6-1996

Private school choice and the returns to private schooling

John Z. Smith Jr.
Vassar College, jzsmith@vassar.edu

Private School Choice and the Returns to Private Schooling*

Vassar College Working Paper No. 36
John Z. Smith Jr.

Completed: June 1996

Address Correspondence to:
Vassar College
Department of Economics
124 Raymond Ave., Box 708
Poughkeepsie, NY 12604-0708
914.437.7669 (voice)
914.437.7576 (fax)
jzsmith@vassar.edu

This paper analyzes the magnitude of earnings differentials and differences in the rates of return to education for individuals who attended private and public high schools in the U.S. from 1976 to 1983, controlling for self-selection into school sector, as well as observed differences in family background, ability, and characteristics of the high school attended. I develop a model of school choice, educational attainment, and earnings determination that allows both the returns to observable characteristics and potential selection biases to vary across private and public school attendees. If private high schools exhibit higher productivity in educating students, the rate of return to private high school attendance should exceed the rate of return to public high school attendance. Using data from the National Longitudinal Survey of Youth and data on private high school affiliations and tuition and fees at the county level, I estimate an endogenous switching regressions model of school choice and wage determination in which years of schooling is endogenous. I find that private school students enjoy significantly higher rates of return to education at all levels

---

*I have benefitted from the comments of Andrew Samwick, Caroline Minter Hoxby, Yacine Ait-Sahalia, Lee Lillard, Jim Hosek, James Dertouzos and seminar participants at Colgate, Colorado, Mathematica Policy Research, MIT, RAND and William and Mary. This research was supported by a grant from the Bradley foundation. Any and all remaining errors are mine.

Private School Choice and the Returns to Private Schooling

"Pencey Prep is this school that's in Agerstown, Pennsylvania. You probably heard of it. You've probably seen the ads anyway. They advertise in about a thousand magazines, always showing some hot-shot guy on a horse jumping over a fence. Like as if all you ever did at Pencey was play polo all the time. I never even once saw a horse anywhere *near* the place. And underneath the guy on the horse's picture, it always says: 'Since 1888 we have been molding boys into splendid, clear-thinking young men.' Strictly for the birds. They don't do any damn more *molding* at Pencey than they do at any other school. And I didn't know anybody there that was splendid and clear-thinking and all. Maybe two guys. If that many. And they probably *came* to Pencey that way."

Holden Caulfield
The Catcher in the Rye
— J.D. Salinger

..."The Dairy School was not a bastion of secondary school education.

It might have once wished for a status equal to Exeter's or Andover's, but it had settled, in the early 1900s, for a future of compromise. Near to Boston, it admitted a few hundred boys who had been turned down by Exeter and Andover, and a hundred more who shouldn't have been admitted anywhere, and it gave them a curriculum that was standard and wise — and more rigorous than most of the faculty who were employed to teach there; most of them had been turned down elsewhere, too. But, even second-rate among New England prep schools, it was far better than the area public schools and especially better than the only high school in the town of Dairy.

...My mother, who (being a girl) could never go to the Dairy school, attended the private female seminary also in town. This was another second-rate education that was nonetheless an improvement over the public high school, and the only choice of the town's parents who wished their daughters to be educated without the presence of boys."

The Hotel New Hampshire
— John Irving

"As a selection and training place of the upper classes, both old and new, the private school is a unifying influence, a force for the nationalization of the upper classes. The less important the pedigreed family becomes in the careful transmission of moral and cultural traits, the more important the private school. The school — rather than the upper-class family — is the most important agency for transmitting the traditions of the upper social classes, and regulating the admission of new wealth and talent. It is the characterizing point in the upper-class experience. In the top fifteen or twenty such schools, if anywhere, one finds a prime organizing center of the national upper social classes. For in these private schools for adolescents, the religious and family and educational tasks of the upper social classes are fused, and in them the major tasks of upholding such standards as prevail in these classes are centered."

The Power Elite
— C. Wright Mills

1. Introduction

One of the major wedge issues in current American policy debates is that of school choice. The school choice debate focuses on whether public provision of education must also entail public production of education. Advocates of school choice cite the U.S. secondary school system as an example of the differential efficacy of public versus private production of educational services. School choice advocates — as well as supporters of the private schools — point to higher achievement test scores, lower dropout rates, and higher college attendance rates for private school students as evidence that private schools deliver a "better education" to students.[1] Critics of school choice policies argue that private schools enroll a decidedly non-representative sample of students through the use of selective admissions policies and the charging of substantial tuition rates.[2] Hence, it is argued that private schools educate students from relatively advantaged backgrounds rather than a representative cross-section of high school-aged youth.

The crux of the debate centers on whether private schools are more effective in educating students, or are simply educating students of higher average quality. Both sides to the debate cite facts that are in agreement with most data collected on the backgrounds and characteristics of private and public school students.[3] As a result, any test of these competing views of private high schools

---

[1]See Coons and Sugarman [1978], Coleman, Hoffer, and Kilgore [1982], Levy [1986], Coleman and Hoffer [1987], Kearns and Doyle [1988], Nathan [1989], and Chubb and Moe [1990] for arguments in favor of school choice and the advantages of private production of education.

[2]These issues are addressed in Coleman and Hoffer [1987], Chapter 2.

[3]See Tables 2,3, and 4, and Coleman, Hoffer, and Kilgore [1982].

must control for observable differences in family background, student ability, and measures of school quality in gauging the effectiveness of private high schools *vis-a-vis* public high schools.

From a policy perspective, it is important to identify separately the effect of student background and school quality when measuring the effect of private school attendance on earnings. If the private-public differential in the rate of return to education is explained entirely by differences in students' family background, then education policies designed to expand parental choice to the private sector may not generate significant welfare improvements to society as a whole.

Further, if the differential return to education across school sectors is attributable to measured differences in school quality, policies designed to lower pupil-teacher ratios in the public schools may allow the government to replicate the private school "effect" without implementing school choice policies such as tuition vouchers.

If the differential return to education across school sectors is not explained by observable differences in the quality of students or school quality measures, then tuition voucher policies may allow for a welfare-improving reallocation of students across school sectors; specifically, an inflow of students into the private schools. The higher return to education realized by individuals switching to the private school sector would increase these individuals' welfare, and, in turn, societal welfare, given the external benefits of education to society as a whole.[4]

Any investigation of the effect of private schooling, however, is complicated by the fact that students are not assigned randomly to private and public schools. Rather, parents choose to send their children to private or public schools, based on the costs of attendance and the perceived quality of schooling in both sectors. As a result, there is likely to be endogenous selection into private and public schooling. Unless one controls for the endogeneity of school choice, any results concerning relative earnings and rates of return to education will be biased.

Moreover, the choice of school sector might represent differences in desired years of schooling across individuals. If unobservable characteristics that increase one's propensity to attend private school also augur for increased educational attainment, there exists potential selection bias in the schooling equations. I test for the presence of selection bias in the schooling equations, and also treat educational attainment as an endogenous variable in the log-wage equations.

Finally, one needs to define a yardstick by which "success" will be measured. Much of the existing literature has focused on the comparison of achievement test scores and gains in achievement test scores across public and private school students (e.g. so-called achievement differentials) in measuring the differential productivity or value-added from private schooling. Given the ambiguity as to precisely what standardized "ability" and "achievement" tests actually measure, a preferable metric for evaluating the productivity of private and public high schools is the rate of return to education estimated from canonical log-wage equations. If private high schools are more productive in educating students, then one would expect to see a significantly higher return to a year of private

---

[4]This paper does not assert that the only circumstance under which one would advocate a voucher system is the finding of a differential return to education for private school attendees. There are many reasons to advocate a school voucher program (e.g. a shift toward school-based management and decentralized decision making within the public schools), only one of which is the labor market returns to private and public school attendance.

school education than to a year of education received in a public high school, controlling for self-selection and differences in observables.

I estimate a model of private school choice and earnings that allows both the returns to observable characteristics, such as years of schooling, and the nature of any potential selection bias to vary across private and public school students. The rate of return to a year of schooling is significantly higher for private school students (0.131) than for public school students (0.066).[5] Finally, the average public school student, given the treatment of a private high school education, would attain an additional 1.457 years of schooling and enjoy a large, but not significant, increase in log-wages (+0.134). Roughly seventy percent of this wage differential is attributable to the combination of additional educational attainment and the higher return to schooling provided by the private school treatment effect.

Including controls for achieved ability does not alter the differences in the returns to schooling across school sectors significantly. The inclusion of the age-adjusted AFQT score results in the return to a year of private schooling increasing from 0.131 to 0.145 and the return to a year of public schooling declining from 0.066 to 0.058. Hence, the inclusion of the AFQT score does not diminish the private-public differential in the return to schooling. The higher return to education for private school students is quite robust across model specifications.

This paper is organized as follows. The returns to education literature — including studies that have measured the role of school quality and private school attendance on wage determination — is reviewed in Section 2. Section 3 develops models of school choice, educational attainment, and earnings determination. Section 4 describes the data used in this study. Empirical results on school choice, educational attainment, log-wages, log-wage differences, and returns to education are presented in Section 5. Section 6 concludes, and discusses directions for future research.

2. Returns to Education and School Quality Literature

Students who attend private high school come from families of higher socioeconomic status, as evidenced by the summary statistics reported in Table 2. Parents of private school children have mean annual incomes that are $6700 higher than the mean income of public school students.[6] Parents of private school children also have higher educational attainment, with the private-public differential on father's years of schooling almost two years. Further, there is some evidence that parents of private school children are more likely to work in professional or managerial occupations, particularly for mothers (see Table 2).

---

[5]Table 13 reports regression results in which years of secondary and post-secondary education are entered separately in the log-wage equations. The coefficients on years of secondary and post-secondary education are not significantly different for either the private school students or the public school students. (e.g. Education may be treated as a linear function in the wage equation. This finding is supported by the results of Card and Krueger [1992].)

[6]For a detailed description of the construction of the net family income measure, see Data Appendix A. Median incomes for the sample were: $20,383 for the entire sample, $29,000 for families with private school children, and $20,000 for families of public school children.

Since the work of Behrman *et al.* [1980], Taubman [1975], and Taubman and Wales [1974] suggests that the sum of genetics and family environment accounts for about one-half of the variance in earnings, and three-fourths of the variance in schooling, it is crucial to include measures of family background when estimating the labor market return to private and public school education.[7] I control for parental income, parents' years of schooling and occupational status, two-parent family, family size, and birth-order in estimating both the school choice and log-wage equations.[8]

Controlling for innate ability is more difficult. To some extent, there is uncertainty as to precisely what IQ test scores measure, and how this measured ability is related to labor market "success."[9] Griliches [1977] provides an excellent discussion of the problems associated with interpreting measured ability, noting that: "ability, in the sense of being able to earn higher wages, has little to do with IQ. It is an unobserved latent variable that both drives people to get relatively more schooling and earn more income, given schooling, and perhaps also enables and motivates people to score better on various tests."[10]

An additional criticism of the use of standardized test scores — particularly standardized achievement tests — as measures of ability arises in that these tests are measures of achieved ability (sometimes referred to as late ability) rather than innate ability. As such, test score differentials may reflect differential opportunities afforded the respondent as a result of family socioeconomic status rather than true ability differences. This problem is particularly worrisome in the NLSY data I employ in this study, since the respondents in the sample were 15-20 years old when they took the battery of tests contained in the Armed Services Vocational Aptitude Battery (ASVAB). Thus, all of the respondents had received at least one year of private high school or public high school (many had received four) by the time they were administered the ASVAB in the summer and fall of 1980. If private schooling and public schooling differ in their value-added, then a portion of the observed ability differential may be attributable to the productivity differential of private vs. public schooling.[11]

---

[7]Behrman *et al.* [1980], p. 194.

[8]Although it is possible to difference out common family environment using information on siblings in the NLSY (by matching the household identifier), only 60% of the respondents in my selected sample had siblings in the NLSY. Since the estimation of the fixed effects model would have resulted in a further reduction in sample size, I did not pursue this estimation strategy.

[9]Bowles and Gintis [1976] argue that IQ scores reflect differences in socioeconomic status rather than differences in innate ability.

[10]Griliches [1977], p. 7.

[11]To obtain some measure of both the presence and magnitude of differential value-added in public vs. private schooling, I defined a measure of ability that normed one's raw AFQT score by the average score in one's age cohort. For example, a 15 year-old taking the AFQT test receives a raw score, which I then divide by the average AFQT score of all 15 year-olds in the NLSY (not just those in the selected sample). A crude test of the differential productivity of private and public schooling can be constructed by looking at the age-adjusted AFQT scores, and the private-public school differential, for both the 15 year-old cohort and the 18 year-old cohort in 1980. The 15 year-olds had received, at most, one year of private or public schooling by 1980, whereas the 18 year-olds had likely received three or four years of private school education, conditional upon remaining in school.

Finally, private school students, independent of any true ability differential, may be "test-smart", since over 53% of all private high schools require an entrance examination as a criterion for admission.[12,13]

Given that most private schools practice selective admissions that are based largely on academic achievement, one might expect that omission of a measure of ability would induce a positive bias in the estimated returns to schooling coefficient in the private school log-wage equation, and thereby overestimate the "treatment effect" of private schooling as well. The commonly-held view that ability produces a substantial upward bias in the rate of return to education, however, hinges critically on the least-squares log-wage equation as the correct specification and the absence of measurement error, particularly in variables such as schooling. Griliches [1977] illustrates that both the sign and the magnitude of ability bias are highly sensitive to correct specification of the earnings function and no measurement error. Instrumenting for both years of schooling and test scores, Griliches [1977] finds that the "implied net ability bias is negative or nil."[14]

2.1 The Role of School Quality

An alternative explanation for the differential effectiveness of private vs. public schools is rooted in differences in the quality of private and public high schools. Some measures of school

Mean Age-Adjusted AFQT Scores

Age-cohort, 1980	Entire Sample	Private	Public	Pvt.-Pub. Differential
15 year-olds	1.177	1.282	1.168	0.114
	(0.298)	(0.338)	(0.293)	(0.447)
18 year-olds	1.167	1.261	1.163	0.098
	(0.289)	(0.192)	(0.292)	(0.349)

This cut of the data does not support the view that private schools are more productive in educating students, as the differential does not change significantly across age cohorts. Moreover, I did not remove those who had dropped out of high school from the 18 year-old cohort, which likely generates a downward bias in the mean age-adjusted AFQT score for the public school attendees.

[12]Benson and McMillen [1990], p. 79. Further, if one looks at the standard private school admissions tests, such as the SSAT, ISEE, COOP, and STS, they are strikingly similar to the achievement test proxy in the ASVAB, the AFQT.

[13]One might also be concerned that the presence of selective admissions induces a bias in the average abilities of private and public school students. To test for the presence of ability-based selection into the private schools, I included the age-adjusted AFQT score in the private school choice probit equation. The test score variable entered with a positive sign, but the T-statistic for the estimated coefficient on the age-adjusted AFQT score was +0.589.

