogenous variables (the current Catholic share and its squared term) and two instruments, the critical value for rejecting the test that the quality of the instruments is below the highest level is 7.03. We computed the Cragg-Donald statistic using the Stata procedure IVREG2 set out by Baum et al. (2006).

[13] See Wooldridge (2002), p. 236, for a discussion on how to obtain IV estimates when the relationship between the dependent variable and the endogenous explanatory variable is non-linear.

Column 5 of Table 7 also presents the 2SLS estimates. All the variables are significant, and the regression explains about 35% of the variance in the private enrollment rate. The results also show that the marginal effect (at the mean) of the Catholic share is almost 40% lower under OLS than under 2SLS.

To test and control for the endogeneity of percent Catholics and percent Catholics squared under a Tobit regression, we use the two-stage procedure of Smith and Blundell (1986). The only difference between this procedure and the Hausman (1978) procedure is that in the former we estimate the second stage using a Tobit regression. The results, reported in Column 6 of Table 7, again indicate that the exogeneity of percent Catholics and percent Catholics squared is strongly rejected. Column 6 also reports the IV Tobit estimates, which are computed using the STATA procedure set out by Harkness (2000). In this procedure, the reported *t*-statistics take into account the pre-estimation of percent Catholics and percent Catholics squared, but fail to correct for heteroscedasticity. The results indicate that the marginal effect at the mean of the current Catholic share is about 32% lower under the regular Tobit estimation than under the IV Tobit one. Taken together, our results demonstrate that the current Catholic share in the local population is endogenous to the local private enrollment rate, and that the endogeneity bias associated with treating the current Catholic share as exogenous is more than 30%. This implies that the current Catholic share in the population cannot serve as a valid instrument for private school competition. However, this result should be interpreted with caution as in some of our specifications the quadratic relationship between the *current* Catholic share and the private enrollment rate appears to rest on the few observations with very high Catholic shares, as shown in Table 8. Although Columns 1 and 2 show that excluding the two observations with a Catholic share higher than 100% still yields a quadratic relationship between the current Catholic share and private school competition, excluding the 82 observations with a Catholic share higher than 60% implies that the quadratic term is significant only in the Tobit specification (Column 4). Moreover, excluding the two observations with a Catholic share higher than 100% and estimating the same equation by either 2SLS or IVTobit yields a non-significant quadratic term (Columns 5 and 6). Thus, we suggest that more weight should be given to testing the exogeneity of the current Catholic share in a linear specification.

The results of a simple OLS regression of the private enrollment rate on the current Catholic share and other demographic variables appear in Column 1 of Table 9. The Catholic share is found to be a very significant determinant of the private enrollment rate, and the regression as a whole explains about 36% of the variance in the private enrollment rate. We then compare the results of the OLS regression with four different 2SLS estimations, differing by the set of instruments used. The results of the four reduced form regressions are presented in Table 10, each of which explains more than two-thirds of the variance in the current Catholic share. In addition, all the instruments have a very significant effect on the current Catholic share, and the *F*-statistics on the excluded instruments are always much above any critical value suggested by Stock and Yogo (2005). This indicates that we can very easily reject the null hypothesis that the instruments are weak.[14]

The 2SLS results are reported in Columns (2)-(5) of Table 9. They show that the size of the Catholic share effect, which is very significant in all the

[14] As the first three reduced form specifications do not include at least two degrees of over-identification, we apply the Stock and Yogo (2005) test only under the 'maximum Wald test size distortion' definition. However, the fourth reduced form specification includes three degrees of over-identification, which allows us to implement the Stock and Yogo (2005) test also under the 'relative bias' definition. The *F*-statistic statistic on the excluded instruments in our regression equals 288.3, which indicates that the Stock and Yogo (2005) test rejects the null hypothesis that the instruments are weak under both definitions. That is, the bias in the 2SLS estimates are not greater than 5% of the bias in the OLS estimates, and the maximal size of 5 percent Wald test of the current Catholic share does not exceed 10%.

regressions, is very similar in all the specifications, and is about 50% larger than in the OLS regression. Stated differently, all the specifications imply that OLS yields a Catholic share effect that is biased down by approximately 33%, which is definitely a substantial bias. In addition, in all the specifications, the Hausman (1978) test indicates that the maintained assumption that the current Catholic share is exogenous is strongly rejected.

Next, we test whether the Catholic share in 1890 and its squared term are valid instruments for the current Catholic share, by testing the identifying restrictions. This is done only for the specifications for which the private enrollment rate equation is over-identified (Columns (3)-(5) of Table 9). The results show that in all the specifications, we cannot reject the joint null hypothesis that the instruments are valid, i.e., uncorrelated with the error term. Thus, while the assumption that the current Catholic share and its squared term are exogenous is strongly rejected, we could not reject the assumption that the Catholic share in 1890 and its squared term are exogenous.

To test and control for the endogeneity of the current Catholic share under a Tobit regression, we use Smith and Blundell's (1986) two-stage procedure (see Table 11). The results again indicate that in all the specifications, the commonly-used assumption that the current Catholic share is exogenous is strongly rejected. In addition, all the IV Tobit estimations yield a similar Catholic share effect, which is about 50% larger than in the regular Tobit regression.

