Evaluating Private School Quality in Denmark

Beatrice Schindler Rangvid*

Preliminary version: February 2002

Abstract

We estimate the effect of attending private rather than public school on educational attainment measures. We suggest a new instrument variable reflecting the potential dissatisfaction with the local public school offer and thus the potential demand for private schooling. This instrument is especially designed to capture a part of the variation in demand for private schooling that is not related to educational outcomes. We find the instrument to be strong and (informal) tests show that it can be validly excluded from the outcome equation. Intuitively, the validity of the instrument is probably due to peer sorting across either school types or residential location. Preliminary results of a two-stage least squares model (as a linear probability model) do not indicate any significant private school effects for either outcome measures.

*The Aarhus School of Business, Fuglesangs Allé 4, 8210 Århus V and Institute of Local Government Studies, Nyropsgade 37, DK-1602 Copenhagen V, Denmark. Phone: (45) 3311 0300, fax: (45) 3315 2875, and e-mail: bs@akf.dk.

This work is part of the research of the Graduate School for Integration, Production and Welfare. Financial support from the Danish Social Science Research Council is gratefully acknowledged.

1 Introduction

There is substantial controversy regarding private school effects on academic outcomes. Critics claim that seemingly positive private school effects could be the result of selection rather than causation. In the existing studies trying to correct for selection into the private school sector, the predominantly used identification variables are related to religious affiliation (e.g., Evans and Schwab 1995; Sander and Krautmann 1995; Sander 1996, 1999, 2000; Dee 1998; Neal 1997). In the literature of private-public school choice, in contrast, typically the approach is to use variables describing the local population composition (in addition to socio-economic family background) to explain choice of school (see Downes and Greenstein 1996, Epple and Romano 1998, 2000 and Benabou 1993 for theoretical models and Lankford, Lee and Wyckoff 1995; Lankford and Wyckoff 1992, 2000 for empirical analyses). In the present paper, we suggest an instrument variable[1], constructed especially for the purpose of capturing a very specific part of the variation in private school attendance that is not related to the educational outcome. The proposed instrument variable is shown to satisfy the requirements for validity: on the one hand, it is significantly related to school choice, and on the other hand, informal tests[2] show that the instrument is transmitted to educational outcomes only through choice of school.

To estimate the model, we use data from Danish administrative registers. Private comprehensive schools (combining primary and lower secondary school) are heavily subsidized in Denmark (on an average, 85% of the expenditures are covered by subsidies[3]), but no thorough assessment of school quality has been conducted, yet. Over the years, several examples of inferior private schools have become known. Recently some arabic private schools have become the focus of public attention. On the other hand, a study conducted for a major Danish newspaper suggests that private schools as a whole are superior to public schools[4]. The present study tries to give some answers to these conflicting results.

Most of the existing studies rely on test scores as the outcome of interest, but measures of post-secondary educational attainment are potentially more important indicators of school quality in the light of the economic consequences of obtaining more education. In this paper, we consider two measures of post-secondary educational achievement: (i)

[1] We estimate a linear probability model as suggested by Angrist (1991, 2001) and Angrist and Krueger (2001).

[2] Formal testing is not possible in my just-identified model.

[3] There is a fixed state subsidy of 75% of average public school expenditures and, on average, local authorities subsidies private schools with another 10%.

[4] Published in Jyllandsposten København (May, 27th/ 31st and June, 1st 2001) who report that private school students are superior with respect to both educational attainment and wage, even after controlling for the parents' social status. However, their estimates are not corrected for selection bias.

having passed upper secondary school or finished a vocational education at age 25 and (ii) being enrolled in or having finished college/university at age 25. Preliminary work on this paper suggests that the instrument works best in urban areas. We therefore restrict the analysis to the metropolitan areas of the three biggest cities (Copenhagen, Århus and Odense) and other bigger cities, comprising about 2/3 of all students in the sample.

Literature The central challenge in estimating the effects of private schools is to control for unmeasured factors that may be correlated with both educational outcome and the family's choice of school sector. Therefore, in the recent literature, school choice is treated as an endogenous determinant of educational outcomes. Most of the existing research on the effect of private schools relies on variables related to religious affiliation as identifying restrictions. Recently, variation in state laws reflecting the unionization of the public sector have been used to identify private school choice (Figlio and Stone 1999).