[14]Griliches [1977], p. 18.

quality, such as type of curriculum offered, percentage of teachers with advanced degrees, and teacher salaries, are directly measurable. Measures such as discipline, freedom from bureaucracy, and parental support, however, are considerably more difficult to identify, and are rather unwieldy to quantify. The discussion below focuses primarily on the "standard suspects" used to define school quality, although intangible factors — such as the benefit of a value consensus among the constituents of a private religious school — are potentially important.[15]

The primary findings concerning school quality measures are obtained from Coleman [1966] and Hanushek [1986]. Coleman [1966] found that, controlling for the student's own background and the background of other students in the school, measures of school quality mattered little for the performance of students on standardized tests.[16] Hanushek [1986] reviewed a number of studies that estimated education production functions, where output was typically defined as student achievement or average school achievement on standardized tests and inputs to the educational production process were teacher salaries, pupil-teacher ratios, and per pupil expenditures. He found that these measures of school quality were generally insignificant.[17] Moreover, he found that when these measures were significant in the educational production function, they frequently entered the equation with the "wrong" sign.[18] Hanushek concludes that most available measures of school quality and teacher quality are poor indicators of the true effects of schools.

Altonji [1988] provides arguably the most careful and thoroughgoing analysis of the role of school quality, community characteristics, and family background on earnings and educational attainment, using the National Longitudinal Survey of the High School Class of 1972 (NLS72). Since the NLS72 was a stratified random sample of high schools, there are multiple individuals who attended the same high school in the sample. Altonji [1988] finds that 12.4% of the variance in permanent wage rates is across high schools.[19] He concludes that "high school and community factors that are common to students who attend the same high school play a relatively minor role in differences in labor market success."[20]

---

[15]Coleman and Hoffer [1987], Chapter 8, discuss the role of social capital, both within the family and the local community, in strengthening the effectiveness of religious schools, particularly Catholic schools.

[16]Coleman [1966], Chapter 1.

[17]Hanushek [1986], p. 1160.

[18]Hanushek [1986], p. 1160.

[19]Altonji [1988], Chapter 1.

[20]Altonji [1988], Chapter 1, pp. 9-11.

Altonji [1988] also analyzes the effect of family background and school characteristics on the rate of return to education. His principal findings are as follows:

1.) Controlling for family background leads to a substantial reduction in the estimates of the rate of return to education. The rate of return to a year of post-secondary education falls from 0.073 to 0.058 when one controls for family background.[21]

2.) Controlling for observed and unobserved high school characteristics has only a modest effect on estimates of the rate of return to education, lowering the rate of return by five-tenths of a percentage point.[22]

Based on Altonji's findings from the NLS72 data, controls for differences in school quality have little effect on parameter estimates in the log-wage equation, particularly the rate of return to education.[23] Since the NLSY data used in this paper does not contain a high school identifier that would allow one to conduct the fixed-effects estimation used by Altonji, it is somewhat reassuring to have external evidence on the magnitude of any resulting bias.[24,25] Instead, I include a parsimonious set of variables that are reasonable proxies for school quality, namely: pupil-teacher ratios, school enrollment, and sociodemographic characteristics of the aggregate school enrollment.

2.2 Studies of the Return to Private Schooling

Three previous studies have considered the effect of private school attendance on earnings. The first two studies, Taubman [1975] and Altonji [1988], did not focus specifically on the effect of private schooling on earnings. Chipty and Kessler [1992] investigate the effect of private schooling on earnings, controlling for the endogeneity of school choice. They estimate a log-wage equation that does not include years of education. Below, I review the findings of these studies.

---

[21]Altonji [1988], Chapter 2, p. 4. Further, Altonji finds this result to be invariant to whether the individual was enrolled in a college preparatory (academic) curriculum.

[22]Altonji [1988], Chapter 2, pp. 4-5.

[23]Altonji [1988], Chapter 2, p. 5.

[24]Card and Krueger [1992] use state-level information on measures of public school quality, such as average teacher salaries, average pupil-teacher ratios, and average per pupil expenditures to measure the effect of school quality on earnings and the rate of return to education using male cohorts from the 1980 Census. They find that: a 10% increase in teachers' relative pay is associated with a 0.1 percentage point increase in the rate of return to a year of education; and
a decline in the pupil teacher ratio from 25 to 20 is estimated to raise the rate of return to education by 0.4 percentage points.

[25]Earlier studies of the effect of school quality on earnings are Johnson and Stafford [1973] and Wachtel [1976]. Wachtel's study uses the Thorndike-Hagen data and provides a thorough analysis of educational quality by using data on per pupil expenditures for the high school and college the respondent attended. Unfortunately, Wachtel excludes from the sample those who attended private high schools.

Using the Thorndike-Hagen data, Taubman [1975] finds that individuals who attended private high schools earned 25% more in 1955 than those who went exclusively to public schools. By 1969, those who attended private high schools earned 29% more than those who had attended only public schools. The indicator for attending private high school entered the log-annual earnings equation with a positive and significant sign in both years.[26] Using the NLS72 data, Altonji [1988] regressed the high school-specific mean of the log-wage, adjusted for background, aptitude, and high school characteristics, on a set of regressors normally used in log-wage equations. Altonji found that private high school students earned a five to six percent wage premium, although this premium was not significant once controls for ability and the type of high school curriculum in which one was enrolled were included.[27] While the studies cited above suggest that private school attendees earn more than public school attendees, none control for the possibility of selection bias.

Chipty and Kessler [1992] use the NLSY data to explore the effect of private schooling on earnings. Using all individuals who reported wages from 1982-1986, they find virtually no evidence of selection bias in log-wage equations for private and public school attendees.[28] Further, they find that the average public school student would experience no significant change in earnings if she received a private school education.[29]

In the next section, I develop an econometric model of school sector choice and earnings, in which returns to all observables are allowed to vary across private and public school attendees. Further, I test for both the presence and the nature of selection biases in log-wage and log-wage difference equations estimated separately for private and public school attendees. The endogenous switching regressions model outlined below allows the rates of return to education to vary across

---

[26]Taubman made no attempt to control for the endogeneity of private school attendance, and acknowledges this point. Taubman [1975], p. 44, footnote 27. Moreover, the private school variable in the Thorndike-Hagen data set excluded those who attended religious private schools. Of the men in the Thorndike-Hagen data set, 99 attended private high schools.

[27]Altonji [1988], Chapter 1, Table 1.10. Again, Altonji notes that "this finding should be viewed cautiously given the likelihood of unobserved differences in the backgrounds of students who attend public and private schools." Chapter 1, pp. 9-10.

[28]Chipty and Kessler [1992], p. 22.

[29]Chipty and Kessler [1992] report that they identify their system of school choice and log-wage equations by the exclusion of the Catholic faith indicator from the log-wage equations which contain an indicator variable if the respondent reported attending religious services at least once per week. Chipty and Kessler [1992], pp. 21-22. Previous work, however, has found that religion is significant in the wage equation. Taubman [1975] finds that Catholics earn three to nine percent more than Protestants. (p. 39) Further, he found that Catholics had five percent higher growth rates in earnings. (pp. 131-133) Chipty and Kessler are also mistaken in their interpretation of the frequency of religious attendance question asked in the 1979 NLSY. The question asked respondents the frequency of religious attendance in the past calendar year, not at age 14. Moreover, use of the frequency of religious attendance measure is likely to be highly correlated with private school attendance, as 52% of Catholic high schools require weekly attendance at Mass, and 83% require attendance at Mass one to three times a month. Yeager *et al.* [1985], p. 209. Townsend [1989] reports that slightly over half of all boarding schools require regular attendance at chapel services.

private and public school attendees. Moreover, allowing educational attainment to be endogenous controls for differences in years of schooling completed across school sector.[30]

3. Empirical Methodology

3.1 Model of School Choice

Below, I develop an index function model of school choice. Private school attendance is determined by the utility a child and her or his parents derive from attending a private high school.[31] While these utility levels are not observed, the level of utility (I_i^*) can be modeled as a function of observable characteristics, contained in the vector Z_i, and an error term u_i:

$$I_i^* = Z_i'\gamma - u_i, \quad where \quad u_i \sim N(0,1) \tag{1}$$

Hence, an individual attends private school if:

$$I_i^* \geq 0 \quad \rightarrow \quad Z_i'\gamma \geq u_i. \tag{2}$$

Since we observe whether an individual attends private high school, we can construct an indicator I_i, which is a function of the latent index I_i^*, as follows:

$$I_i = 1 \quad if \quad I_i^* \geq 0 \quad (Z_i'\gamma \geq u_i), \tag{3}$$

$$I_i = 0 \quad if \quad I_i^* < 0 \quad (Z_i'\gamma < u_i). \tag{4}$$

---

[30]Willis and Rosen [1978] develop a switching regressions framework to test for selection biases in the earnings of high school and college graduates, using the Thorndike-Hagen data. For late earnings (1969), they find no selection bias for college graduates, but positive selection bias for those with high school educations. Willis and Rosen [1978], S27-S31. Kenny, Lee, Maddala, and Trost [1979] estimate an endogenous switching model for the earnings of high school and college graduates using the Project Talent data. The criterion function specified is a tobit function for desired years of post-secondary education. Kenny *et al.* [1979] also find no evidence of selection bias for the earnings of college graduates, but positive selection bias for those with only a high school education.

[31]Note that the above model is based only on utility-maximizing behavior. A stronger form of the school choice model, based on maximizing the present value of lifetime income, would define choice of school sector to be based on the present value of earnings streams under private and public school attendance, with individuals choosing the school sector that provided the highest present value of life-cycle income.

Given that we observe the private school enrollment decision and that we have imposed the condition that the u_is have unit variance, we can recover parameter estimates of the γs from probit maximum likelihood estimation of:

$$L = \prod_{i=1}^{N} [\Phi(Z_i'\gamma)]^{I_i} \cdot [1 - \Phi(Z_i'\gamma)]^{1 - I_i} . \quad (5)$$

3.2 School Choice and Educational Attainment

Given that an individual's choice of private or public school may, in part, be influenced by *ex ante* differences in desired years of schooling, the private school enrollment decision may be correlated with educational attainment. Hence, I estimate separate years of schooling equations for private and public school attendees, in which I test for selection bias. Further, even if one does not find self-selection in the educational attainment equations, it is unlikely that years of schooling is truly an exogenous variable in the log-wage equations. Using number of siblings in the family, an indicator variable for growing up in a two-parent household, and an indicator for having access to reading materials at home at age fourteen as instruments, I estimate second-stage log-wage equations in which years of schooling are treated as endogenous.

I specify years of education attained as S_i, and use subscripts N to denote private school attendees and P to denote public school attendees. The vector J_i comprises family background measures and school quality measures for the high school attended by the respondent. Below, I list the separate schooling equations for private and public school attendees, in which I test for endogenous selection in educational attainment.

Private School Attendees:

$$S_{Ni} = J_{Ni}'\zeta_N + \sigma_{N_S u}\frac{-\phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)} + \xi_{N_S i} . \quad (6)$$

Public School Attendees:

$$S_{Pi} = J_{Pi}'\zeta_P + \sigma_{P_S u}\frac{\phi(Z_i'\gamma)}{[1 - \Phi(Z_i'\gamma)]} + \xi_{P_S i} . \quad (7)$$

If the Heckman specification tests for endogenous selection reveal no significant selection biases in the educational attainment of private and public school attendees, one can estimate linear schooling equations for each school sector.

Private School Attendees:

$$S_{Ni} = J_{Ni}^{\prime}\zeta_N + \epsilon_{N_S i} \,. \tag{8}$$

Public School Attendees:

$$S_{Pi} = J_{Pi}^{\prime}\zeta_P + \epsilon_{P_S i} \,. \tag{9}$$

3.3 The Effect of Private Schooling on Earnings and the Returns to Education

The simplest method of estimating the effect of private high school attendance on earnings is the inclusion of the indicator variable I_i in the log-wage equation:

$$\ln[w_i] = X_i^{\prime}\beta + \delta I_i + u_{wi} \,, \tag{10}$$

where the u_{wi} are assumed to be distributed *iid*, and the X_i vector includes observables and educational attainment. In the specification of Equation (10), the effect of private schooling serves only to shift the intercept, as the returns to schooling and other observables are constrained to be equal across private and public school students. Moreover, the dummy-variable specification assumes that selection into private schools is exogenous; that there is zero correlation of the disturbance terms in the school choice and log-wage equations.[32] If the correlation of the disturbance terms in the choice and earnings equations is non-zero, then Equation (10) above will produce biased estimates of the return to schooling.

Extension of the simple OLS-indicator variable model of the effect of private schooling may take two forms: allowing interactions between the private school choice indicator and observed characteristics, to allow the returns to observables to vary across private and public school students; or testing for selection bias by incorporating selection-correction terms in the log-wage equation.

To address the problem of potential selection bias, I first outline the source of misspecification in the OLS log-wage equations, and then derive selection-correction terms.

---

[32]Technically, the least-squares indicator variable treatment of the effect of private schooling will hold if: 1) the correlation of disturbance terms across equations is zero; 2) conditional upon the X_is, there is zero correlation in the disturbance terms across equations; or 3) the choice of school sector equation is exactly specified, such that $u_i=0$.

If the error terms in the school choice equation (the u_is) are correlated with the error terms in the wage equation (the u_{wi}s), then the expected value of u_{wi} conditional upon attending private school is non-zero:

$$_i'\gamma \geq u_i] \neq 0 \rightarrow E[u_{wi} | Z_i'\gamma \geq u_i] = E[\sigma_{wu} u_i | u_i \leq Z_i'\gamma] = -\sigma_{wu} \frac{\phi}{\phi} \tag{11}$$

where σ_{wu} is the covariance (u_{wi}, u_i), $\phi(\cdot)$ is the standard normal pdf, and $\Phi(\cdot)$ is the standard normal cdf. Note that since we have assumed u_i to have unit variance, we can rewrite σ_{wu} as:

$$\rho_{wu} = \frac{\sigma_{wu}}{\sigma_w \sigma_u} = \frac{\sigma_{wu}}{\sigma_w} , \tag{12}$$

$$\sigma_{wu} = \rho_{wu} \sigma_w , \tag{13}$$

where ρ_{wu} is the correlation between the disturbance terms. Since $-\phi(\cdot)/\Phi(\cdot)$ is, by definition, less than zero, the expected value of u_{wi} for private school attendees is greater than or less than zero as $\sigma_{wu}<0$ or $\sigma_{wu}>0$.