Finally, we check the robustness of our results when all the observations with a Catholic share higher than 60% are excluded from the regression. That is, we estimate again the four specifications reported in Table 9, but only among the observations for which the Catholic share is lower than 60%. Table 12 reports the results, which are very similar to those that include all the observations. They still indicate that the current Catholic share is endogenous to private school competition, and that OLS yields a Catholic share effect that is biased down by approximately 33%. In addition, we still cannot reject the null hypothesis that the historical Catholic shares are valid instruments, i.e., uncorrelated with the error term.

6. Concluding remarks

This paper shows that while the current Catholic share is endogenous to the private enrollment rate, the Catholic shares in 1890 is exogenous. Also, we find that the Catholic share in 1890 and its squared term are very strong instruments for private school competition. These results were found to be robust to several specifications and estimations methods. Our instruments may also be relevant to studies that estimate the treatment effect of Catholic schools, which have used the current share of Catholics in the local population as an instrument for Catholic school attendance (Evans and Schwab, 1995; Sander 1996; Neal, 1997; Dee 2005, among others). Altonji et al. (2005a) recently argued that the current share of Catholics in the local population is not a valid instrument for Catholic school attendance as it influences student outcomes. As our instrument is more exogenous to student outcomes, and strongly correlates with Catholic school attendance, it can serve as a better instrument for this purpose as well.

References

[1] Altonji, J., Elder, T. and Taber, C., 2005a, An evaluation of instrumental variable strategies for estimating the effects of Catholic schools, *Journal of Human Resources* 40, 791-821.

[2] Altonji, J., Elder, T. and Taber, C., 2005b, Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools, *Journal of Political Economy* 113, 151-184.

[3] Andrews, D. and Stock, J., 2005, Inference wit weak instruments, NBER Technical Working Paper No. 313.

[4] Baum, C., Schaffer, M., and Stillman, S., 2006, Ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression.

[5] Belfield, C. and Levin H. 2002, The effects of competition on educational outcomes, *Review of Educational Research* 72, 279-341.

[6] Bound, J., Jaeger, D. and Baker, R., 1995, Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak, *Journal of the American Statistical Association* 90, 443-450.

[7] Buddin, R., Cordes, J. and Kirby, S. N., 1998, School choice in California: Who chooses private schools?, *Journal of Urban Economics* 44, 110-134.

[8] Chiswick, B. and Koutroumanes, S., 1996, An econometric analysis of the demand for private schooling, *Research in Labor Economics* 15, 209-237.

[9] Clotfelter, C., 1976, School desegregation, "tipping" and private school enrollment, *Journal of Human Resources* 11, 28-50.

[10] Cohen-Zada, D. and M. Justman, 2003, The political economy of school choice: Linking theory and evidence, *Journal of Urban Economics* 54, 277-308.

[11] Cohen-Zada, D., 2006, Preserving religious identity through education: Economic Analysis and evidence from the United States, *Journal of Urban Economics* 60, 372-398.

[12] Cragg, J. and Stephen D., 1993, Testing identifiably and specification in instrumental variables models, *Econometric Theory* 9, 222-240.

[13] Dee, T. S., 2005, The effect of Catholic schooling on civic participation, *International Tax and Public Finance* 12, 605-625.

[14] Dee, T. S., 1998, Competition and the quality of public schools, *Economics of Education Review* 17, 419-427.

[15] Evans, R. and R. Schwab, 1995, Finishing high school and starting college: Do Catholic schools make a difference? *Quarterly Journal of Economics* 110, 947-974.

[16] Finke R. and R. Stark, 2005, The Churching of America, 1776-2005: Winner and Losers in our Religious Economy, Rutgers University Press.

[17] Friedman, M., 1962. The role of government in education. In *Capitalism and Freedom*, Chapter VI. Chicago: University of Chicago Press.

[18] Geller, C., Sjoquist, D. and Walker, M, 2006, The effect of private school competition on public school performance in Georgia, *Public Finance Review* 34, 4-32.

[19] Gemmello, J and Osman J. W., 1984, Estimating the demand for private school enrollment, *American Journal of Education* 92, 262-279.

[20] Glenmary Research Center, Religious Congregations and Membership in the United States 2000 (Nashville, Tennessee).

[21] Green, K. and Byung-Goo, K., 2004, The effect of public and private competition on high school outputs in New York State, *Economics of Education Review* 23, 497-506.

[22] Grogger, J. and Neal, D., 2000, Further evidence on the effects of Catholic secondary schooling, Brookings-Wharton: Papers on Urban Affairs, 151-202.

[23] Hamilton, B. and M. Macauley, 1991, Determinants and consequences of the private-public school choice, *Journal of Urban Economics* 29, 282-294.

[24] Hahn, J. and Huasman, J., 2002, A new specification test for the validity of instrumental variables, *Econometrica* 70, 163-189.

[25] Harkness, J., 2000, Ivprob-Ivtobit: Stata modules to estimate instrumental variables probit and tobit, Boston College Department of Economics, Statistical software component, number s415801.

[26] Hausman, J., 1978, Specification tests in Econometrics, *Econometrica* 46, 1251-1271.

[27] Hausman, J., Stock, J. and Yogo, M., 2005, Asymptotic properties of the Hahn-Hausman test for weak instruments, *Economic Letters* 89, 333-342.

[28] Hoxby, C., 1994, Do private schools provide competition for public schools? NBER Working Paper No. 4978.

[29] James, E., 1987, The public-private division of responsibility for education: An international comparison, *Economics of Education Review* 6, 1-14.

[30] Jepsen, C., 2002, The role of aggregation in estimating the effects of private school competition on student achievement, *Journal of Urban Economics* 52, 477-500.