Results of the efforts to measure the quality of private schooling in the existing literature are mixed. The available evidence seems to suggest that private schools may on average produce slightly better educational outcomes than public schools - particularly for students with only low-quality public alternatives and students least likely to choose private school (Neal 1997; Figlio and Stone 1999; Evans and Schwab 1995; Sander 1996). Only a few papers use educational attainment as indicator of school quality, e.g. Sander and Krautmann 1995; Neal 1997; Evans and Schwab 1996. The existing literature is exclusively on US data, except for one study on British data (Wright 1999). The present study extends the discussion to Danish data.

Private school sector in Denmark In Denmark, over the period covered by our data (1980-1985 for information about type of school attended), about 9.5 percent of all primary and lower secondary students attend private schools. Private school students are geographically concentrated in urban areas, especially in the capital's inner city where 14 percent of students attend private schools. In the remaining metropolitan area of the capital and the other greater cities the number is down to an average of 8 percent, whereas in rural areas only 7 percent attend private school. The difference between the parents' educational background of private and public school students (we will use this measure as a proxy for school quality later on) is greatest in urban areas, too. In inner-city Copenhagen, the average parental education of private school students is two years higher than for their public school peers, in contrast to a one-year difference for the country as a whole, indicating stronger selection by socio-economic background in the capital center.

Selection issues Private school students have better educational results[5]. The share of students who have passed upper secondary school or taken a vocational degree at age 25 is significantly greater for private schools than for public schools (81% and 76%, respectively). The difference of the share of students enrolled in or having finished college is even more pronounced (39% and 29%, respectively). Thus, according to the raw data, educational attainment for private school students is higher than for public school students. But private and public school students differ in systematic ways. Superior results from private school students may be explained by the more advantageous parental background and ability of private school students. The summary statistics show that while public school and private school students are similar with respect to individual characteristics (gender, ethnicity), they come from different socio-economic backgrounds. The most significant differences in means are for parents' educational attainment. Parents of private school students have a considerably higher education than parents of public school students. Thus, output quality might not be due to better schools, but due to "higher quality" inputs (students) in private schools. As the summary statistics show, there is important selection on observables in the data which we account for by including a rich set of control variables in the empirical analysis.

However, observable characteristics are unlikely to fully capture the nonrandom differences in sector selection. If so, then models that do not account for this nonrandom selection will lead to biased estimates of the treatment effect of private schools. Consistent estimates can be obtained using an instrumental-variables procedure, assuming we can find suitable instruments. The instrument variable estimate is consistent if the identifying instrument is correlated with sector selection, but any effect on educational attainment is transmitted only through sector selection.

Contribution In this paper, we use the findings of the school choice literature and exploit the heterogeneity in the socio-economic composition of the population as exogenous variation for school choice. The idea is that these community-level variables proxy for local public school quality which has been shown to affect the demand for private schooling (e.g. Lankford and Wyckoff 1992; Lankford, Lee and Wyckoff 1995; West and Palsson 1988; Downes and Greenstein 1996; Goldhaber 1996). In this paper, we construct the instrument in a way, such that it captures only a part of the variation in private school attendance, which, as is shown, is uncorrelated with students educational outcome. The suggested variable turns out to be a strong and valid instrument for private school attendance.

[5]The variable means are shown along with the significance level of a t-test for identical means in Table 1 in the appendix.

Our instrument exploits variation across municipalities in the potential amount of parental dissatisfaction with the local public school. Basically, there are two main reasons for parents to choose private schools: (i) the public school the family is assigned to by their residential location may be of "bad quality" (as perceived by the family and relative to the family's preferences) and (ii) the family has a preference for private schooling (because of a special curriculum or pedagogy offered by private schools). We cannot use variables for the individual families preference for schooling as they tend to be related to educational outcome as well. Thus, we base the search for an instrument on (i), i.e.

we construct an instrument on the basis of the families' potential dissatisfaction with the local public school. Such an instrument can be constructed either on an individual basis (i.e. the individual family's degree of dissatisfaction) or can be aggregated to e.g. the municipal level. Prior research has shown that individual level variables tend to be correlated not only with sector selection, but with outcomes, too, and are therefore not valid as instruments. We will show that the instrument we construct is both a strong predictor of private school attendance and is not correlated with educational outcome apart from the effect on school choice.