A similar argument applied to the conditional expectation of u_{wi} for public school attendees yields:

$$E[u_{wi} | u_i > Z_i'\gamma] = \sigma_{wu} \frac{\phi(Z_i'\gamma)}{[1 - \Phi(Z_i'\gamma)]} . \tag{14}$$

Since $\phi(\cdot)/[1 - \Phi(\cdot)]$ is positive, the conditional expectation of u_{wi} is greater than or less than zero as σ_{wu} is positive or negative. Hence, we have a log-wage equation where the residuals may be rewritten as:

$$u_{wi} = I_i \left[-\sigma_{wu} \frac{\phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)} \right] + [1 - I_i] \left[\sigma_{wu} \frac{\phi(Z_i'\gamma)}{[1 - \Phi(Z_i'\gamma)]} \right] + \epsilon_{wi} , \tag{15}$$

where ϵ_{wi} has zero conditional mean.

To correct the wage equation for possible selection bias, we define a selection-correction term:

$$h_i = I_i \left[\frac{-\phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)} \right] + [1 - I_i] \left[\frac{\phi(Z_i'\gamma)}{[1 - \Phi(Z_i'\gamma)]} \right] , \tag{16}$$

where the $Z_i'\gamma$s are estimated from a first-stage probit ML estimation of the school choice equation.[33] With the inclusion of h_i in the log-wage equation, one can estimate:

$$\ln[w_i] = X_i'\beta + \sigma_{wu}h_i + \epsilon_{wi} . \quad (17)$$

The inclusion of a selection correction term in the log-wage equation provides a specification test, where the OLS-indicator variable specification (and hence, random assignment) is rejected if the estimated σ_{wu} is significant. This specification, however, restricts the correlation between the disturbance terms in the school choice equation and the log-wage equation to be the same for both private and public school attendees.

Inclusion of the selection-correction term in the log-wage equation both introduces heteroskedasticity and requires a correction to the variance-covariance matrix to account for the fact that the estimated, rather than true, $Z_i'\gamma$s have been included in the second-stage estimation of the log-wage equation. The procedure used to derive the asymptotic variance-covariance matrix is that of Lee, Maddala, and Trost [1980].[34]

An important remaining issue is that of identification, which requires the exclusion of a variable from the X vector that is contained in the Z vector. More intuitively, at least one variable that affects the choice of school sector must be excluded, legitimately, from the log-wage equation.[35] Above, I have indicated that both the measures of the availability of private schools and the costs of private school attendance are candidates for instruments. I estimate the model of school sector choice and wage determination using two-stage estimation, instrumenting for private school attendance with the following private school access measures: number of private high schools per 10,000 population in the county; an indicator for a non-sectarian private high school in the county; and an indicator for a private high school of one's religious affiliation in the county.

Treating the private school attendance decision as endogenous allows one to control for self-sorting of students into private and public high schools. Nevertheless, these specifications restrict the returns to observables to be equal across both private and public school students. Moreover, this model constrains the correlation of the disturbance terms in the school choice and wage equations to be the same for both private and public school attendees.

The most general model with which the private-public earnings and returns to education differentials may be analyzed allows the intercept, returns to observable characteristics, and the disturbance term to vary across private and public school students.[36] If individuals do not self-select into the private and public school sectors, then consistent parameter estimates will be obtained from

---

[33]Note that I write $-\phi(\cdot)/\Phi(\cdot)$ as the selection correction term for private school attendees so that I estimate σ_{wu} across all observations.

[34]The derivation of the asymptotic variance-covariance matrix is contained in Appendix A, which appears at the end of the paper.

[35]Technically, identification may rest solely on functional form.

[36]Such a model is sometimes referred to as a two-population model.

estimating the following system of equations, where N denotes private school attendees and P denotes public school attendees:

$$\textit{Private School Attendees:} \quad \ln[w_{Ni}] = X_{Ni}'\beta_N + u_{Ni} , \tag{18}$$

$$\textit{Public School Attendees:} \quad \ln[w_{Pi}] = X_{Pi}'\beta_P + u_{Pi} . \tag{19}$$

Introducing the possibility of endogenous selection into private and public high schools yields:

$$\ln[w_{Ni}] = X_{Ni}'\beta_N + u_{Ni} \quad \textit{iff} \quad Z_i'\gamma \geq u_i , \tag{20}$$

$$\ln[w_{Pi}] = X_{Pi}'\beta_P + u_{Pi} \quad \textit{iff} \quad Z_i'\gamma < u_i . \tag{21}$$

The disturbance terms in the school choice equation (u_i), and the private and public school log-wage equations (u_{Ni}, u_{Pi}) are assumed to have a trivariate normal distribution with mean vector zeroand variance-covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_N^2 & \sigma_{NP} & \sigma_{Nu} \\ \sigma_{NP} & \sigma_P^2 & \sigma_{Pu} \\ \sigma_{Nu} & \sigma_{Pu} & 1 \end{bmatrix} \tag{22}$$

In Equations 20 and 21, we need to introduce the selection-correction terms (the h_is) derived above in the private and public school log wage equations. Controlling for self-selection yields:

$$\ln[w_{Ni}] = X_{Ni}'\beta_N + \sigma_{Nu}\left[\frac{-\phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)}\right] + \epsilon_{Ni} \quad \textit{iff} \quad Z_i'\gamma \geq u_i , \tag{23}$$

$$\ln[w_{Pi}] = X_{Pi}'\beta_P + \sigma_{Pu}\left[\frac{\phi(Z_i'\gamma)}{[1 - \Phi(Z_i'\gamma)]}\right] + \epsilon_{Pi} \quad \textit{iff} \quad Z_i'\gamma < u_i , \tag{24}$$

where ϵ_{Ni} and ϵ_{Pi} are given by:

$$\epsilon_{Ni} = v_{Ni} + \sigma_{Nu}[h_i - \hat{h}_i] , \tag{25}$$

$$\epsilon_{Pi} = v_{Pi} + \sigma_{Pu}[h_i - \hat{h}_i] . \tag{26}$$

The disturbance terms v_{Ni} and v_{Pi} have zero mean, and ϵ_{Ni} and ϵ_{Pi} have zero mean, conditional on X_i, I_i, and the estimated $\hat{h}$. As noted above, the usual OLS standard errors are not consistent, since inclusion of the selection-correction terms introduces heteroskedasticity into the log-wage equations,

and also necessitates correcting the variance-covariance matrix to account for the fact that the h_is are estimated.

3.4 Measuring the Private-Public Earnings Differential

In the OLS-indicator variable specification, the effect of private schooling on wages is measured by the coefficient on the private school indicator.

Under the two-population models, I calculate the increment in log-wages that would accrue to the "average" public school student if she had received a private school education. Hence, the private-public wage differential is given as follows:

$$\textit{Private-public wage differential: } \bar{X}_P^{/} * [\beta_N - \beta_P] \; . \qquad (27)$$

The private-public wage differential uses the consistent parameter estimates derived from the two-stage log-wage equations to illustrate the effect of private school attendance on the earnings of the hypothetical average public school student's earnings.[37]

4. Description of Data

4.1 National Longitudinal Survey of Youth

The primary data source used in this paper is the National Longitudinal Survey of Youth (NLSY), 1979-1990. The NLSY is a panel survey comprising 12,686 individuals at its inception in 1979, all of whom were between the ages of 14 and 21 on 1 January 1979. The NLSY consists of three subsamples: the cross-sectional sample of 6111 youths, designed to represent the civilian segment of American youth aged 14-21; the supplemental sample of 5295 youth, designed to over-sample black, Hispanic, and economically disadvantaged white youths; and a military sample of 1280 youths, designed to represent the 17-21 year-old population serving in the active-duty military. The survey provides sampling weights which allow one to weight properly over-sampled sub-groups and derive population estimates. Respondents have been interviewed annually since 1979, and non-response rates and attrition from the panel have been quite low. The most recent year for which survey data are available, 1990, provides a panel of 11 years of observations on individuals' educational attainment and labor market histories.

---

[37]These calculations are presented in Table 10. The hypothetical average public school student is a white Catholic male who grew up in an SMSA in the New England Census region. This man is a "middle-born" child who lived with both parents at age 14, had access to reading materials in the home, whose mother worked outside the home when he was fourteen, and who lived in a county that contained both a Catholic and a non-sectarian private high school. All remaining indicator variables were set to zero, and all remaining continuous variables were set to the mean values over all public school students.

In addition to usual measures of parental socioeconomic status, such as income, educational attainment, and occupational status, the NLSY survey includes information on the number of siblings and information which allows one to construct measures of birth-order. Moreover, the NLSY provides information on a subset of these family background measures and area of residence when the respondent was 14.[38] These variables are particularly useful, since they characterize family demographics at the time (the eighth grade for most respondents) when high school attendance decisions are decided, or at least under consideration. Finally, the NLSY asks a series of questions regarding the dates of previous changes of residence.[39] Knowledge of previous moves and the dates of attendance at the high school allow me to exclude those individuals who did not attend the same school for the duration of their high school enrollment.[40]

4.2 County-Level Demographic Information

Use of the Geocode supplement to the NLSY provides state and county identifiers for all respondents in the survey, as well as a wealth of demographic and local labor market variables at the county level. Given that one's county of residence is known, county level data from external sources may be merged into the data set. From the *1983 and 1988 City and County Data Books*, I extracted information on median family income and median house values at the county level, and merged this information into the data. Finally, I use the state identifiers to merge information from the *Digest of Educational Statistics* concerning state-average measures of public school quality (e.g. current per pupil expenditures).

4.3 Private School Data

In the 1979 NLSY survey, respondents were asked:

> *"Is the high school you are currently attending or the high school in which you were last enrolled a public or private (parochial) school?"*

The response to this question represents the extent of information I have regarding private school attendance in my sample. As a result, I restrict my analysis to only those youths who were enrolled in the same high school for the duration of their high school attendance.

---

[38]The NLSY asked survey respondents for the following retrospective information in 1979: did father/mother work for pay when respondent was age 14; occupational status of mother's/father's job at age 14; did respondent live in a two-parent family at age 14; did respondent have access to newspapers, magazines, and a library card at home at age 14; and geographic location of residence at age 14.

[39]Questions concerning all previous residence changes were asked in 1979, 1980, and 1982.

[40]Given that I know only whether the last high school attended by a respondent was a private or public school, this exclusion allows me to consider only those individuals who received their high school education in the same school (and same school sector).

While it is not possible to determine the exact identity or the affiliation of the private high schools attended by respondents in the NLSY with the public-use version of the tape, it is possible to construct measures of the availability of private high schools in one's county of residence. The Department of Education's *1980-1981 Private High Schools Data File* provides a universe list of 4982 private high schools in the United States that were in operation during the 1979-1980 and 1980-1981 academic years, as well as each school's location and religious affiliation, if any.

Further, I collected information on reported day tuition and fees at these private high schools from private school handbooks, particularly the *Porter-Sargent Handbook of Private Schools* and *Official Catholic Directories*. Data Appendix B provides a detailed description of the construction of the tuition information.

Individual high schools were aggregated to the county level to create variables indicating: the presence of private school(s) within the county; the presence of private high school schools, by school affiliation; and median day tuition and fees at private schools in the county, by school affiliation. Merging these data with the NLSY data enabled me to determine whether one's county of residence contained a private high school of one's religious affiliation.

4.4 The Selected Sample

From the initial NLSY sample of 12,686 youths, I selected those individuals aged 14-19 in 1979 who were not in the active-duty military and were enrolled in high school at the time of the 1979 interview. Further, I excluded those individuals who reported moving during the time period 1 June of the year in which they turned 14 until the date at which they reported having left or completed high school.[41] Inclusion in the sample required at least two consecutive years of "participation" in the labor market, where participation required one to have worked a positive number of hours and precluded enrollment in regular schooling.[42]

The final sample consists of 2330 youths, all of whom were enrolled in high school during the fall of 1979, and had entered the labor market before the 1989 NLSY survey was administered.[43] Of this group, 119 were enrolled in private high schools for the duration of their high school attendance.

---

[41]Individuals who were administered Version B of the Household Interview (15 individuals in my sample, four private, eleven public) were not excluded, as they may have been living in boarding school dorms or juvenile centers.

[42]See Table 1 and Data Appendix D for the precise definition of an observed entry into the labor market.

[43]Hence, a potential source of censoring bias arises in that a respondent who graduated from high school in 1982, attended college, and subsequently enrolled in medical school or a Ph.D. program in Economics would be excluded from the sample. As noted in Table 1, 18 respondents remained enrolled in school "continuously" from 1988 to 1990, two of whom were graduates of private high schools. Any potential bias resulting from this exclusion is likely to be small.

5. Empirical Results

In this section of the paper, I discuss the empirical results obtained from estimating the private school choice and earnings models developed in Section 3. Table 1 presents the rules describing the evolution of the sample. Tables 2, 3, and 4 provide summary statistics for the variables used to estimate the school choice and log-wage equations. Since these statistics were reviewed previously in the text, I move to a discussion of the school choice model results, which are contained in Table 5. Years of schooling equations and the concomitant tests for self-selection in educational attainment, are reported in Table 7. I then focus on the various log-wage equations estimated to analyze the effect of private schooling on earnings and earnings growth (Tables 9-12). This section concludes with a discussion of findings on the rates of return to education for private and public school attendees (Table 13).

5.1 School Choice

Table 5 reports parameter estimates from the private school choice equation, estimated using probit maximum likelihood (LIML). The first column of Table 5 contains parameter estimates from the probit school choice equation. The next two columns contain two alternative measures of marginal effects estimates — the change in the probability of attending private school due to a change in one of the regressors. Column two contains average derivative estimates, whereas column three reports derivative of the average estimates, where the derivative of the fitted values is evaluated at the vector of mean values of the regressors. Due to the numerous indicator variables in the probit equation, column four reports derivative of the average marginal effects estimates for an individual with the following characteristics: a white Catholic male living in an SMSA in a county in New England that contains a nonsectarian private high school and a Catholic private high school. This man is a "middle-born" child who lived with both parents at age 14, had access to newspapers, magazines and a library card at age 14, and whose mother worked outside the home when he was fourteen. The interpretation of the marginal effects is as follows. For the average derivative estimates of column two, the presence of a private high school of one's religious affiliation increases the likelihood of one's attendance at a private high school by 4.15 percentage points. Further, if current per pupil expenditures in one's state of residence increase by one thousand dollars, the likelihood of attending a private high school falls by 7.95 percentage points. The interpretation of the derivative of the average marginal effects estimates is similar.

All of the variables reflecting the ease of access or "availability" of private schools are positive and significant, with the exception of the indicator for the presence of a private school of any type in the county. One would expect that a parent would prefer to send her or his child to a private school of the same religious denomination, or at least a non-sectarian school, rather than a private religious school of a different denominational affiliation. As a result, I include separate indicators for non-sectarian schools and private schools of one's religious denomination in the school choice probit equation.[44]

---

[44] Admittedly, the presence of a private school in a county, and its continued operation, signals dissatisfaction with the public schools, at least among some constituency. The dissatisfaction theory, however, fails to provide a compelling explanation as to why Catholic schools continue to remain open. While the impetus for founding Catholic schools in the mid- and late-1800s was dissatisfaction

The results from Table 5 provide insights into the determinants of private school enrollment, and serve to illustrate the relative importance of family background, religion, and the availability of private schooling in influencing the private school enrollment decision. Further, bringing additional data to bear on the school choice decision provides a number of promising instruments to identify the school choice and earnings equation system.