[31] Long, J. and Toma, E., 1988, The determinants of private school attendance, 1970-1980, *Review of Economics and Statistics* 70, 351-356.

[32] McCormick, R., Moore, C. and Yandle, B., 1994, Private and public choices in public education: An investigation of the trustee effect, *Public Choice* 78, 219-230.

[33] Neal, D., 1997, The effects of Catholic secondary schooling on educational achievement, *Journal of Labor Economics* 15, 98-123.

[34] Poterba, J., 1997, Demographic structure and the political economy of public education, *Journal of Policy Analysis and Management* 16(1), 48-66.

[35] Sander, W., 1999, Private schools and public schools achievement, *Journal of Human Resources* 34, 697-709.

[36] Sander, W., 1996, Catholic grade schools and Academic Achievement, *Journal of Human Resources* 31, 540-548.

[37] Sander, W. and Krautmann A., 1995, Catholic schools, dropout rates, and educational attainment, *Economic Inquiry* 33, 217-233.

[38] Sander, W., 2000, Parochial schools and student achievement: Findings for older adults, *Education Economics* 8, 259-268.

[39] Smith, R. and R. Blundell, 1986, An exogeneity test for a simultaneous tobit model with an application to labor supply, *Econometrica* 54, 679-686.

[40] Sonstelie, J., 1979, Public school quality and private school enrolment, *National Tax Journal* 132, 343-351.

[41] Sonstelie, J., 1982, The welfare cost of free public schools, *Journal of Political Economy* 90, 794-808.

[42] Staiger, D. and Stock, J., 1997, Instrumental variables regression with weak instruments, *Econometrica* 65, 557-586.

[43] Stark, R., 1992, The reliability of historical United States census data on religion, *Sociological Analysis* 52, 91-95.

[44] Stock, J. and Yogo, M., 2005, Testing for weak instruments in linear IV regressions. In Andrews, D. and Stock, J. (Eds.), *Identification and Inference for Econometric Models*, pp.80–108, Cambridge University Press, Cambridge, New York.

[45] U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Public Elementary/Secondary School Universe Survey, 1989-1990 and 1999-2000.

[46] U.S. Department of Education, National Center for Education Statistics, Common Core of Data, Private School Survey, 1989-1990 and 1999-2000.

[47] U.S. Bureau of the Census, 1890, Census of the population, Washington, DC, Available from the Geospatial and Statistical Data Center at the University of Virginia at http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html.

[48] U.S. Bureau of the Census, 1910, Census of the population, Washington, DC, Available from the Geospatial and Statistical Data Center at the University of Virginia at http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html.

[49] U.S. Bureau of the Census, Religious bodies, 1890, Part 1, Washington, DC. Available at http://216.122.145.46/AGGREGATE.asp.

[50] U.S. Bureau of the Census, Religious bodies, 1906, Part 1, Washington, DC. Available at http://216.122.145.46/AGGREGATE.asp.

[51] U.S. Census Bureau, County and City Data Book on CDROM (1994).

[52] U.S. Census Bureau, 2000, County and City Data Book.

[53] West, E. and H. Palsson, 1988, Parental choice of school characteristics: Estimation using state-wide data, *Economic Inquiry* 26, 725-740.

[54] White, H., 1980, A heteroscedasticity consistent covariance matrix estimator and a direct test for heteroscedasticity, *Econometrica* 48, 817-838.

[55] Wooldridge, J., 2002, Econometric Analysis of Cross Section and Panel Data, Cambridge, MIT Press.

Table 1. Correlation between the Catholic share in the local population in 2000 and at earlier points in time

Variable	Correlation
1990	0.89
1952	0.81
1936	0.80
1926	0.78
1916	0.74
1906	0.68
1890	0.65

Table 2. Number of observations in each 10% range in the distribution of the Catholic share in the population in 1890, 1906 and 1990-2000

	1890	1906	1990-2000
0%-10%	4,304	4,478	3,352
10%-20%	618	722	1305
20%-30%	236	280	709
30%-40%	64	170	376
40%-50%	20	62	174
50%-60%	20	24	129
60%-70%	4	18	48
70%-80%	10	8	28
80%-90%	2	6	14
90%-100%	0	2	6
>100%	2	4	1
Total	5,280	5,774	6,142

Table 3. Descriptive statistics

	Year	Mean	St. Dev	Min	Max
Private enrollment share, %	1990	5.33	7.30	0	79.24
	2000	5.54	6.41	0	52.23
% Catholics	1890	5.50	9.26	0	127.32
	1906	6.66	11.03	0	142.24
	1990	12.88	14.90	0	116.30
	2000	13.41	14.45	0	94.68
Mean income ($000s)	1990	30.14	7.20	15.96	71.38
	2000	44.73	10.28	24.62	105.21
% Hispanics	1990	3.52	9.41	0	97.2
	2000	5.08	10.19	0.1	97.5
Percent of population rural	1990	64.17	28.86	0	100
	2000	60.25	30.08	0	100
Percent of population 5-17	1990	19.56	2.44	11.2	33.5
	2000	19.03	2.12	10.5	29.6
% African-Americans	1990	8.96	14.63	0	86.3
	2000	8.93	14.61	0	86.5
Density (000s per square mile)	1990	0.17	0.90	0.00	32.40
	2000	0.19	0.95	0.00	34.92