Our work extends the set of potential instruments suggested in the existing literatur by suggesting an instrument that exploits the exogenous variation of a students potential public school peers.

Results This is preliminary. We show that there is no private school effect for our

Danish data.

Overview The next section motivates the construction of the instrument variable. Sec-

tion 3 presents the empirical model and results. Concluding remarks follow.

2 Geographical variations in demand for private schooling

To correct for potential selection bias, we need to find a variable that is closely related to the probability of attending private school, but which does not directly affect educational attainment. The general problem is that (i) variables that are successful in predicting school choice (typically variables at the individual level), tend to be significant predictors of educational outcomes as well and (ii) variables at a more aggregated level are generally

less (or, at best, not) related to the educational outcome, but are on the other hand "not so well" predictors of school type.

In the literature, the most prominent instrument has long been religious affiliation, specifically Catholic religion[6]. However, this variable has been criticized for not being a valid instrument, as it apparently also enters the outcome equation significantly (Sander 1996, Figlio and Stone 1999) Therefore, later analyses have constructed various community-level variables (e.g. variation in the concentration of Catholics across communities) arguing that it should not affect your educational attainment whether your neighbor is Catholic or not (Sander and Krautmann 1995; Dee 1998; Neal 1997; Sander 1996, 2000; Evans and Schwab 1996). Another strand of literature uses ideas from the theory of local public goods when modelling school choice, the idea being that the quality of a locally provided school critically depends on the local population composition; primarily because it affects the composition of the student body (Lankford, Lee and Wyckoff 1995; Lankford and Wyckoff 1992, 2000)[7].

With respect to the research strategy of this paper we have two remarks - (i) we cannot use religious affiliation variables in the Danish context as religious schools play almost no role in the Danish private school sector and (ii) as the existing international research has shown, individual level variables are of little use as identifying variables. The research strategy becomes to look for instruments at the community level, and to borrow ideas from work on the demand of local public goods.

The main hypothesis motivating the instrument used in this paper is that demand for private schooling is a function of how well public school quality fits the needs of the local service population, because private schools offer additional options for those unhappy with the public school system. In a residence based public school system, the location of a family's residence directly determines which public school the family's children are eligible to attend. When dissatisfied with the public offer, some parents are taking advantage of the one minimum degree of freedom that they possess: the freedom to move from the public to the private sector[8]. Thus, low public school quality raises demand for private schools as substitutes for public schools (see Dee 1998). But what makes parents unhappy about the local public school? The school environment is composed of many factors: teachers, peers, financial resources. While the empirical evidence of the effects of educational inputs such

[6]Catholic affiliation has been useful, because a major part of the US private school sector, constists of Catholic schools. Therefore, most US studies are exclusively on Catholic private schools.

[7]Generally, the argument is that variation in the quality of local public goods is due to differences to tax-payments from the residents. But in the case of schooling, school resources do not seem to matter near as much as the socio-economic composition of the student body.

[8]According to the law, there is free school choice within municipalities in Denmark, but in practice this choice is quite restricted.

as per pupil spending, teacher experience, and teacher degree level have been shown to be relatively unimportant predictors of outcomes, and the impact of any particular input to be inconsistent across studies (see Hanushek 1986, Betts and Morell 1999, Goldhaber and Brewer 1997; Graversen and Heinesen 2000 for Danish results), other analyses provide evidence of the importance of peers in family's choice of schooling sector (Lankford and Wyckoff 1992; Lankford, Lee and Wyckoff 1995; Downes and Greenstein 1996; Goldhaber 1996; West and Palsson 1988; Figlio and Stone 1999). Peers are important since they, together with a student's ability, motivation and family inputs, affect his academic performance (Hoxby 2000b; Betts and Morell 1999; Hanushek et al. 2001; Goldhaber and Brewer 1997). Peers may affect the classroom process - aiding learning through questions and answers, contributing to the pace of instruction, or hindering learning through disruptive behavior. Peer group differences are clear: on average, private schools have students with significantly higher than average parental socioeconomic status levels, which is an important predictor for educational performance.