5.2 Private School Attendance, Selection, and Earnings

There are three primary concerns which must be addressed in the empirical estimation of a "private school effect" on subsequent labor market outcomes:

1) Students who attend private and public schools have backgrounds that are markedly different. (To continue the production function analogy, the quality of inputs differs.)

2) Students do not "arrive" at private schools by random assignment; there is, potentially, self-selection of students into private and public schools.[45]

3) Educational attainment differs substantially across school sectors. Forty-six percent of all private school students complete four years of college, *vis-a-vis* twenty-four percent of public school students.

---

with the Protestant-controlled public school system, the continued existence of the Catholic schools may be the result of tastes for religious schooling among Catholics and the efforts of the Catholic church to provide affordable non-public education. For example, parish and diocesan Catholic schools receive 15-30% of their operating income in the form of subsidies from the parish they serve or from the diocese in which they operate. Yeager *et al.* [1985], pp. 61-62. This subsidization, along with the cost savings from the use of contributed services, enables Catholic schools to charge lower tuitions. Moreover, nearly 40% of Catholic schools report that students enroll in courses at the local public high school. Hence, many Catholics do not appear to be averse to receiving a "partially-public" education.

The presence of a non-sectarian private school, however, may be a more accurate signal of dissatisfaction with the public schools, as 72% of all private non-sectarian schools were founded in the last 25 years. Benson and McMillen [1991], p. 14. Moreover, many of the recently opened "academies" in the South are organized as non-sectarian schools.

[45]An interesting issue that has not as yet been considered is the double-selection inherent in **admission** to a private school as opposed to **application** to a private high school. Data constraints preclude the estimation of such a bivariate probit model of private school admissions. Nonetheless, the question of queuing is potentially important, given that many private schools practice selective admissions. In 1982, the average Catholic high school admitted 88% of applicants. Moreover, 36% of all Catholic high schools reported maintaining a waiting list for admissions; for private Catholic schools, 51% reported a waiting list. Yeager *et al.* [1985], p. 213. At elite private schools, the selectivity of admissions is much higher, with less than one of seven applicants admitted. Townsend [1989].

I first consider the schooling equations, investigating the possibility of selection biases in educational attainment. Then, treating schooling as endogenous in the log-wage equations, I address the remaining concerns sequentially, looking first at the simplest specification of a log-wage equation that does not control for endogenous selection into private schooling, and then incorporating these "complications" into the empirical analysis.

Tables 6 and 7 report parameter estimates from educational attainment equations. I estimate both one- and two-population models to allow the returns to observables to vary across school sector. Further, I test for endogenous selection in the educational attainment equations. An examination of Table 7 reveals that although returns to observables differ across school sectors, there is no evidence of selection bias in the schooling equations.

Table 8 presents estimates of several OLS log-wage equations in which the private school control is entered as an indicator variable. Again, educational attainment is treated as endogenous. These results, however, do not include controls for the possible endogeneity of private school choice. Column 1 of Table 8 is a baseline log-wage equation that excludes controls for private school attendance, and provides a means of evaluating the parameter estimates obtained from a garden variety log-wage equation estimated with the data. The introduction of controls for private school attendance (Column 2) yields a private school attendance effect that is negative and significant. Treating private school attendance as an exogenous indicator variable suggests that the effect of attending private school reduces wages by 9.1 percent.[46]

The third column of Table 8 presents a direct specification test for the presence of endogenous selection into school sector by replacing the private school indicator with the selection-correction terms (h_is). Since the model restricts $\rho_{Nu}\sigma_N=\rho_{Pu}\sigma_P=\rho_{wu}\sigma_w$, I find evidence of positive selection bias in the log-wages of public school attendees, and hence negative selection bias in the log-wages of private school students. More intuitively, the finding that $\rho_{wu}\sigma_w>0$ implies that individuals who are unlikely to enroll in private school (those with large values of u_i in the school choice equation) have higher residual earnings potential. Hence, those who are more likely to attend private high schools (those with smaller values of u_i) have lower residual earnings capacity. More importantly, the significance of the coefficient on the selection correction terms, particularly in the log-wage equation for public school attendees, provides evidence of selection bias in the log-wage equations. Given the finding of selection bias, the OLS-indicator variable approach to measuring the effect of private schooling on earnings is invalid.

To this point, I have not allowed the observed returns to background to vary across the samples of private and public high school students. One of the main channels through which private schooling may affect wages, however, is by increasing the return to a year of schooling, especially if one believes private schools are more effective in educating students. To consider the differential return to education across sectors, I estimate "two-population" models of school choice and log-wages that free all estimated parameters to vary across sectors. The estimation of two-population models further allows the covariance of the disturbance terms to vary across school sector, so that evidence of positive selection bias in the earnings of public high school attendees need not imply negative selection bias in the earnings of private high school attendees.

---

[46]This result may arise from the fact that many respondents in my sample (average years of labor market experience are approximately seven, cf. Table 3) are to the left of the Mincerian "overtaking point", which is near 12 years of labor market experience.

Table 9 reports two alternative specifications of the two-population model estimated using two-stage techniques: the first two columns treat years of schooling as exogenous, whereas the last two columns allow educational attainment to be an endogenous variable in the log-wage equations. In all of the log wage equations, a first-stage probit estimation of school sector choice is used to generate selection-correction terms for the separate log-wage equations for private and public high school attendees (e.g. an endogenous switching regressions framework).

The finding of positive selection bias in the earnings of public school attendees is consistent across all model specifications, but I find no evidence of selection bias in the earnings of private school attendees. Moreover, treating years of schooling as an endogenous variable results in an increase in the rate of return to schooling for private school attendees (0.097 to 0.131), and no change in the rate of return to schooling for public school attendees. Finally, none of the school quality measures exhibit a significant effect on the earnings of individuals in either school sector.

5.3 Private-Public Wage Differentials

Table 10 presents a summary of estimates of the private-public wage differential. The first page of Table 10 reviews the private-public wage differentials calculated from the various specifications presented above. The private-public wage differential is calculated as the increment to log-wages the average public school student would have received had she attended a private high school.[47] Controlling for the endogeneity of selection into school sector and the endogeneity of years of schooling, I find that the average public school student would have attained nearly one and one-half years (1.457 years) of additional education and received a positive wage increment of 0.133 had she attended private schools, although this difference is not significant. For the average public school student who completed twelve years of schooling, the private school treatment effect results in an additional 0.39 years of education and a +0.037 increment to wages, although this differential is again not significant. For the average public school student who completed college, however, the private school treatment effect results in an additional 1.36 years of schooling and a +0.122 increment to wages, which is not significant. The combination of the additional schooling attained and the higher rate of return to education in the private school sector explains approximately seventy percent of the wage differential for all of these hypothetical students.

Finally, Table 10 summarizes the results of specification tests for the presence and nature of selection biases in the log-wage equations. I find consistent evidence of positive selection bias in the earnings of public school attendees. For private school attendees, the sign of the estimated covariance term is negative, suggesting positive selection bias in the log-wages of private school attendees, but this selection bias is not significant.

---

[47]In this calculation, the first step inserts the characteristics of the representative public school student into the schooling equation for private school students, which provides an estimate of the educational attainment of the representative public school student given a private high school education. Then, this estimated educational attainment is included in the vector of the representative public school student's characteristics, and the treatment effect of private schooling on earnings is calculated.

Table 11 reports results of overidentification tests performed on the instrument set. None of the possible overidentifying restrictions could be rejected. Both the private school availability variables and the home environment measures were found to be valid instruments.[48]

5.4 Private-Public Differences in Wage Growth Rates

Since my sample comprises relatively young workers, one possible concern is that private school attendees may exhibit lower log-wages in levels, but enjoy higher growth rates of log-wages. As an example, a private school attendee who graduated from college in 1987 with a bachelor's degree in Business Administration and currently works as an entry-level management trainee may earn lower wages than a 1983 public high school graduate working as a welder, but the management trainee may be on a much steeper wage profile. To test for possible selection biases in wage-growth rates, I estimate log-wage change equations for each school sector, in which the dependent variable is the average annual growth in log-wages from the year of labor market entry to 1990. Once again, educational attainment is treated as an endogenous variable in the wage growth equations.

I find positive selection bias in the growth rate of log-wages for private school students and negative selection bias in the growth rate of log-wages for public school students. Hence, individuals who are more likely to attend private schools (those with small values of u_i) have higher residual earnings growth, and *vice versa*. Further, years of schooling has a positive and significant effect on wage growth for private school students, but no effect on wage growth for public school students. Public school students' wage growth is more closely aligned with a job-matching story, given the significance of the number of jobs held measure in the wage-growth equation. This finding is in agreement with earlier work by Mincer [1986], who finds that approximately two-thirds of the wage growth of young men arises as a result of job mobility.

5.5 Private-Public Differences in the Rates of Return to Education

To this point, I have found mixed evidence in support of the contention that private schools are more effective "producers" of education. Until now, however, I have not examined differences in the rates of return to education across private and public school attendees. If one is looking for a measure of the private-public high school differential in productivity or value-added, it seems reasonable to assume that "the body is buried" in the rate of return to education.

Table 13 reports estimates of the rates of return to education and covariances of the school choice and log-wage residuals from several different specifications of the log-wage equations. Specifically, the first page of Table 13 reports the change in the rates of return to education as additional explanatory variables are included in the log-wage equations. All models treat years of schooling as endogenous. The inclusion of family background measures raises the return to schooling for private school students slightly, and has little effect on the return to education for public school students. More surprising is that the inclusion of school quality and ability measures results in a slight increase in the return to schooling for private school students (0.130 to 0.145) and no change

---

[48]To further test the validity of these instruments, I am conducting local average treatment effects tests on the set of instruments, as suggested in Imbens and Angrist [1991].

in the return to education for public school students. Interestingly, the rate of return estimates in Table 13 illustrate that the effects of family background operate primarily through educational attainment rather than directly diminishing the return to education in the log-wage equations.

The second page of Table 13 considers an alternative specification of the log-wage equations, in which years of schooling are treated as an additive spline function, with a discontinuity at twelve years of education. This specification provides a direct test of the effect of private school attendance for those who terminated their schooling at or before high school graduation. In these models, the private-public differential in returns to schooling is even more pronounced, and does not diminish at higher levels of education. These specifications also reveal that the returns to education do not differ across the level of educational attainment. The results in Table 13 provide corroborating evidence that the returns to private high school attendance are significantly higher than those for public high school attendance, and this differential cannot be attributed to differences in years of schooling completed across school sectors.

One finding has been robust across all of the models estimated: students who attend private high schools realize significantly higher rates of return to education than their public school counterparts, independent of the level of schooling attained. Moreover, the private school differential in the returns to schooling is not an artifact of differences in student background or observable, quantifiable measures of school quality in the two school sectors.

6. Conclusions

Initially, I posed the following set of questions to address in this paper: controlling for self-selection into private and public schools and the endogeneity of educational attainment, do students who attend private high schools have higher earnings potential? Do they enjoy greater returns to schooling? Are these differentials due to differences in student quality, such as family background and observable measures of school quality, or are private schools truly more productive in educating students?

To address these questions, I estimate a model of school choice and wage determination that allows the returns to background and education to vary across school sector. I find consistent evidence that private school students realize higher rates of return to schooling, controlling for the endogeneity of educational attainment. Including measures of family background, ability, and school quality do not diminish this differential. The return to a year of private high school (0.162) is three times greater than the return to a year of public high school education (0.051). More generally, the return to a year of education for private school students (0.131) is twice the return to a year of education for public school students (0.066). While provocative, these findings surely require corroboration from additional data sources, as unobserved differences in high school and secondary school quality may provide an alternative explanation of these results.

The principal finding of this paper — that private school students exhibit higher rates of return to schooling — makes it tempting to conclude that school choice policies that allow parents access to the private schools of their choice would result in welfare improvements to society as a whole. Nonetheless, it is important to realize that the analysis in this paper is partial equilibrium, in that it does not consider the effects of a non-trivial redistribution of students across the two school sectors. Part of the private school advantage in educating students may stem from their enrollment

of students from relatively advantaged backgrounds, which creates a population of students who face fewer challenges and, quite likely, more support and assistance at home.

In particular, the implementation of a tuition tax credit policy or voucher system might not generate substantial benefits to all participants. Undoubtedly, some clienteles would benefit from increased access to the private schools, but this does not suggest that all students would be better off receiving a private high school education. Any analysis of a proposed voucher program should consider carefully those groups most likely to benefit from increased scope for choice, and hence those most likely to switch school sectors. Moreover, future analyses should weigh the potential effects of voucher programs on changing the size and composition of the private and public school sectors.

Given that the rate of return to private high school education is significantly higher than the rate of return to public high school education, and this differential is robust to the inclusion of family background and school quality measures in the log-wage equations, future research should move toward a more careful consideration of the dimensions along which private and public schools differ. For example, the curricula in the private schools may consist of more units of "advanced" courses in mathematics, English, science, and foreign languages than those found in public schools. Alternatively, private school attendance during high school may provide a credential that allows one access to post-secondary institutions of higher quality. Many of these questions can be addressed using data with more complete information on individuals' post-secondary education, such as the National Longitudinal Study of the High School Class of 1972. Use of this data provides detailed information on the post-secondary institutions attended, and also allows one to control for unobservable family and high school effects. Estimating a fixed-effects model that parses out unobservable differences in family background and high school attended would represent a significant contribution to furthering our understanding of the private school effect.

In closing, it is important to recognize that the provision and production of secondary schooling in America is too important to warrant continued disinterest and neglect. The private high school system in the U.S. provides many examples of schools that differ in both their organization and educational mission. Admittedly, attendance at these schools is not the outcome of random assignment, hence no natural experiments are available. Nonetheless, the private schools offer a wealth of information concerning the private production of education. Clearly, more data needs to be collected on both private high schools in the U.S. and the students they serve. Most importantly, future analyses should include a careful consideration of the differences in the educational production processes within the private and public high schools.

APPENDIX A

Derivation of the Asymptotic Covariance Matrices: Endogenous Switching Regressions Model for School Choice and Earnings Determination

In Section 3 of the text, the models of school sector choice and earnings determination were developed for the endogenous switching regressions model specification as follows:

Private School Choice Probit Equation:

$$I_i^* = Z_i'\gamma - u_i \ . \tag{1}$$

Second-stage Log-wage Equations:

Private School Attendees:

$$\ln[w_{Ni}] = X_{Ni}'\beta_N + u_{Ni} \quad if \quad Z_i'\gamma \geq u_i \ , \tag{2}$$

Public School Attendees:

$$\ln[w_{Pi}] = X_{Pi}'\beta_P + u_{Pi} \quad if \quad Z_i'\gamma < u_i \ . \tag{3}$$

The residuals for the above system of equations have a multivariate normal distribution with mean vector zero and covariance matrix cited in Equation (22) of the text.