Table 4. Instrument relevance

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)
	POOLED OLS	POOLED OLS	POOLED OLS	POOLED OLS	POOLED OLS	POOLED OLS
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	3.61 (32.07)	3.63 (1.06)	3.94 (1.13)	2.49 (0.71)	4.06 (1.16)	5.08 (1.47)
% Catholics 1890	0.43 (16.61)	0.30 (12.88)	0.15 (6.05)	0.15 (8.92)	0.06 (3.38)	
% Catholics 1890, squared	-4.7e-3 (-7.62)	-2.8e-3 (-8.18)	-1.6e-3 (-6.23)			
% Catholics 1906			0.16 (7.07)		0.12 (7.37)	
% Catholics 1906, squared			-1.0e-3 (-3.65)			
% Hispanics		-0.06 (-3.19)	-0.05 (-2.70)	-0.07 (-3.37)	-0.03 (-1.43)	-0.07 (-3.37)
% Hispanics, squared		1.6e-4 (0.51)	1.1e-4 (0.38)	2.9e-4 (0.93)	-3.7e-4 (-1.07)	8.1e-4 (2.76)
% Rural		-0.09 (-6.08)	-0.08 (-5.98)	-0.09 (-6.64)	-0.09 (-6.38)	-0.10 (-7.23)
% Rural, squared		3.0e-4 (2.72)	3.0e-4 (2.73)	3.3e-4 (3.04)	3.3e-4 (2.98)	3.8e-4 (3.41)
Density		1.29 (5.66)	1.24 (5.50)	1.41 (6.04)	1.34 (5.86)	1.68 (7.14)
Density, squared		-0.04 (-6.10)	-0.04 (-5.99)	-0.05 (-6.39)	-0.04 (-6.30)	-0.05 (-7.16)
Mean income		0.27 (7.15)	0.27 (7.17)	0.29 (7.51)	0.28 (7.32)	0.29 (7.23)
Mean income, squared		-1.9e-3 (-5.42)	-1.8e-3 (-5.36)	-2.0e-3 (-5.67)	-1.9e-3 (-5.47)	-2.0e-3 (-5.24)
% African-Americans		0.06 (2.74)	0.06 (2.88)	0.05 (2.59)	0.06 (2.83)	0.053 (2.50)
% African-Americans, squared		9.6e-4 (2.61)	9.4e-4 (2.57)	1.0e-3 (2.76)	9.7e-4 (2.63)	9.9e-4 (2.69)
Share of population at school- age		-0.86 (-2.45)	-0.88 (-2.43)	-0.60 (-1.65)	-0.77 (-2.14)	-0.70 (-2.01)
Share of population at school- age, squared		0.02 (2.37)	0.02 (2.34)	0.02 (1.56)	0.02 (2.04)	0.02 (1.97)
Number of observations	5280	5280	5254	5280	5254	5280
F-statistic on the excluded instruments	248.8	96.0	59.6	79.5	77.4	
$\overline{R}^2$ =	0.13	0.35	0.36	0.33	0.35	0.31

Table 5. Tobit results

(Huber/White corrected t statistics are in parentheses)

	(1)	(2)	(3)	(4)
	POOLED TOBIT	**POOLED TOBIT**	**POOLED TOBIT**	**POOLED TOBIT**
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	-2.49 (-0.50)	-1.73 (-0.34)	-3.95 (-0.77)	-1.52 (-0.30)
% Catholics 1890	0.37 (10.08)	0.18 (4.68)	0.18 (9.33)	0.07 (2.99)
% Catholics 1890, squared	-3.7e-3 (-4.86)	-2.0e-3 (-3.80)		
% Catholics 1906		0.22 (6.79)		0.14 (7.21)
% Catholics 1906, squared		-1.6e-3 (-3.56)		
% Hispanics	-0.10 (-3.24)	-0.09 (-3.06)	-0.11 (-3.46)	-0.06 (-1.86)
% Hispanics, squared	5.7e-4 (1.28)	6.5e-4 (1.52)	7.5e-4 (1.69)	-8.7e-5 (-0.17)
% Rural	-0.04 (-2.25)	-0.04 (-2.23)	-0.05 (-2.86)	-0.05 (-2.69)
% Rural, squared	-4.5e-4 (-3.01)	-4.3e-4 (-2.95)	-4.0e-4 (-2.70)	-3.9e-4 (-2.66)
Density	1.17 (4.73)	1.10 (4.51)	1.31 (5.19)	1.22 (4.98)
Density, squared	-0.04 (-5.36)	-0.04 (-5.19)	-0.04 (-5.74)	-0.04 (-5.61)
Mean income	0.41 (8.04)	0.41 (8.12)	0.43 (8.32)	0.42 (8.21)
Mean income, squared	-2.9e-3 (-6.61)	-2.9e-3 (-6.60)	-3.1e-3 (-6.79)	-3.0e-3 (-6.64)
% African-Americans	0.08 (2.93)	0.09 (3.11)	0.08 (2.79)	0.09 (3.05)
% African-Americans, squared	1.1e-3 (2.39)	1.1e-3 (2.31)	1.2e-3 (2.52)	1.1e-3 (2.37)
Share of population at school-age	-0.95 (-1.82)	-1.01 (-1.91)	-0.62 (-1.16)	-0.88 (-1.67)
Share of population at school-age, squared	0.03 (1.84)	0.03 (1.93)	0.02 (1.19)	0.02 (1.69)
Number of observations	5280	5254	5280	5254
Chi square statistic on the excluded instruments	185.7	251.2	87.03	159.3

Table 6. Sensitivity Analysis
(Huber/White corrected *t* statistics are in parentheses)