According to these arguments, we would expect each student's probability of choosing private school to be a function of the (average) peer group in the local public school and the student's own educational ambitions. We expect that the probability of attending private school increases both as parental education increases and public school peer quality falls. But we expect high educated families to be more sensitive to low peer quality in public schools than low educated homes. To see if this pattern holds in the data, we model the probability of attending private school as a function of the (average) peer group in the local public school, the student's own educational ambitions and the interaction of both, and the usual set of controls as described in section 3 (probabilities are calculated at the sample mean of the control variables). We proxy the quality of the public school peer group for each municipality by average parental education of public school students and we proxy the student's educational ambitions by the educational level of his own parents.

As can be seen from figure 1, a peer group dominated by students from low educated homes in the local public school is an important factor driving private school attendance for students from high educated parents. When public peer group quality rises, private school attendance of students from high educated families falls. The figure also shows that the higher the level of parental education, the more sensitive these families are with respect to public peer group quality. The gap between high and low educated family's choice of private school closes as public school peer quality peaks at an average of 13 years of schooling. When public peer quality is at its maximum, high educated families are no more likely to choose private school than are low educated homes. This result is useful, because while there may be many different reasons for choosing private school (individual

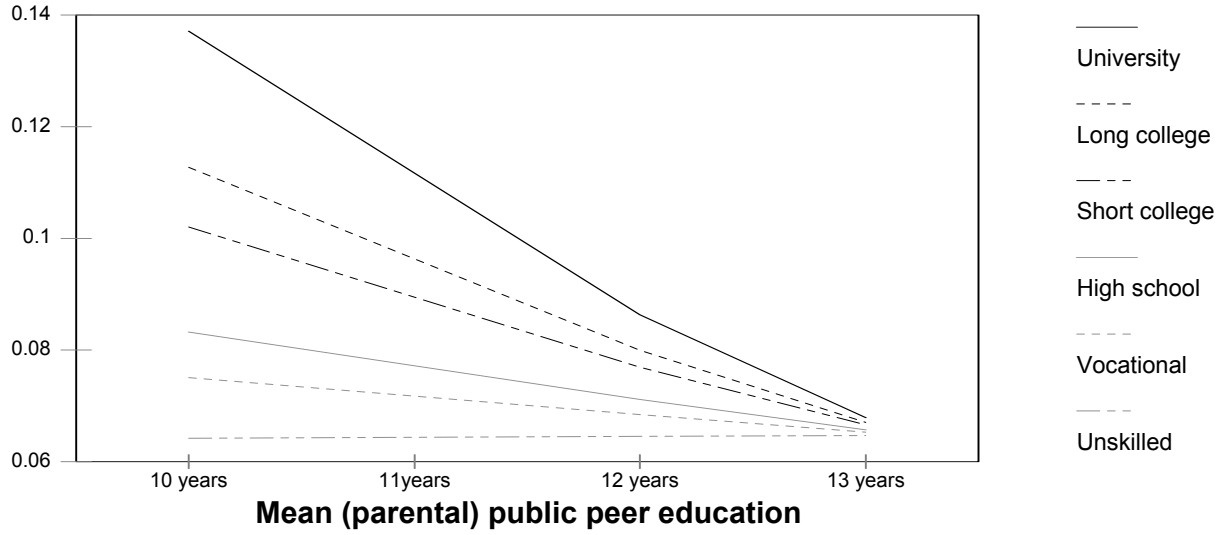

Figure 1: Estimated probability of private school attendance

taste for special curricula, pedagocial methods etc.), the quality of public school peers can be treated as exogenous variation in the public school offer.

Construction of the instrument We construct a measure of potential dissatisfaction with public school as a determinant of private school attendance at the municipal level based on the following idea: if all residents are satisfied with the local public school offer, there is no need for private schools, because the more homogeneous the local population is with respect to for example education, income and race, the more homogeneous their educational needs and preferences tend to be and the easier it is to meet the demand of all residents in the public school. In contrast, suppose that a municipality has a very heterogeneous population. In that case, demand for private schools is high: students who would otherwise attend public school might attend private school instead because their parents prefer them to have somewhat homogeneous peers.