I follow the convention developed in the text in denoting observations on private school attendees with the subscript N and observations on public school attendees with the subscript P. Further, assume that there are N_N observations for which $I_i \geq 0$ and N_P observations for which $I_i < 0$, yielding a total sample size of $N=N_N+N_P$. For notational simplicity, let us define:

$$\phi_i = \phi(Z_i'\gamma) \ , \tag{4}$$

and

$$\Phi_i = \Phi(Z_i'\gamma) \ , \tag{5}$$

where ϕ is the standard normal density and Φ is the cumulative normal distribution.

Now, define the matrix Z such that

$$Z = \begin{bmatrix} Z_1' \\ Z_2' \\ \vdots \\ Z_N' \end{bmatrix}, \tag{6}$$

where Z is an NxL matrix which can be partitioned as follows:

$$Z = \begin{bmatrix} Z_N \\ Z_P \end{bmatrix}. \tag{7}$$

Further, let us define the following diagonal matrices:

$$\Lambda = diagonal\left[\frac{\phi_i^2}{\Phi_i[1 - \Phi_i]}\right] \quad (NxN) , \tag{8}$$

$$A = diagonal\left[Z_i'\gamma\frac{\phi_i}{\Phi_i} + \left[\frac{\phi_i}{\Phi_i}\right]^2\right] \quad (N_N x N_N) \text{ , and} \tag{9}$$

$$B = diagonal\left[\left[\frac{\phi_i}{[1 - \Phi_i]}\right]^2 - Z_i'\gamma\left[\frac{\phi_i}{[1 - \Phi_i]}\right]\right] \quad (N_P x N_P) . \tag{10}$$

Rewriting Equations (2) and (3) above to control for self-selection yields:

Private School Attendees:

$$\ln[w_{Ni}] = X_{Ni}'\beta_N + \sigma_{Nu}\left[\frac{\phi_i}{\Phi_i}\right] + \nu_{Ni} , \tag{11}$$

Public School Attendees:

$$\ln[w_{Pi}] = X_{Pi}'\beta_P + \sigma_{Pu}\left[\frac{\phi_i}{[1 - \Phi_i]}\right] + \nu_{Pi} . \tag{12}$$

We first obtain an estimate of γ by probit maximum likelihood estimation. We then estimate the following second stage log-wage equations:

$$\ln[w_{Ni}] = X_{Ni}'\beta_N + \sigma_{Nu}\left[\frac{-\phi_i}{\Phi_i}\right] + \epsilon_{Ni} , \textit{ and} \tag{13}$$

$$\ln[w_{Pi}] = X_{Pi}'\beta_P + \sigma_{Pu}\left[\frac{\phi_i}{[1 - \Phi_i]}\right] + \epsilon_{Pi} , \tag{14}$$

where:

$$\phi_i = \phi(Z_i'\gamma) , \tag{15}$$

$$\Phi_i = \Phi(Z_i'\gamma) , \tag{16}$$

$$\epsilon_{Ni} = \nu_{Ni} - \sigma_{Nu}\left[\left[\frac{\phi_i}{\Phi_i}\right] - \left[\frac{\phi_i}{\Phi_i}\right]\right] , \textit{ and} \tag{17}$$

$$\epsilon_{Pi} = \nu_{Pi} + \sigma_{Pu}\left[\left[\frac{\phi_i}{[1 - \Phi_i]}\right] - \left[\frac{\phi_i}{[1 - \Phi_i]}\right]\right] , \tag{18}$$

where ν_{Ni} and ν_{Pi} have mean zero.

Now, let us consider the log-wage equation for private school attendees, and further define the following matrices:

$$W_N = \left[X_N, \left[\frac{-\phi}{\Phi}\right]\right], \tag{19}$$

$$\mathbb{W}_N = \left[X_N, \left[\frac{-\phi}{\Phi}\right]\right] . \tag{20}$$

Hence, the two-stage estimator may be written as:

$$\begin{bmatrix} \beta_N \\ \sigma_{Nu} \end{bmatrix} = (\mathbb{W}_N' \mathbb{W}_N)^{-1} \mathbb{W}_N' \ln[w_N] . \tag{21}$$

It then follows that

$$\left[\begin{bmatrix} \beta_N \\ \sigma_{Nu} \end{bmatrix} - \begin{bmatrix} \beta_N \\ \sigma_{Nu} \end{bmatrix}\right] \underline{D} (W_N' W_N)^{-1} W_N' [v_N + \sigma_{Nu} A Z_N (\gamma - \gamma)] . \tag{22}$$

The asymptotic covariance matrix is written as:

$$\begin{aligned} \Big| &= (W_N' W_N)^{-1} W_N' [var(v_N) + \sigma_{Nu}^2 A Z_N var(\gamma) Z_N' A' + \sigma_{Nu} A Z_N cov \\ &\quad + \sigma_{Nu} cov(\gamma', v_N) Z_N' A'] W_N (W_N' W_N)^{-1} . \end{aligned} \tag{23}$$

Below, I derive the variance and covariance terms contained in Equation (23) above.

First, let us define an indicator variable for private school attendance such that:

$$H_i = \begin{cases} 1 \text{ if } I_i \geq 0 , \\ 0 \text{ if } I_i < 0 . \end{cases} \tag{24}$$

To derive the variance of the estimated γ, we first note that:

$$\gamma - \gamma) \stackrel{D}{=} \left[\sum_{i=1}^{N}\left[\frac{\phi_i^2}{\Phi_i[1 - \Phi_i]}Z_i Z_i'\right]\right]^{-1}\sum_{i=1}^{N}\left[\frac{\phi_i}{\Phi_i[1 - \Phi_i]}Z_i(H_i - \Phi_i)\right] \tag{25}$$

Equation (25) may be rewritten as:

$$(\gamma - \gamma) \stackrel{D}{=} (Z'\Lambda Z)^{-1}Z'\Lambda_N\begin{bmatrix} H_N - \Phi_N \\ H - \Phi \end{bmatrix}, \tag{26}$$

and hence it follows that:

$$var(\gamma) = (Z'\Lambda Z)^{-1} . \tag{27}$$

From Equation (26), we can show that the covariance terms in Equation (23) are equal to zero, since:

$$E(\gamma - \gamma)\nu_N' =$$
$$Z'\Lambda_N E\left[\begin{bmatrix} H_N - \Phi_N \\ H - \Phi \end{bmatrix} x\left[\epsilon_{N1} + \sigma_{Nu}\frac{\phi_{N1}}{\Phi_{N1}},\ldots,\epsilon_{NN_N} + \sigma_{N_u}\frac{\phi_{NN_N}}{\Phi_{NN_N}}\right]\right| \begin{matrix} D_N \\ D_{N_N} \end{matrix} \tag{28}$$

Lastly, we have:

$$var(\nu_N) = \sigma_N^2 I_{N_N} - \sigma_{Nu}^2 A . \tag{29}$$

Substituting the above variance terms into Equation (23) yields:

$$\sigma_N^2(W_N'W_N)^{-1} - \sigma_{Nu}^2(W_N'W_N)^{-1}W_N'[A - AZ_N(Z'\Lambda Z)^{-1}Z_N'A]W_N(\tag{30}$$

Following a similar procedure, one can derive the asymptotic covariance of the second-stage estimator for public school attendees to be:

$$= \sigma_P^2(W_P'W_P)^{-1} - \sigma_{Pu}^2(W_P'W_P)^{-1}W_P'[B - BZ_P(Z'\Lambda Z)^{-1}Z_P'B]W_P(\tag{31}$$

APPENDIX B
Derivation of Wegge's Overidentification Test

Given the log wage equation

$$\ln[w_i] = S_i\zeta_i + \sigma_u \cdot fitted\,value_i + X_i\beta_i + X_i^*\beta_i^* + \epsilon_i \,, \tag{1}$$

in which

$$\beta_i^* = 0 \tag{2}$$

overidentifies the equation, we can define the following terms:

$$M_i = I - X_i(X_i'X_i)^{-1}X_i' \,, \tag{3}$$

$$M = I - X(X'X)^{-1}X' \,, \tag{4}$$

$$W^* = M_i - M \,, \tag{5}$$

$$M^* = W^* - W^*S_i(S_i'W^*S_i)^{-1}S_iW^*\ln[w] \,. \tag{6}$$

The 2SLS-IV estimator of β_i^* is given by:

$$\hat{\beta}_i^* = (X_i^{*\prime}M^*X_i^*)^{-1}X_i^{*\prime}M^*\ln[w_i] \,. \tag{7}$$

Now, since:

$$\sqrt{N}(\hat{\beta}_i^* - \beta_i^*) \xrightarrow{D} N(0,\sigma_{ii}plim(X_i^{*\prime}M^*X_i^*/N)^{-1}), \tag{8}$$

the following quadratic form is obtained:

$$(\hat{\beta}_i^* - \beta_i^*)'(X_i^{*\prime}M^*X_i^*)(\hat{\beta}_i^* - \beta_i^*)/\sigma_{ii} \xrightarrow{D} \chi^2_{k_i^*} \,, \tag{9}$$

where σ_{ii} is the 2SLS-IV estimate of σ_{ii} and k_i^* equals the number of excluded exogenous variables. Under the null hypothesis, $\beta_i^*=0$, in which case the test statistic may be rewritten as:

$$\beta_i^{*\prime}(X_i^{*\prime} M^* X_i^*)\beta_i^*/\sigma_{ii} \ . \qquad (10)$$

The above derivation is based on Wegge [1978] and Hwang [1980].

References

Abramowitz, S. and E.A. Stackhouse 1980. *The Private High School Today.* Washington: U.S. Department of Education.

Altonji, Joseph G. 1988. "Controlling for Personal Characteristics, School and Community Characteristics, and High School Curriculum when Estimating the Return to Education." (mimeo, Columbia University).

Amemiya, Takeshi 1985. *Advanced Econometrics.* Cambridge: Harvard University Press.

Angrist, Joshua D. 1990. "Instrumental Variables Estimation of Qualitative Response Models with Dummy Endogenous Regressors." (mimeo, Harvard University).

Angrist, Joshua D. and Alan B. Krueger 1991. "Does Compulsory Schooling Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106: 979-1014.

Angrist, Joshua D. and Alan B. Krueger 1991. "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." (mimeo, Harvard University).

Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, 2d ed.* New York: National Bureau of Economic Research.

Becker, Gary S. 1981. *A Treatise on the Family.* Cambridge: Harvard University Press.

Behrman, Jere R., Zdenek Hrubec, Paul Taubman, and Terence Wales 1980. *Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family, Environment, and Schooling.* New York: North Holland.

Benson, Peter and Marilyn McMillen 1991. *Private Schools in the United States: A Statistical Profile, with Comparisons to Public Schools.* Washington: National Center for Education Statistics.

Bowles, Samuel and Herbert Gintis 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life.* New York: Basic Books.

Card, David and Alan Krueger 1992. "Does School Quality Matter? Returns to Education and Characteristics of Public Schools in the United States." *Journal of Political Economy* 100: 1-40.

Carper, James C. and Thomas C. Hunt 1984. *Religious Schooling in America.* Birmingham, AL: Religious Education Press.

Chipty, Tasneem and Daniel Kessler 1992. "Do Private Schools Educate Students Better? The Private-Public Achievement Differential." (mimeo, MIT).

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

Coleman, J.S. 1966. *Equality of Educational Opportunity.* Washington: U.S. Government Printing Office.

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

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

Coons, John E. and Stephen D. Sugarman 1978. *Education by Choice: The Case for Family Control.* Berkeley: University of California Press.

Devins, Neal E. 1989. *Public Values, Private Schools.* New York: The Falmer Press.

Digest of Education Statistics, annual eds., 1978-1984. Washington: U.S. Government Printing Office.

Freeman, Richard B. 1986. "Who Escapes? The Relationship of Churchgoing and Other Background Factors to the Socioeconomic Performance of Black Male Youths from Inner-City Tracts" in Freeman, Richard B. and Harry Holzer (eds.), *The Black Youth Unemployment Crisis.* Chicago: University of Chicago Press.

Freeman, Richard B. 1986. "Demand for Education" in Ashenfelter, O. and R. Layard (eds.), *Handbook of Labor Economics, vol. I.* New York: Elsevier.

Greene, William 1990. *Econometric Analysis.* New York: MacMillan.

Greene, William 1981. "Sample Selection Bias as a Specification Error: Comment." *Econometrica* 49: 795-798.

Griliches, Zvi 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* 45: 1-21.

Handbook of Private Schools, 59th-61st eds. 1978-1980. Boston: Porter-Sargent Publishers, Inc.

Hanushek, Eric 1981. "Throwing Money at Schools." *Journal of Policy Analysis and Management* 1: 19-41.

Hanushek, Eric 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24: 1141-1177.

Hausman, J.A. 1978. "Specification Tests in Econometrics." *Econometrica* 46: 1251-1271.

Heckman, J.J. 1978. "Dummy Endogenous Variables in a Simultaneous Equations System." *Econometrica* 46: 931-960.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153-162.

Imbens, G.W. and J.D. Angrist 1991. "Identification and Estimation of Local Average Treatment Effects." (mimeo, Harvard University).

Judge, G. *et al.* 1985. *The Theory and Practice of Econometrics, 2d ed.* New York: John Wiley and Sons.

Juster, F. Thomas 1975. *Education, Income, and Human Behavior.* New York: McGraw Hill.

Kearns, David T. and Denis P. Doyle 1988. *Winning the Brain Race: A Bold Plan to Make our Schools Competitive.* San Francisco: ICS Press.

Kenny, Lawrence W., Lung-Fei Lee, G.S. Maddala, and R.P. Trost 1979. "Returns to College Education: An Investigation of Self-Selection Bias Based on the Project Talent Data." *International Economic Review* 20: 775-789.

Kraushaar, Otto F. 1972. *American Nonpublic Schools: Patterns of Diversity.* Baltimore: Johns Hopkins University Press.

Layard, Richard and George Psacharopoulos 1974. "The Screening Hypothesis and the Returns to Education." *Journal of Political Economy* 82: 985-998.

Lee, Lung-Fei 1978. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *International Economic Review* 19: 415-433.

Lee, Lung-Fei 1979. "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables." *Econometrica* 47: 977-995.

Lee, Lung-Fei, G.S. Maddala, and R.P. Trost 1980. "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity." *Econometrica* 48: 491-503.

Levy, Daniel C. 1986. *Private Education: Studies in Choice and Public Policy.* New York: Oxford University Press.

Lillard, Lee A. and Robert J. Willis 1978. "Dynamic Aspects of Earnings Ability." *Econometrica* 46: 985-1012.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York: National Bureau of Economic Research.

Mincer, Jacob. 1986. "Wage Changes in Job Changes" in *Research in Labor Economics, vol. 8.* Greenwich, CT: JAI Press.

Murnane, Richard J. 1984. "Comparisons of Public and Private Schools: Lessons from the Uproar." *Journal of Human Resources* 19: 263-278.