		(1)	(2)	(3)	(4)
Instruments		**Pcath1890 Pcath1890^2**	**Pcath1890 Pcath1890^2 Pcath1906 Pcath1906^2**	**Pcath1890**	**Pcath1890 Pcath1906**
OLS Catholic share <100	*% Catholics 1890*	0.33 (13.16)	0.17 (5.65)	0.17 (10.87)	0.07 (3.94)
	% Catholics 1890, squared	-3.6e-3 (-8.71)	-1.9e-3 (-4.21)		
	% Catholics 1906		0.17 (6.29)		0.13 (8.38)
	% Catholics 1906, squared		-1.3e-3 (-2.86)		
	F-statistic on the excluded instruments	*100.1*	*59.6*	*118.2*	*98.5*
OLS Catholic share <60	*% Catholics 1890*	0.34 (10.68)	0.19 (4.98)	0.20 (12.33)	0.08 (3.89)
	% Catholics 1890, squared	-3.9e-3 (-4.68)	-3.0e-3 (-2.81)		
	% Catholics 1906		0.16 (4.52)		0.14 (7.99)
	% Catholics 1906, squared		-7.7e-4 (-0.90)		
	F-statistic on the excluded instruments	*99.5*	*56.4*	*152.0*	*96.8*
Tobit Catholic share <100	*% Catholics 1890*	0.42 (13.18)	0.20 (5.00)	0.19 (10.77)	0.07 (3.08)
	% Catholics 1890, squared	-4.9e-3 (-8.17)	-2.5e-3 (-3.64)		
	% Catholics 1906		0.25 (6.93)		0.16 (8.18)
	% Catholics 1906, squared		-2.4e-3 (-4.06)		
	Chi^2-statistic on the excluded instruments	*206.0*	*255.5*	*116.1*	*203.1*
Tobit Catholic share <60	*% Catholics 1890*	0.45 (11.31)	0.24 (4.62)	0.23 (12.18)	0.08 (3.04)
	% Catholics 1890, squared	-6.1e-3 (-5.63)	-4.4e-3 (-2.90)		
	% Catholics 1906		0.26 (5.70)		0.18 (8.09)
	% Catholics 1906, squared		-2.5e-3 (-2.40)		
	Chi^2-statistic on the excluded instruments	*208.6*	*250.4*	*148.4*	*209.4*

Table 7. Regression Results

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)
	POOLED OLS	TOBIT	POOLED OLS	POOLED OLS	POOLED 2SLS	POOLED IVTOBIT
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	%CATHOLICS	%CATHOLICS SQUARED	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	4.58 (1.32)	-0.87 (-0.17)	-2.40 (-0.43)	35.12 (0.09)	4.76 (1.36)	-0.74 (-0.16)
% Catholics	0.25 (10.93)	0.34 (11.96)			0.43 (9.50)	0.52 (8.64)
% Catholics, squared	-1.4e-3 (-3.96)	-2.2e-3 (-5.36)			-3.1e-3 (-4.37)	-4.0e-3 (-3.95)
% Catholics 1890			0.99 (10.91)	40.86 (5.28)		
% Catholics 1890, squared			-6.5e-3 (-3.01)	5.1e-3 (0.03)		
% Hispanics	-0.13 (-6.17)	-0.19 (-5.85)	0.25 (4.07)	-4.98 (-1.01)	-0.18 (-7.13)	-0.24 (-5.93)
% Hispanics, squared	5.5e-4 (1.73)	1.1e-3 (2.47)	1.4e-3 (1.30)	0.38 (3.75)	7.5e-4 (1.98)	1.4e-3 (2.44)
% Rural	-0.09 (-6.28)	-0.04 (-2.28)	-0.03 (-1.18)	-0.30 (-0.18)	-0.07 (-5.37)	-0.03 (-1.47)
% Rural, squared	3.1e-4 (2.92)	-4.3e-4 (-2.96)	7.9e-5 (0.44)	1.8e-3 (0.14)	2.7e-4 (2.51)	-4.7e-4 (-3.23)
Density	1.37 (6.49)	1.24 (5.37)	0.51 (1.17)	48.48 (1.44)	1.22 (5.98)	1.10 (3.96)
Density, squared	-0.04 (-6.64)	-0.04 (-5.72)	-0.03 (-1.95)	-1.93 (-1.94)	-0.04 (-6.14)	-0.03 (-3.54)
Mean income	0.24 (5.94)	0.36 (6.66)	0.13 (2.17)	-4.34 (-1.07)	0.20 (4.72)	0.32 (5.78)
Mean income, squared	-1.8e-3 (-4.62)	-2.7e-3 (-5.57)	-4.1e-4 (-0.68)	0.05 (1.34)	-1.5e-3 (-3.93)	-2.5e-3 (-5.02)
% African-Americans	0.07 (3.59)	0.11 (3.81)	-0.09 (-4.56)	-2.73 (-2.09)	0.09 (4.21)	0.12 (4.60)
% African-Americans, squared	8.2e-4 (2.23)	9.0e-4 (1.93)	5.7e-4 (1.87)	-9.9e-3 (-0.47)	6.9e-4 (1.86)	7.6e-4 (1.84)
Share of population at school-age	-0.77 (-2.16)	-0.89 (-1.70)	-0.40 (-0.73)	-62.24 (-1.49)	-0.88 (-2.42)	-1.01 (-2.19)
Share of population at school-age, squared	0.02 (1.88)	0.02 (1.57)	0.02 (1.59)	2.51 (2.17)	0.02 (2.06)	0.02 (2.04)
Exogeneity test P-Value					0.0000	0.0000
F-statistic on the excluded instruments			258	119		
Number of observations	5280	5280	5280	5280	5280	5280
$\bar{R}^2$ =	0.37		0.67	0.53		

Table 8. Sensitivity Analysis

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)
	POOLED OLS	POOLED TOBIT	POOLED OLS	POOLED TOBIT	POOLED 2SLS	POOLED IVTOBIT
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	5.20 (1.49)	-0.53 (-0.10)	5.97 (1.69)	-0.18 (-0.04)	5.17 (1.45)	-0.63 (-0.13)
% Catholics	0.24 (9.89)	0.33 (10.79)	0.20 (6.71)	0.32 (8.37)	0.31 (3.16)	0.49 (3.92)
% Catholics, squared	-1.1e-3 (-2.96)	-2.2e-3 (-4.44)	-3.8e-4 (-0.63)	-2.0e-3 (-2.76)	-8.6e-4 (-0.48)	-3.5e-3 (-1.51)
% Hispanics	-0.13 (-5.97)	-0.19 (-5.80)	-0.12 (-5.73)	-0.18 (-5.52)	-0.15 (-3.74)	-0.23 (-4.27)
% Hispanics, squared	4.6e-4 (1.41)	1.1e-3 (2.35)	5.1e-4 (1.43)	1.1e-3 (2.14)	3.3e-5 (0.05)	1.2e-3 (1.36)
% Rural	-0.09 (-6.35)	-0.04 (-2.29)	-0.09 (-6.30)	-0.04 (-2.15)	-0.08 (-5.52)	-0.03 (-1.50)
% Rural, squared	3.2e-4 (2.96)	-4.3e-4 (-2.94)	3.1e-4 (2.90)	-4.4e-4 (-3.02)	2.9e-4 (2.66)	-4.6e-4 (-3.19)
Density	1.36 (6.48)	1.24 (5.38)	1.35 (6.47)	1.25 (5.47)	1.19 (5.67)	1.09 (3.90)
Density, squared	-0.04 (-6.64)	-0.04 (-5.73)	-0.04 (-6.66)	-0.04 (-5.83)	-0.04 (-5.92)	-0.03 (-3.51)
Mean income	0.24 (5.96)	0.36 (6.65)	0.25 (6.03)	0.36 (6.58)	0.23 (4.86)	0.32 (5.40)
Mean income, squared	-1.8e-3 (-4.64)	-2.7e-3 (-5.58)	-1.8e-3 (-4.68)	-2.7e-3 (-5.49)	-1.7e-3 (-4.15)	-2.5e-3 (-4.85)
% African-Americans	0.07 (3.59)	0.11 (3.81)	0.08 (3.74)	0.11 (3.96)	0.08 (3.93)	0.12 (4.50)
% African-Americans, squared	8.1e-4 (2.21)	8.9e-4 (1.92)	7.9e-4 (2.14)	8.5e-4 (1.82)	7.6e-4 (2.04)	7.8e-4 (1.86)
Share of population at school-age	-0.83 (-2.32)	-0.93 (-1.76)	-0.90 (-2.46)	-0.96 (-1.79)	-0.86 (-2.33)	-1.01 (-2.19)
Share of population at school-age, squared	0.02 (2.05)	0.02 (1.63)	0.02 (2.17)	0.02 (1.64)	0.02 (1.95)	0.02 (2.04)
Number of observations	5278	5278	5196	5196	5278	5278
$\overline{R}^2$ =	0.37		0.36		0.35	

Table 9. IV Estimates

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)
	POOLED OLS	**POOLED 2SLS**	**POOLED 2SLS**	**POOLED 2SLS**	**POOLED 2SLS**
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	5.21 (1.49)	5.17 (1.45)	5.17 (1.44)	5.24 (1.43)	5.23 (1.43)
% Catholics	0.17 (15.02)	0.26 (12.53)	0.26 (14.55)	0.25 (15.19)	0.25 (15.90)
% Hispanics	-0.11 (-5.23)	-0.13 (-5.62)	-0.13 (-5.65)	-0.12 (-5.43)	-0.12 (-5.42)
% Hispanics, squared	1.2e-4 (0.37)	-2.3e-4 (-0.62)	-2.5e-4 (-0.67)	-3.6e-4 (-0.97)	-3.8e-4 (-1.01)
% Rural	-0.09 (-6.53)	-0.08 (-5.89)	-0.08 (-5.86)	-0.08 (-5.81)	-0.08 (-5.79)
% Rural, squared	3.3e-4 (3.05)	3.0e-4 (2.75)	2.9e-4 (2.73)	2.9e-4 (2.67)	2.9e-4 (2.66)
Density	1.35 (6.40)	1.18 (5.66)	1.18 (5.67)	1.21 (5.83)	1.20 (5.80)
Density, squared	-0.04 (-6.59)	-0.04 (-5.91)	-0.04 (-5.92)	-0.04 (-6.03)	-0.04 (-6.01)
Mean income	0.26 (6.45)	0.24 (5.90)	0.24 (5.87)	0.24 (5.92)	0.24 (5.89)
Mean income, squared	-1.9e-3 (-4.96)	-1.8e-3 (-4.73)	-1.8e-3 (-4.72)	-1.8e-3 (-4.71)	-1.8e-3 (-4.69)
% African-Americans	0.07 (3.44)	0.08 (3.88)	0.08 (3.90)	0.08 (3.88)	0.08 (3.90)
% African-Americans, squared	8.6e-4 (2.35)	8.0e-4 (2.16)	8.0e-4 (2.16)	8.0e-4 (2.16)	7.9e-4 (2.15)
Share of population at school-age	-0.79 (-2.20)	-0.83 (-2.24)	-0.83 (-2.24)	-0.83 (-2.20)	-0.84 (-2.20)
Share of population at school-age, squared	0.02 (1.93)	0.02 (1.86)	0.02 (1.85)	0.02 (1.82)	0.02 (1.82)
Exogeneity test P-Value		0.00	0.00	0.00	0.00
Over-identification test P-Value			0.65	0.55	0.57
Instruments		Pcath1890	Pcath1890 Pcath1890sq	Pcath1890 Pcath1906	Pcath1890 Pcath1906 Pcath1890sq Pcath1906sq
Number of observations	5278	5278	5278	5248	5248
$\overline{R}^2$ =	0.36				