Murnane, Richard J., Stuart Newstead, and Randall J. Olsen 1985. "Comparing Public and Private Schools: The Puzzling Role of Selectivity Bias." *Journal of Business and Economics Statistics* 3: 23-35.

Nathan, Joe 1989. *Public Schools by Choice.* Minneapolis: Free Spirit Publishing.

Riley, John G. 1979. "Testing the Educational Screening Hypothesis." *Journal of Political Economy* 79: S27-S52.

Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 1: 135-146.

Taubman, Paul 1975. *Sources of Inequality in Earnings.* New York: North-Holland.

Taubman, Paul and Terence Wales 1974. *Higher Education and Earnings.* New York: McGraw Hill.

Taubman, Paul and Terence Wales 1973. "Higher Education, Mental Ability, and Screening." *Journal of Political Economy* 73: 28-55.

Tomes, Nigel 1984. "The Effects of Religion and Denomination on Earnings and the Return to Human Capital." *Journal of Human Resources* 19: 33-48.

Topel, Robert H. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Seniority." *Journal of Political Economy* 99: 145-176.

Topel, Robert H. and Michael Ward 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics* 107: 439-479.

Townsend, Kiliaen V. 1989. *The Boarding School Guide.* Athens, GA: Agee Publishers, Inc.

U.S. Bureau of the Census, Current Population Reports, Series P-23, No. 121. Private School Enrollment, Tuition, and Enrollment Trends: October 1979. 1982. Washington: U.S. Government Printing Office.

Wachtel, Paul 1976. "The Effect on Earnings of School and College Investment Expenditures." *Review of Economics and Statistics* 58: 326-331.

Wales, Terence 1973. "The Effect of College Quality on Earnings: Results from the NBER-Thorndike Data." *Journal of Human Resources* 8: 306-317.

Wegge, Leon L. 1978. "Constrained Indirect Least Squares Estimators." *Econometrica* 46: 435-449.

Willis, Robert J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions" in Ashenfelter, O. and R. Layard (eds.), *Handbook of Labor Economics, vol. I.* New York: Elsevier.

Willis, Robert J. and Sherwin Rosen 1978. "Education and Self-Selection." *Journal of Political Economy* 79: S7-S36.

Yeager, Robert J., Peter L. Benson, Michael Guerra, and Bruno V. Manno 1985. *The Catholic High School: A National Portrait, vol. I.* Washington, DC: National Catholic Education Association.

TABLE 1

Sample Exclusion Criteria

(Conditions resulting in exclusion from the sample)

Condition Generating Exclusion:	Number of Observations
Full NLSY sample	12,686
Non-response to *private school* and *currently enrolled in high school* questions:	(46)
Not enrolled in grades 9-12 at time of 1979 interview:	(7371)
Inconsistencies in responses to schooling questions, 1979 and later:	(170)
Refusals and invalid skips on religion question(s):	(45)
Missing information concerning: county, state, urban, and SMSA of residence at age 14 (which could not be recovered from matching to 1979 or later responses), or living outside the US at age 14:	(522)
Respondent moved across counties during the period: June of the calendar year of the respondent's 14th birthday to date reported last enrolled in grades 9-12, or did not remain enrolled in same school	(236)
Missing ASVAB test score information:	(129)
Missing/incomplete information necessary to construct labor market work history:	(116)
Had not made transition into the labor force by date of 1989 interview:[1]	(196)[2]
Not interviewed in 1990:	(532)
Out of labor force, 1990:	(508)

[1]I define a labor market transition as occurring during the initial year of a three-year period in which the respondent does not enroll (attend) regular school. Further, the respondent must have labor market activity in at least two of the three years. Exceptions to this criterion are made for respondents for whom the year of entry is coded as 1989. Here, a transition is said to occur if the respondent has not enrolled in regular school since the 1988 interview, and reported labor force activity in both 1989 and 1990. Further details regarding the definition of labor market entry and the distribution of entry years is contained in Data Appendix D.

[2]Of these 196 individuals, 18 were enrolled "continuously" from 1988-1990; sixteen were public school students, and two were private school students.

Sample Exclusion Criteria

(Conditions resulting in exclusion from the sample)

Condition Generating Exclusion:	Number of Observations
Enrolled in school between 1989 and 1990 interviews:	(141)
Missing labor market information in 1990 (occupation, industry, number of jobs held) which could not be imputed from 1989 survey due to job change(s), wage outliers, and/or no reported wage in 1989:	(247)
Lived in state in 1979 with less than 10 respondents in survey	(37)
Missing demographic information in 1990 and inconsistencies in reported information across 1989 and 1990 surveys:	(60)
Number of respondents in selected sample:	2330
Public school attendees:	2211
Private school attendees:	119

TABLE 2

Summary Statistics: Family Background and Demographics

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Attended Private High School	0.065	1	0
Father's years of schooling*[1]	12.002 (3.194)	13.718 (3.172)	11.883 (3.163)
Father's occupation professional or managerial*	0.242	0.360	0.233
Mother's years of schooling*	11.821 (2.443)	12.932 (2.608)	11.744 (2.413)
Mother's occupation professional or managerial*	0.100	0.213	0.092
Mother worked outside household when respondent was age 14	0.572	0.518	0.576
Access to newspaper(s), magazine(s), and library card at age 14	0.535	0.681	0.525
Lived with both parents at age 14	0.784	0.843	0.780
Black	0.130	0.081	0.133
Hispanic	0.053	0.049	0.053
White	0.817	0.870	0.814
Female	0.429	0.524	0.422
Live in SMSA	0.704	0.857	0.694
Live in South*White	0.226	0.252	0.225
Number of siblings	3.106 (2.110)	3.168 (2.117)	3.102 (2.110)
Only child	0.035	0.042	0.034
First-born child (One or more siblings)	0.164	0.132	0.166
"Middle-born" child (Two or more siblings)	0.566	0.489	0.571

[1]Asterisk (*) denotes a variable for which some respondents had missing values. Values were imputed for these observations, and missing value dummies were created. Data Appendix A describes the procedure used to impute missing values.

Summary Statistics: Family Background and Demographics

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Last-born child (One or more siblings)	0.235	0.337	0.229
Net family income in 1979 ($1000s)*	22.176 (12.522)	28.443 (13.308)	21.742 (12.350)
Raised in the Catholic faith[2]	0.321	0.703	0.295
Raised in the Baptist faith	0.210	0.072	0.220
Raised in the Episcopalian faith	0.014	0.003	0.015
Raised in the Lutheran faith	0.081	0.004	0.087
Raised in the Methodist faith	0.087	0.014	0.092
Raised in the Presbyterian faith	0.035	0.030	0.035
Raised in the Jewish faith	0.015	0.006	0.016
Raised in the Pentecostal faith	0.009	0.014	0.009
Raised in the Church of Christ faith	0.021	0.022	0.021
Raised in the Seventh Day Adventist faith	0.002	0.006	0.002
Private high school in county of residence[3]	0.868	0.988	0.860
Number of private high schools per 10,000 population in cty. of residence	0.236 (0.217)	0.282 (0.175)	0.233 (0.219)
Non-sectarian private high school in county of residence	0.595	0.840	0.578
Private HS in cty. of residence with same religious affiliation as respondent	0.402	0.751	0.378

[2]Recorded response to the questions: *1) In what religion were you raised?* (If response is Protestant or Christian, probe:) *2) What denomination was that?* Further details on religious faith and denomination variables are provided in Data Appendix C.

[3]Within the sample, 174 counties (representing 1956 respondents) had private schools located in the county. Of these 174 counties with private schools, 131 counties, comprising 1492 respondents, contained Catholic high schools, and 1432 respondents lived in the 124 counties with non-sectarian private high schools. For more information, see Data Appendix B: Private High School Location and Tuition Information.

Summary Statistics: Family Background and Demographics

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Median Catholic high school tuition and fees in county of residence ($100s)*[4]	10.546 (3.894)	11.470 (4.712)	10.483 (3.824)
Median non-sectarian private HS tuition and fees in cty. of res. ($100s)*	24.837 (8.428)	25.901 (7.541)	24.764 (8.483)
Median other-religious HS tuition and fees in cty. of residence ($100s)*	20.757 (8.398)	21.451 (8.078)	20.710 (8.419)
Median family income in county, 1980 ($1000s)	20.050 (3.709)	20.859 (3.038)	19.994 (3.745)
Median house value in county, 1980 ($1000s)	47.719 (17.441)	52.662 (17.421)	47.377 (17.395)
Pupil-teacher ratio in public schools, state average (1979-1980)	18.739 (1.830)	18.411 (2.084)	18.762 (1.810)
Public teacher annual salary in public schools, state average (1979-1980) ($1000s)	15.941 (2.060)	16.121 (2.008)	15.928 (2.064)
Current per-pupil expenditure in public schools, state average (1979-1980) ($1000s)	2.302 (0.494)	2.368 (0.535)	2.297 (0.491)

[4]There were a total of 2080 private schools in the 174 counties with private schools, and 815 of these reported tuition. For those counties which had no private schools, the state average median county tuition figure for each of the three "private school types" was assigned to the county. In counties which had private schools that did not report tuition information, tuition was imputed from a regression of characteristics of private schools on average median tuition by state and by religious affiliation of school. Further details are contained in Data Appendix B.

TABLE 3

Summary Statistics: Wage Determinants

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Log(hourly wage), 1990[1]	2.260 (0.464)	2.317 (0.405)	2.256 (0.468)
Average annual wage change, entry year to 1990	0.087 (0.343)	0.113 (0.307)	0.085 (0.345)
Number of jobs held, entry year to 1990	8.344 (4.327)	9.134 (4.803)	8.289 (4.289)
Years of schooling	13.360 (2.021)	14.242 (2.111)	13.299 (2.001)
Age, 1990	27.375 (1.230)	27.098 (1.100)	27.394 (1.236)
Actual labor market experience, (years), year of entry to 1990 interview	7.069 (2.378)	6.389 (2.115)	7.116 (2.389)
Wages covered by collective bargaining	0.141	0.132	0.141
Married	0.520	0.442	0.526
Married*Female	0.240	0.266	0.238
One-digit Occupations:[2]			
Managerial or Professional	0.237	0.345	0.229
Technical	0.038	0.071	0.035
Sales	0.093	0.122	0.090
Administrative/Clerical	0.183	0.171	0.184
Service	0.115	0.100	0.116
Crafts	0.149	0.113	0.152

[1]Natural log of hourly rate of pay on current or most recent job, 1990 interview. Wages were symmetrically trimmed with the lowest (0.0025) and highest (0.9975) one-quarter percentiles of the wage distribution excluded in an attempt to minimize the effect of outliers.

[2]One-digit (1980 Census classification) occupations. Excluded one-digit occupation group: Farming, Forestry, and Fishing.

Summary Statistics: Wage Determinants

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Operatives	0.167	0.079	0.174
One-digit Industries:[3]			
Construction	0.077	0.049	0.079
Manufacturing	0.224	0.198	0.225
Transportation and Communication	0.071	0.092	0.070
Wholesale and Retail Trade	0.189	0.130	0.193
Financial, Insurance, and Real Estate	0.081	0.155	0.076
Business and Repair Service	0.088	0.084	0.088
Professional Service	0.160	0.164	0.160

[3]One-digit (1980 Census classification) industries. Excluded industries are: Agriculture, Forestry, and Fishing; Mining; Personal Services; Entertainment and Recreation; and Public Administration.

TABLE 4

Summary Statistics: Achievement and School Quality Measures

Sample Means are weighted to reflect population estimates
(standard deviations of non-categorical variables in parentheses)

Variable Description	All observations (N=2330)	Private School Attendees (N=119)	Public School Attendees (N=2211)
Years of Schooling, 1990	13.360 (2.021)	14.242 (2.111)	13.299 (2.001)
Armed Forces Qualifier Test Score (AFQT Score)[1]	71.020 (18.556)	77.163 (17.252)	70.602 (18.573)
AFQT score, age-adjusted[2]	1.160 (0.304)	1.271 (0.284)	1.153 (0.304)
Type of High School Curriculum last enrolled in during high school:*			
Commercial	0.017	0	0.018
Vocational	0.118	0.055	0.123
General	0.497	0.282	0.512
College Preparatory	0.331	0.632	0.311
Book-pupil ratio in school (books per pupil)*	13.757 (10.016)	19.098 (10.917)	13.388 (9.847)
Proportion of students in school classified as disadvantaged*[3]	0.182 (0.199)	0.074 (0.099)	0.202 (0.158)
Total school enrollment divided by 100	13.324 (5.943)	7.762 (3.607)	13.709 (9.847)
Proportion of students in school classified as minorities	0.193 (0.158)	0.136 (0.137)	0.221 (0.200)
Pupil-teacher ratio (Teachers measured in full-time equivalent units)	19.059 (3.300)	17.919 (3.704)	19.138 (3.253)

[1]Score on Armed Forces Qualifier Test. The AFQT score is the sum of the scores on the following four portions of the Armed Services Vocational Aptitude Battery Test (ASVAB): Paragraph Composition, Word Knowledge, Arithmetic Reasoning, 0.5*(Numerical Operations). The AFQT approximates a standard achievement test, such as the SAT or the ACT.

[2]Raw AFQT score divided by mean score of all individuals in the sample of the same age at the date the test was administered.

[3]Disadvantaged is based on Elementary and Secondary Education Act (ESEA) Guidelines of proportion of students from families with incomes below the poverty level.

TABLE 5

Private School Choice Probit Equations

coefficient estimate
(Heteroskedasticity robust standard errors in parentheses)

Variable Description	Probit Coefficients LIML Estimation	Marginal Effects, Average Derivatives	Marginal Effects, Deriv. of the Average	Marg. Eff., Deriv of the Avg., Avg. Individual
Private high school in county of residence[1]	0.337 (0.434)	0.0286	0.0159 (0.0158)	0.0670 (0.0738)
Number of private high schools per 10,000 pop. in county of residence	0.395**[2] (0.228)	0.0335	0.0190 (0.0149)	0.0800 (0.0682)
Non-sectarian private high school in county of residence	0.326** (0.182)	0.0277	0.0154** (0.0091)	0.0648 (0.0439)
Private religious school of respondent's faith in county of residence	0.489* (0.188)	0.0415	0.0236* (0.0092)	0.0994** (0.0523)
Median Catholic high school tuition and fees in county of residence	0.038* (0.020)	0.0033	0.0018* (0.0008)	0.0077 (0.0047)
Median Non-sectarian high school tuition and fees in county of residence	-0.007 (0.014)	-0.0006	-0.0003 (0.0006)	-0.0013 (0.0028)
Median tuition and fees in other religious high schools in county of res.	-0.005 (0.012)	-0.0004	-0.0002 (0.0005)	-0.0010 (0.0022)
Median family income (1980) in county of residence ($1000s)	-0.045 (0.028)	-0.0038	-0.0022** (0.0013)	-0.0090 (0.0062)
Median house value (1980) in county of residence ($1000s)	0.005 (0.006)	0.0005	0.0003 (0.0003)	0.0011 (0.0012)
Average pupil-teacher ratio in public high schools in state of residence	-0.098 (0.090)	-0.0083	-0.0047 (0.0039)	-0.0197 (0.0124)
Average annual salary for public school teachers in state of residence	0.186 (0.124)	0.0158	0.0089 (0.0059)	0.0374** (0.0215)

[1]All models include: eight indicators for Census region of residence at age 14, indicators for parents' education level missing, indicator for mother working at age 14 response missing, and an indicator for private school tuition in the county of residence imputed from state average tuition for private schools of same religious affiliation. Data are weighted by sample weights to adjust for the inclusion of members of the supplemental (poverty) subsample in sample.