Table 10. Reduced form regressions
(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)
	POOLED OLS	POOLED OLS	POOLED OLS	POOLED OLS
Dependent variable	%CATHOLICS	%CATHOLICS	%CATHOLICS	%CATHOLICS
Constant	-4.72 (-0.91)	-8.88 (-1.72)	-3.77 (-0.71)	-6.37 (-1.20)
% Catholics 1890	0.65 (15.49)	1.25 (21.59)	0.23 (4.85)	0.53 (6.90)
% Catholics 1890, squared		-0.01 (-10.21)		-5.8e-3 (-3.58)
% Catholics 1906			0.56 (14.03)	0.73 (11.46)
% Catholics 1906, squared				-4.8e-3 (-3.89)
% Hispanics	0.24 (3.83)	0.26 (4.32)	0.30 (4.69)	0.29 (4.72)
% Hispanics, squared	1.7e-3 (1.52)	1.7e-3 (1.59)	1.4e-3 (1.12)	1.6e-3 (1.39)
% Rural	-0.05 (-2.04)	-6.8e-3 (-0.29)	-0.03 (-1.46)	-5.9e-3 (-0.26)
% Rural, squared	1.7e-4 (0.95)	-3.9e-5 (-0.22)	1.5e-4 (0.88)	-3.1e-8 (-0.00)
Density	0.77 (1.71)	0.40 (0.97)	0.45 (1.09)	0.22 (0.57)
Density, squared	-0.03 (-2.38)	-0.02 (-1.80)	-0.03 (-1.99)	-0.02 (-1.55)
Mean income	0.18 (2.97)	0.12 (1.99)	0.16 (2.72)	0.12 (1.97)
Mean income, squared	-7.7e-4 (-1.33)	-2.5e-4 (-0.41)	-4.2e-4 (-0.76)	-7.7e-5 (-0.13)
% African-Americans	-0.10 (-4.75)	-0.09 (-4.67)	-0.08 (-4.12)	-0.08 (-4.25)
% African-Americans, squared	7.0e-4 (2.27)	6.8e-4 (2.28)	6.1e-4 (2.05)	6.1e-4 (2.03)
Share of population at school-age	0.15 (0.28)	0.15 (0.28)	0.07 (0.12)	-1.9e-3 (-0.00)
Share of population at school-age, squared	8.9e-3 (0.63)	7.3e-3 (0.52)	0.01 (0.74)	0.01 (0.84)
F-statistic on the excluded instruments	239.8	336.8	305.4	288.3
Number of observations	5278	5278	5248	5248
$\overline{R}^2$ =	0.66	0.68	0.71	0.72

Table 11. IV Tobit estimates

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)
	POOLED TOBIT	**POOLED IV TOBIT**	**POOLED IV TOBIT**	**POOLED IV TOBIT**	**POOLED IV TOBIT**
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	-0.67 (-0.13)	-0.71 (-0.15)	-0.74 (-0.16)	-0.49 (-0.10)	-0.48 (-0.10)
% Catholics	0.20 (15.11)	0.29 (13.56)	0.30 (16.08)	0.29 (16.98)	0.30 (18.22)
% Hispanics	-0.15 (-4.78)	-0.17 (-4.69)	-0.17 (-4.77)	-0.16 (-4.37)	-0.16 (-4.42)
% Hispanics, squared	4.9e-4 (1.07)	1.7e-4 (0.35)	9.6e-5 (0.20)	2.1e-5 (0.04)	-3.3e-5 (-0.07)
% Rural	-0.05 (-2.65)	-0.04 (-2.04)	-0.04 (-1.96)	-0.04 (-1.98)	-0.04 (-1.91)
% Rural, squared	-4.0e-4 (-2.76)	-4.3e-4 (-2.97)	-4.4e-4 (-3.00)	-4.4e-4 (-3.01)	-4.4e-4 (-3.03)
Density	1.22 (5.28)	1.06 (3.80)	1.03 (3.71)	1.07 (3.87)	1.05 (3.78)
Density, squared	-0.04 (-5.67)	-0.03 (-3.45)	-0.03 (-3.36)	-0.03 (-3.48)	-0.03 (-3.40)
Mean income	0.39 (7.29)	0.37 (6.96)	0.37 (6.91)	0.37 (6.96)	0.37 (6.92)
Mean income, squared	-2.9e-3 (-6.02)	-2.8e-3 (-5.84)	-2.8e-3 (-5.82)	-2.8e-3 (-5.81)	-2.8e-3 (-5.79)
% African-Americans	0.10 (3.59)	0.11 (4.22)	0.11 (4.26)	0.11 (4.27)	0.11 (4.31)
% African-Americans, squared	9.9e-4 (2.12)	9.2e-4 (2.24)	9.1e-4 (2.22)	9.0e-4 (2.20)	8.9e-4 (2.18)
Share of population at school- age	-0.84 (-1.57)	-0.87 (-1.92)	-0.88 (-1.93)	-0.90 (-1.96)	-0.91 (-1.97)
Share of population at school- age, squared	0.02 (1.44)	0.02 (1.75)	0.02 (1.74)	0.02 (1.78)	0.02 (1.79)
Exogeneity test P-Value		0.00	0.00	0.00	0.00
Instruments		Pcath1890	Pcath1890 Pcath1890sq	Pcath1890 Pcath1906	Pcath1890 Pcath1906 Pcath1890sq Pcath1906sq
Number of observations	5278	5278	5278	5248	5248

Table 12. IV Estimates

(Huber/White corrected *t* statistics are in parentheses)

	(1)	(2)	(3)	(4)	(5)
	POOLED OLS	**POOLED 2SLS**	**POOLED 2SLS**	**POOLED 2SLS**	**POOLED 2SLS**
Dependent variable	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE	PRIVATE ENROLLMENT RATE
Constant	6.04 (1.71)	6.16 (1.72)	6.15 (1.72)	6.58 (1.81)	6.58 (1.81)
% Catholics	0.18 (14.71)	0.28 (11.91)	0.27 (13.94)	0.26 (13.82)	0.26 (14.58)
% Hispanics	-0.12 (-5.73)	-0.15 (-6.31)	-0.15 (-6.34)	-0.15 (-6.17)	-0.15 (-6.18)
% Hispanics, squared	4.7e-4 (1.33)	3.6e-4 (0.87)	3.7e-4 (0.90)	4.9e-4 (1.09)	5.0e-4 (1.10)
% Rural	-0.09 (-6.32)	-0.08 (-5.52)	-0.08 (-5.58)	-0.08 (-5.64)	-0.08 (-5.65)
% Rural, squared	3.1e-4 (2.91)	2.7e-4 (2.51)	2.7e-4 (2.54)	2.8e-4 (2.57)	2.8e-4 (2.57)
Density	1.34 (6.46)	1.16 (5.63)	1.17 (5.74)	1.19 (5.80)	1.19 (5.84)
Density, squared	-0.04 (-6.66)	-0.04 (-5.89)	-0.04 (-6.00)	-0.04 (-6.01)	-0.04 (-6.05)
Mean income	0.25 (6.28)	0.23 (5.66)	0.23 (5.69)	0.23 (5.65)	0.23 (5.65)
Mean income, squared	-1.8e-3 (-4.82)	-1.8e-3 (-4.56)	-1.8e-3 (-4.57)	-1.7e-3 (-4.49)	-1.7e-3 (-4.49)
% African-Americans	0.08 (3.73)	0.09 (4.25)	0.09 (4.24)	0.09 (4.26)	0.09 (4.25)
% African-Americans, squared	7.9e-4 (2.16)	6.9e-4 (1.87)	7.0e-4 (1.89)	6.9e-4 (1.88)	6.9e-4 (1.88)
Share of population at school- age	-0.89 (-2.45)	-0.96 (-2.59)	-0.96 (-2.59)	-0.98 (-2.58)	-0.98 (-2.58)
Share of population at school- age, squared	0.02 (2.16)	0.02 (2.22)	0.02 (2.22)	0.02 (2.20)	0.02 (2.20)
Exogeneity test P-Value		0.00	0.00	0.00	0.00
Over-identification test P-Value			0.62	0.42	0.86
Instruments		Pcath1890	Pcath1890 Pcath1890sq	Pcath1890 Pcath1906	Pcath1890 Pcath1906 Pcath1890sq Pcath1906sq
Number of observations	5196	5196	5196	5159	5159
$\overline{R}^2$ =	0.36				

APPENDIX: DATA SOURCES

Variable	Years	Source
Per capita money income	1989	County and City Data Book 1994
	1999	2000 Census SF3
Persons per household	1990	CCDB 1994
	2000	CCDB 2000
Density of population	1990	CCDB 1994
	2000	CCDB 2000
Percent of population at school-age (5 through 17)	1990	CCDB 1994
	2000	CCDB 2000
% Blacks in the population	1990	CCDB 1994
	2000	CCDB 2000
% Catholics in the population	1990 2000	Religious Congregations and Membership in the US 2000
Public enrollment	1989-1990	School and Agency Survey 1989-90
	1989-1990	School and Agency Survey 1989-90
Private enrollment	1989-1990	Private School Survey 1989-90
	1999-2000	Private School Survey 1999-2000
Catholic members	1890	United States Census of Religious Bodies 1890. Available at the American Religion Data Archive web-site http://216.122.145.46/AGGREGATE.asp
Catholic members	1906	United States Census of Religious Bodies 1906. Available at the American Religion Data Archive web-site (http://216.122.145.46/AGGREGATE.asp)
Population	1890	Decennial Census of the Unites States 1890. Available at the Geospatial and Statistical Data Center at the University of Virginia (http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html)
Population	1910	Decennial Census of the Unites States 1910. Available at the Geospatial and Statistical Data Center at the University of Virginia (http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html)