[2]Asterisks denote the significance of the estimated coefficient as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

Private School Choice Probit Equations

coefficient estimate
(Heteroskedasticity robust standard errors in parentheses)

Variable Description	Probit Coefficients LIML Estimation	Marginal Effects, Average Derivatives	Marginal Effects, Deriv. of the Average	Marg. Eff., Deriv of the Avg., Avg. Individual
Average *current* per pupil expenditure in public schools in state of residence	-0.936** (0.489)	-0.0795	-0.0450** (0.0243)	-0.1892** (0.1032)
Siblings	0.061** (0.033)	0.0052	0.0029** (0.0015)	0.0123 (0.0076)
Only child	0.514 (0.342)	0.0436	0.0246** (0.0142)	---
First-born child (1 or more siblings)	-0.068 (0.232)	-0.0057	-0.0033 (0.0092)	---
Middle-born child (2 or more siblings)	-0.226 (0.189)	-0.0192	-0.0106 (0.0081)	-0.0446 (0.0360)
Last-born child (1 or more siblings)	0.356* (0.140)	0.0302	0.0169* (0.0064)	---
Raised Catholic	0.358 (0.244)	0.0304	0.0163 (0.0102)	0.0685 (0.0500)
Raised Baptist	-0.486* (0.247)	-0.0412	-0.0242* (0.0116)	---
Raised Episcopalian	-1.161* (0.511)	-0.0985	-0.0563 (0.0356)	---
Raised Lutheran	-1.167* (0.444)	-0.0991	-0.0564* (0.0245)	---
Raised Methodist	-0.557 (0.383)	-0.0473	-0.0272** (0.0149)	---
Raised Presbyterian	-0.110 (0.329)	-0.0094	-0.0059 (0.0146)	---
Raised Jewish	-0.985 (0.616)	-0.0836	-0.0475** (0.0248)	---
Raised Pentecostal (includes Assemblies of God)	0.988* (0.469)	0.0839	0.0460* (0.0207)	---
Raised Church of Christ (United)	-0.011 (0.393)	-0.0009	-0.0012 (0.0163)	---
Raised Seventh Day Adventist	0.639 (0.615)	0.0542	0.0342 (0.0580)	---

Private School Choice Probit Equations

coefficient estimate
(Heteroskedasticity robust standard errors in parentheses)

Variable Description	Probit Coefficients LIML Estimation	Marginal Effects, Average Derivatives	Marginal Effects, Deriv. of the Average	Marg. Eff., Deriv of the Avg., Avg. Individual
Black	0.370** (0.204)	0.0314	0.0190** (0.0114)	---
Hispanic	0.112 (0.200)	0.0095	0.0056 (0.0119)	---
Female	0.214** (0.113)	0.0181	0.0101* (0.0050)	---
Lived in SMSA at age 14	0.130 (0.214)	0.0110	0.0061 (0.0090)	0.0256 (0.0391)
(Lived in South at age 14)*White	0.602* (0.245)	0.0511	0.0297* (0.0145)	---
Access to newspaper(s), magazine(s), and library card at age 14	0.101 (0.126)	0.0085	0.0047 (0.0053)	0.0196 (0.0243)
Lived with both parents at age 14	0.036 (0.143)	0.0030	0.0019 (0.0067)	0.0078 (0.0286)
Father's years of schooling	0.046* (0.022)	0.0039	0.0022* (0.0010)	0.0094** (0.0053)
Mother's years of schooling	0.088* (0.031)	0.0075	0.0042* (0.0014)	0.0176* (0.0088)
Mother worked outside home when respondent was 14	-0.114 (0.116)	-0.0097	-0.0052 (0.0050)	-0.0220 (0.0216)
Net family income in 1979, ($1000s)	0.005 (0.005)	0.0004	0.0002 (0.0002)	0.0010 (0.0094)
Constant	-3.290* (1.402)	---	---	---
Log-likelihood	-419.506	---	---	---
$\chi^2(k)$ statistic	$\chi^2(52)=100$ *(p=0.000)*	---	---	---
Number of observations	N=2330	N=2330	N=2330	N=2330

Note to Table 5: Probit Marginal Effects Estimates

In Table 5, two alternative marginal effects estimates are reported: the average derivative and the derivative of the average. Both of the marginal effects estimators measure the change in the value of the index function, given a change in one of the regressors. The average derivative derives a measure of the marginal effect by taking the mean over all individuals of the derivative of the index function.

Average Derivative Marginal Effects Estimator for the *kth* regressor:

$$\sum_{i=1}^{N} N^{-1} \frac{\partial \Phi(Z_i'\gamma)}{\partial Z_{ik}} = \sum_{i=1}^{N} N^{-1} \phi(Z_i'\gamma)\gamma_k .$$

The derivative of the average marginal effects estimator also measures the change in the index function due to a change in one of the regressors, but evaluates the derivative at the vector of means of the regressors.

Derivative of the Average Marginal Effects Estimator for the *kth* regressor:

$$\frac{\partial \Phi(Z_i'\gamma)}{\partial Z_{ik}} | Z_i = \bar{Z} = \phi(Z_i'\gamma)\gamma_k | Z_i = \bar{Z} .$$

Standard errors for the average derivative marginal effects estimates require the use of bootstrapping techniques. For the derivative of the average estimates, however, one can derive standard errors following a linear approximation approach, discussed in Greene [1992], which is listed below.

Asymptotic Variance of the Derivative of the Average Marginal Effects:

$$Asymptotic\ variance[\frac{\partial \Phi(Z_i'\gamma)}{\partial Z_i}] = [\phi(Z_i'\gamma)I - \phi(Z_i'\gamma)[\gamma' Z_i][\gamma Z_i']]Var[\gamma][\phi(Z_i'\gamma)I - \phi(Z_i'\gamma)[\gamma' Z_i][\gamma Z_i']] .$$

TABLE 6

Educational Attainment (Schooling) Equations, One-Population Models

(Dependent variable: Years of schooling completed, 1990)

Variable Description	No Private School Control	Private School Indicator, Exogenous	Two-Stage, Private School Endogenous
Attended private high school[1]	---	-0.143	0.065[2]
		(0.186)	(0.095)
Net family income in 1979 ($1000s)	0.015*[3]	0.015*	0.015*
	(0.004)	(0.004)	(0.004)
Respondent's age (in years), 1990	0.019	0.019	0.019
	(0.032)	(0.032)	(0.032)
Father's years of schooling	0.066*	0.067*	0.066*
	(0.019)	(0.019)	(0.019)
Father employed in professional or managerial occupation, 1979	0.552*	0.549*	0.550*
	(0.122)	(0.122)	(0.122)
Mother's years of schooling	0.093*	0.093*	0.093*
	(0.021)	(0.021)	(0.021)
Mother employed in professional or managerial occupation, 1979	0.415*	0.423*	0.420*
	(0.163)	(0.163)	(0.163)
Mother worked outside home when R was 14	-0.357	-0.354	-0.354
	(0.221)	(0.223)	(0.222)
Access to newspaper(s), magazine(s), and library card at age 14	0.390*	0.392*	0.391*
	(0.090)	(0.090)	(0.090)
Lived with both parents at age 14	0.334*	0.333*	0.333*
	(0.109)	(0.109)	(0.109)

[1]All regression specifications include: eight indicators for Census region of residence at age 14; indicators for parents' education level or occupation missing; indicator for mother working when R was 14 response missing; median family income and house value in one's county of residence, and controls for birth-order.

[2]All standard errors in the table above are heteroskedasticity-robust, and the standard errors reported for the two-stage model of educational attainment, which includes controls for endogenous selection into private schools, are adjusted to account for the use of the estimated selection-correction terms in the second-stage schooling equation.

[3]Asterisks denote the significance of the estimated coefficient as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

Educational Attainment (Schooling) Equations, One-Population Models

(Dependent variable: Years of schooling completed, 1990)

Variable Description	No Private School Control	Private School Indicator, Exogenous	Two-Stage, Private School Endogenous
Number of siblings	-0.060*	-0.059*	-0.060*
	(0.021)	(0.021)	(0.021)
Female	0.299*	0.301*	0.298*
	(0.081)	(0.082)	(0.081)
Black	0.379*	0.380*	0.381*
	(0.119)	(0.119)	(0.119)
Hispanic	0.260**	0.263**	0.260**
	(0.139)	(0.139)	(0.139)
Pupil-teacher ratio in public schools, state avg. 1979-80	0.037	0.037	0.037
	(0.059)	(0.059)	(0.059)
Public teacher annual salary in public schools, state avg. 1979-80, ($1000s)	0.040	0.044	0.042
	(0.081)	(0.081)	(0.081)
Current per-pupil expenditure in public schools, state avg. 1979-80, ($1000s)	-0.017	-0.029	-0.022
	(0.324)	(0.324)	(0.324)
Enrolled in a college prep curriculum in high school	1.374*	1.382*	1.381*
	(0.101)	(0.102)	(0.102)
Proportion of enrollment in high school classified as economically disadvantaged	-0.269	-0.313	-0.303
	(0.290)	(0.293)	(0.292)
Proportion of enrollment in high school classified as minorities	-0.274	-0.271	-0.274
	(0.231)	(0.231)	(0.231)
Book-pupil ratio (books per pupil)	0.002	0.002	0.002
	(0.005)	(0.004)	(0.005)
Pupil-teacher ratio (Teachers measured in full-time equivalent units)	-0.003	-0.002	-0.002
	(0.014)	(0.014)	(0.014)
High school enrollment divided by 100	-0.011	-0.013	-0.012
	(0.008)	(0.009)	(0.009)
Adjusted R^2	0.346	0.346	0.346
Number of Observations	N=2330	N=2330	N=2330

TABLE 7

Educational Attainment (Schooling) Equations, Two-Population Models

(Dependent variable: Years of schooling completed, 1990)

Variable Description	OLS		Two-stage	
	Private	Public	Private	Public
Selection correction terms:[1]				
Private school schooling eq.: $\rho_{S_{pr}u}\sigma_{S_{pr}}$	---	---	-0.124[2] (0.379)	---
Public school schooling eq.: $\rho_{S_{pu}u}\sigma_{S_{pu}}$	---	---	---	-0.604 (0.392)
Net family income in 1979 ($1000s)	0.006 (0.015)	0.014*[3] (0.005)	0.008 (0.014)	0.015* (0.005)
Respondent's age (in years), 1990	0.281* (0.119)	0.003 (0.033)	0.288* (0.116)	0.003 (0.033)
Father's years of schooling	-0.011 (0.069)	0.076* (0.019)	-0.007 (0.070)	0.083* (0.019)
Father employed in professional or managerial occupation, 1979	0.693 (0.437)	0.607* (0.127)	0.708** (0.439)	0.606* (0.127)
Mother's years of schooling	0.308* (0.096)	0.075* (0.022)	0.310* (0.097)	0.080* (0.022)
Mother employed in professional or managerial occupation, 1979	0.746** (0.404)	0.444* (0.175)	0.742** (0.403)	0.448* (0.174)
Mother worked outside home when R was 14	0.808 (0.796)	-0.488* (0.217)	0.811 (0.793)	-0.509* (0.219)
Access to newspaper(s), magazine(s), and library card at age 14	1.457* (0.294)	0.368* (0.093)	1.454* (0.295)	0.384* (0.093)

[1]All regression specifications include: eight indicators for Census region of residence at age 14; indicators for parents' education level or occupation missing; indicator for mother working when R was 14 response missing; median family income and house value in one's county of residence; and controls for birth-order.

[2]All standard errors are both heteroskedasticity-robust and adjusted to account for the use of the estimated selection-correction terms in the second-stage years of schooling equation.

[3]Asterisks denote the significance of the estimated coefficient as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

Educational Attainment (Schooling) Equations, Two-Population Models

(Dependent variable: Years of schooling completed, 1990)

Variable Description	OLS		Two-stage	
	Private	Public	Private	Public
Lived with both parents at age 14	-0.300 (0.543)	0.364* (0.112)	-0.351 (0.558)	0.370* (0.112)
Number of siblings	-0.105 (0.088)	-0.055* (0.021)	-0.102 (0.090)	-0.046* (0.021)
Female	0.483** (0.288)	0.292* (0.084)	0.490** (0.290)	0.321* (0.086)
Black	0.029 (0.592)	0.419* (0.123)	0.018 (0.599)	0.409* (0.123)
Hispanic	0.003 (0.502)	0.260** (0.146)	-0.003 (0.501)	0.285* (0.145)
Pupil-teacher ratio in public schools, state avg. 1979-80	-0.207 (0.184)	0.049 (0.061)	-0.214 (0.187)	0.054 (0.061)
Public teacher annual salary in pub. schools, state avg. 1979-80, ($1000s)	0.654* (0.254)	0.029 (0.083)	0.650* (0.251)	0.037 (0.083)
Current per-pupil expend. in public schools, state avg. 1979-80, ($1000s)	-2.457* (0.937)	0.053 (0.333)	-2.444* (0.922)	0.010 (0.334)
Enrolled in a college prep curric. in high school	1.426* (0.301)	1.352* (0.105)	1.427* (0.303)	1.347* (0.105)
Proportion of enrollment in HS classified as econ. disadvantaged	-0.110 (1.296)	-0.349 (0.298)	-0.089 (1.300)	-0.351 (0.298)
Proportion of enrollment in HS classified as minorities	2.645** (1.582)	-0.282 (0.234)	2.672** (1.605)	-0.261 (0.233)
Book-pupil ratio (books per pupil)	-0.008 (0.010)	0.002 (0.005)	-0.008 (0.010)	0.002 (0.005)
Pupil-teacher ratio	0.017 (0.039)	-0.008 (0.014)	0.017 (0.039)	-0.008 (0.014)
High school enrollment divided by 100	0.055 (0.054)	-0.012 (0.009)	0.057 (0.054)	-0.011 (0.009)
Adjusted R^2	0.556	0.330	0.550	0.332
Number of Observations	N=119	N=2211	N=119	N=2211

TABLE 8

Log-Wage Equations: One-Population Models

Dependent variable: Log(hourly wage, 1990)

Variable Description	OLS, no private school controls	OLS with private school indicator (exogenous)	"Two-stage" estimation with selection correction term
Private	---	-0.091*[1]	0.052*[2]
		(0.037)	(0.020)[3]
Years of schooling[4]	0.068*	0.067*	0.067*
	(0.007)	(0.007)	(0.007)
Years of experience, 1990	0.013	0.014	0.014
	(0.020)	(0.020)	(0.020)
Yrs. of experience squared, 1990	0.0018	0.0018	0.0018
	(0.0014)	(0.0014)	(0.0014)
Married	0.056*	0.054*	0.054*
	(0.024)	(0.024)	(0.024)

[1]Asterisks denote the significance of the estimated coefficient as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

[2]The selection correction used was the standard conditional mean correction used when both "participants" and "non-participants" appear in the second-stage estimation. Let h_i denote the selection correction term, such that:

$$h_i = \frac{-\phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)} \quad \text{if } Private=1, \qquad h_i = \frac{\phi(Z_i'\gamma)}{[1-\Phi(Z_i'\gamma)]} \quad \text{if } Private=0,$$

where $Z_i'\gamma$s are the fitted values from the school choice probit, ϕ is the standard normal pdf, and Φ is the standard normal cdf. Note that the opposite of $[\phi(\cdot)/\Phi(\cdot)]$ was entered as the selection correction term for participants, so that the restriction imposed by this estimation is $\rho_{Nu}\sigma_N=\rho_{Pu}\sigma_P$.

[3]All standard errors reported in Table 8 are heteroskedasticity robust. Further, the standard errors in the third column, the model of self-selection, have been corrected to account for the fact that the $Z_i'\gamma$s are estimated in the first stage. The number of private high schools per 10,000 population, the presence of a non-sectarian private high school in one's county of residence at age 14, and the presence of a private high school of one's religious affiliation are used as instruments to identify the school choice and log-wage equations.

[4]Years of schooling are treated as endogenous, and the regressions reported instrument for educational attainment using: indicator for access to reading materials in the home; indicator for live with both parents at age 14; and number of siblings.

Log-Wage Equations: One-Population Models

Dependent variable: Log(hourly wage, 1990)

Variable Description	OLS, no private school controls	OLS with private school indicator (exogenous)	"Two-stage" estimation with selection correction term
Married*Female	-0.072*	-0.070*	-0.071*
	(0.036)	(0.035)	(0.035)
Female	-0.114*	-0.113*	-0.115*
	(0.027)	(0.027)	(0.027)
Black	-0.085*	-0.085*	-0.084*
	(0.026)	(0.026)	(0.026)
Hispanic	-0.034	-0.032	-0.034
	(0.034)	(0.034)	(0.034)
Wages covered by collective bargaining	0.165*	0.165*	0.165*
	(0.026)	(0.026)	(0.026)
Adjusted R^2	0.386	0.388	0.388
Number of observations	2330	2330	2330

Note to Table 8: All of the regressions reported above included:three census region indicators for 1990 residence; seven indicators for one-digit occupation in which R was employed in 1990; seven indicators for one-digit industry in which one was employed in 1990; firm-size and multi-plant employer measures; parents' education and occupational status, parents' (family) income in 1979, and mother working at age 14; indicators for religious faith in which one was raised; median family income and median house value in county of residence at age 14; and five measures of quality of high school which respondent attended: book-pupil ratio, proportion of school economically disadvantaged, proportion of school minorities, pupil-teacher ratio in high school, total enrollment, and an indicator variable if R was enrolled in a college prep curriculum during high school.

TABLE 9

Log-Wage Equations: Public and Private School Attendees
Two-Population Model, Two-Stage Estimation

(Dependent variable: Log(hourly wage, 1990))

Variable Description	Educational Attainment Exogenous		Educational Attainment Endogenous	
	Private	Public	Private	Public
Mean of Log-wage	2.317	2.256	2.317	2.256
Selection Correction Terms:[1]				
$\rho_{Nu}\sigma_N$	-0.057[2] (0.077)	---	-0.086 (0.087)	---
$\rho_{Pu}\sigma_P$	---	0.189*[3] (0.085)	---	0.160** (0.088)
Years of schooling	0.097* (0.040)	0.066* (0.008)	0.131* (0.038)	0.066* (0.008)
Experience in yrs., 1990	-0.050 (0.072)	0.018** (0.021)	-0.045 (0.074)	0.020 (0.020)
Experience squared	0.004 (0.005)	0.001 (0.001)	0.005 (0.005)	0.001 (0.001)
Married	0.300* (0.070)	0.042** (0.025)	0.264* (0.089)	0.044** (0.025)
Married*Female	-0.370* (0.117)	-0.060 (0.037)	-0.372* (0.107)	-0.061** (0.036)

[1]All of the log-wage equations reported above include: three Census region dummies for 1990 residence; indicator for living in an SMSA in 1990; seven indicators for one-digit industry in which respondent was employed in 1990; seven indicators for one-digit occupation in which respondent was employed in 1990; parents' education and occupational status, family income, and mother worked; religious faith indicators; and median family income and house value in county of residence at age 14. Private high school density and indicators for non-sectarian private high school and private high school of one's religious affiliation were used as instruments for the private school attendance decision. Access to reading materials, living with both parents at age 14, and number of siblings were used as instruments for educational attainment. Overidentification tests are reported in Table 11.

[2]Standard errors are corrected for both the heteroskedasticity introduced by the inclusion of the selection correction terms in the log-wage equation and the fact that the selection correction terms are estimates of the true selection-correction measures.

[3]Asterisks denote the significance of the estimated coefficients as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

Log-Wage Equations: Public and Private School Attendees
Two-Population Model, Two-Stage Estimation

(Dependent variable: Log(hourly wage, 1990))

Variable Description	Educational Attainment Exogenous		Educational Attainment Endogenous	
	Private	Public	Private	Public
Female	0.028 (0.082)	-0.139* (0.028)	-0.017 (0.081)	-0.132* (0.028)
Black	-0.030 (0.118)	-0.097* (0.026)	0.001 (0.105)	-0.088* (0.027)
Hispanic	-0.044 (0.125)	-0.034 (0.036)	0.038 (0.113)	-0.037 (0.036)
Wages covered by collective bargaining	0.201* (0.089)	0.179* (0.026)	0.182* (0.082)	0.174* (0.026)
Father's years of schooling	0.025 (0.016)	0.002 (0.004)	0.029* (0.014)	0.002 (0.004)
Father employed in prof. or mangrl. occ.	0.086 (0.094)	0.071* (0.029)	0.009 (0.105)	0.067* (0.028)
Mother's years of schooling	0.047* (0.018)	0.003 (0.005)	0.044* (0.019)	0.003 (0.005)
Mother employed in prof. or mangrl. occ.	0.081 (0.094)	0.048 (0.035)	0.075 (0.102)	0.045 (0.035)
Family income in 1979, ($1000s)	-0.002 (0.004)	0.004* (0.001)	-0.006 (0.004)	0.004* (0.001)
Enrolled in college prep curriculum in HS	0.057 (0.069)	0.057* (0.023)	0.054 (0.073)	0.058* (0.023)
Prop. of school enrlmt. econ. disadvantaged	-0.027 (0.301)	-0.086 (0.063)	-0.050 (0.307)	-0.086 (0.064)
Prop. of school enrlmt. minorities	0.032 (0.287)	-0.008 (0.051)	0.116 (0.275)	-0.027 (0.053)
Pupil-teacher ratio	0.021 (0.087)	-0.003 (0.003)	0.013 (0.084)	-0.003 (0.003)
School enrollment/100	-0.017 (0.014)	0.002 (0.002)	-0.011 (0.012)	0.002 (0.002)
Book-pupil ratio	0.005* (0.002)	0.001 (0.001)	0.004 (0.003)	0.001** (0.0004)
Adjusted R^2	0.410	0.387	0.392	0.390
Number of observations	N=119	N=2211	N=119	N=2211

TABLE 10

Estimates of Private-Public Wage Differential

Parameter Estimate
(standard error in parentheses)

Description of Measure of Private School Effect or Differential	Estimate of Differential
Two population models: private-public differential evaluated as wage differential accruing to the "average" public school student receiving the treatment effect of a private high school education: $\bar{X}_P' \cdot [\hat{\beta}_N - \hat{\beta}_P]$	
Two-population model, Two-stage estimation:	
Private-public differential evaluated at mean characteristics of avg. public school student:	0.133 (0.182)
Proportion of differential attributable to additional education (+1.457 years):	71.00%
Private-public differential evaluated at mean characteristics of public school student with 12 years of schooling (high school graduate):	0.037 (0.197)
Proportion of differential attributable to additional education (+0.39 years):	67.87%
Private-public differential evaluated at mean characteristics of public school student with 16 years of schooling (college graduate):	0.122 (0.209)
Proportion of differential attributable to additional education (+1.36 years):	72.61%

TABLE 11

Overidentification Tests: Instruments Identifying System of School Choice, Educational Attainment, and Log-Wage Equations

Set of instruments:

Private School Choice and Log-Earnings:

private high schools per 10,000 pop. in county (phsdense),
non-sectarian private high school in county (nspskol),
private religious high school in county of one's faith (psrpref).

Educational Attainment and Log-Earnings:

access to newspapers, magazines, and library card at age 14 (accesrms),
lived with both parents at age 14 (bthpar14),
number of siblings (siblings).

	Exclusion Restrictions Tested		
	$\chi^2_{0.05}(4)=9.49$		
Model	phsdense, psrpref, bthpar14, siblings	nspskol, psrpref, bthpar14, siblings	phsdense, nspskol, bthpar14, siblings
Two-stage estimation, two population model:			
Public school attendee log-wage equation	4.841	5.002	4.077
Private school attendee log-wage equation	3.760	4.318	3.419

Note to Table 10: The test of overidentifying restrictions is that of Wegge [1978]. The test statistic is distributed $\chi^2(k)$, where k is equal to the number of overidentifying restrictions, in this case, four. The Wegge test treats the overidentified log-wage equation as an exactly-identified model that is subject to overidentifying restrictions. In the model above, the null hypothesis is that the coefficients on the phsdense, nspskol, bthpar14, and siblings measures in the log-wage equations are zero. Note that the overidentification restrictions cannot be rejected in any of the models.

TABLE 12

Log-Wage Change Equations: Two-Population Model

Dept. variable: [Log(hourly wage, 1990) - Log(entry year hourly wage)]/years in labor market

Variable Description	Two-stage estimation	
	Private	Public
Avg. annual log-wage change, entry year-1990[1,2]	0.113 (0.307)	0.085 (0.345)
Selection correction terms:		
Private: $\rho_{Nu}\sigma_N$	-0.140**[3] (0.082)	---
Public: $\rho_{Pu}\sigma_P$	---	-0.171** (0.091)
Years of schooling	0.075* (0.030)	-0.003 (0.007)
Experience in years, 1990	-0.081 (0.083)	-0.015 (0.019)
Experience squared	0.007 (0.007)	0.001 (0.001)
Number of jobs, entry year to 1990	-0.007 (0.008)	0.005* (0.002)
Female	-0.082 (0.083)	-0.021 (0.028)
Black	0.156 (0.100)	-0.022 (0.027)
Hispanic	0.129 (0.104)	-0.002 (0.033)
Adjusted R^2	0.069	0.014
Number of observations	N=119	N=2211

[1]Both regression specifications include: indicators for year of entry into labor force; parents' education, occupation, and income; median family income and house value in county of residence at age 14; and school quality measures. Instruments used were the same as those for the wage level equations.

[2]Standard errors are heteroskedasticity-robust, and are corrected to account for the first-stage estimation of the selection correction terms.

[3]Asterisks denote the significance of the estimated coefficients as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the 10% level.

TABLE 13

Returns to Education Estimates: Controls for Ability and School Quality Measures

Two-Population Model, Two-Stage Estimation
Dependent Variable: Log(hourly wage, 1990)

Var. Name	Baseline Model		Baseline, Family Bkgrd.		Baseline, Family Bkgrd., Cty Demog.		Baseline, Family Bkgrd., Cty Demog., School Qual.,		Baseline, Family Bkgrd., Cty Demog., School Qual., AFQT score	
	Private	Public	Private	Public	Private	Public	Private	Public	Private	Public
$\rho_{Nu}\sigma_N$	-0.018[1] (0.061)[2]	---	-0.053[3] (0.086)	---	-0.067 (0.086)	---	-0.086 (0.087)	---	-0.156** (0.081)	---
$\rho_{Pu}\sigma_P$	---	0.257* (0.082)	---	0.178* (0.088)	---	0.157** (0.088)	---	0.160** (0.088)	---	0.152** (0.092)
Yrs. of school	0.131* (0.039)	0.066* (0.008)	0.130* (0.038)	0.066* (0.008)	0.130* (0.038)	0.066* (0.008)	0.131* (0.038)	0.066** (0.008)	0.145* (0.041)	0.058* (0.008)
AFQT score	---	---	---	---	---	---	---	---	0.403* (0.167)	0.203* (0.042)
Obs.	119	2211	119	2211	119	2211	119	2211	119	2211
Adj. R^2	0.382	0.374	0.430	0.384	0.407	0.388	0.392	0.390	0.435	0.398

[1]The baseline specification is the two-population model estimated with two-stage techniques, and includes: experience, experience squared, union indicator, married indicator, married interacted with Female, Black indicator, Hispanic indicator, indicator for living in SMSA in 1990, three Census region dummies for 1990 residence, ten indicator variables for religious faith in which one was raised, seven indicator variables for the one-digit industry in which one was employed in 1990, and seven indicators for one-digit occupation in which respondent was employed in 1990.

[2]All standard errors are corrected for both the heteroskedasticity introduced by the inclusion of the selection-correction terms in the regression equations and the fact that the selection-correction terms are estimates of the true selection-correction terms.

[3]Asterisks denote the significance of the estimated coefficients as follows: a single asterisk (*) denotes significance at the five percent level; two asterisks (**) denote significance at the ten percent level.

Returns to Education Estimates: Years of Schooling Entered as a Fully-Additive Spline Function

Two-Population Model, Two-Stage Estimation
Dependent Variable: Log(Hourly wage, 1990)

Variable Description	Baseline Model, Family Background, County Demographics, School Quality		Baseline Model, Family Background, County Demographics, School Quality, AFQT Score	
	Private	Public	Private	Public
$\rho_{Pu}\sigma_P$	---	0.158**	---	0.155**
		(0.087)		(0.091)
$\rho_{Nu}\sigma_N$	-0.096	---	-0.162*	---
	(0.087)		(0.083)	
Years of schooling, high school (grades 9-12)	0.162*	0.051*	0.163*	0.045*
	(0.052)	(0.012)	(0.055)	(0.013)
Years of schooling, post-secondary (13+ years)	0.151*	0.056*	0.156*	0.049*
	(0.044)	(0.010)	(0.047)	(0.010)
AFQT score	---	---	0.393*	0.201*
			(0.159)	(0.042)
Adjusted R^2	0.387	0.391	0.426	0.398
Number of observations	119	2211	119	2211