We control for the general differences in private school preference by including a wealth of family background variables as regressors (they are included in the outcome equation as well). However, we need to model an expression for the parents' response to public school quality as well. As we favor a municipal-level variable as potential instrument, we need to construct a measure of (i) how many students' families are potentially unsatisfied with the public school offer as a share of all students and (ii) to what degree they are dissatisfied, because the less happy they are with the public school, the more attractive the private alternative will be to them. The construction will be in three steps.

First, we need a measure of public school quality. We proxy public school quality by the average parental education of public students. Unfortunately, we have no data on students' intellectual ability and must therefore use the best predictor which is their parents' education. It seems undisputable that performance in school is highly related to parents' educational status, see Graversen, Heinesen and Madsen (1999) for Danish results. Thus, for estimation purposes, the average parental education of public school students at the municipal level is an indicator for the peer group quality at public schools. One might argue that the average peer group in public schools is endogenous to the individuals' decision whether to choose public or private school. But it is reasonable to assume that each individual views school characteristics as being unaffected by his own actions.

Second, as we do believe that the greater the family's educational distance to public school quality, the more likely they are to opt out of public school, we need to model the (positive) "distance from mean". We do so by summing individual years of (positive) deviation from public school mean, which then is squared and weighted by the total number of students in the municipality. By squaring the deviation, we assume that the dissatisfaction with the public school offer increases non-linearly in excess years of schooling[9].

Third, as we need to determine the share of families potentially unsatisfied with the public school quality, we weight the sums of squared deviations by the total number of students[10]. The variable now expresses both the degree of variance *above* the mean, and the share of potentially dissatisfied students. Formally,

$$DPPR_j = \frac{1}{n_j}\sum_{i=1}^{n^*} D_{ij}^{*}\left[educpar_{ij} - \frac{1}{n_{pub_j}}\left(\sum_{i=1}^{n^*} D_{ij}^{pub} educpar_{ij}\right)\right]^2,$$

where $D_{ij}^{*} = 1$ if $educpar_{ij} > \frac{1}{n_{pub_j}}\left(\sum_1^{n_{pub_j}} educpar_{ij}\right)$, 0 otherwise and $D_{ij}^{pub} = 1$ if i attends public school, 0 otherwise. $DPPR_j$ is the **D**emand-**P**otential for **PR**ivate schooling in municipality j; n_j is the total number of students in the municipality; $educpar_{ij}$ is parental education for student i. $\left[\frac{1}{n_{pub_j}}\left(\sum_1^{n_{pub_j}} educpar_{ij}\right)\right]$ is the average parental education of public school students, which we use as a proxy for public school quality. As the equation shows, the sum of squared deviations is weighted by the total number of students in the municipality, which gives us a measure of the potential of dissatisfied

[9] When we estimate the model using linear rather that squared deviation from mean, the results are very similar.

[10] Note that we assume that less able children do not mind attending a school with on average superior peers.

families. One of the shortcomings of the identification strategy is that it is not possible to test whether $DPPR$ is a valid instrument in this just-identified model. However, we will provide an informal test of the validity of the instrument in the next section.

Actually, already at this stage we can verify one of the propositions with our data. We suggested that in both homogeneous and heterogeneous communities, parents select private schools according to their general preference for private schooling. In addition to that, some parents in heterogeneous municipalities respond to the deviation of public school quality from their own needs, such that parents to a greater extent choose the private alternative. The different choice mechanisms in heterogeneous and homogeneous municipalities should result in stronger selection by ability in heterogeneous areas. Our data provide some evidence for this as the difference in the average parental education of public and private students is more pronounced in heterogeneous municipalities: at an average, in the 20 most homogeneous municipalities, private school students have parents with 6% more (years of) education than public school students (11.5 years as opposed to 10.9 years of education); for the 20 most heterogeneous municipalities, the difference is 12% (12.8 and 11.4 years, respectively). This is evidence of a greater degree of sorting by ability between public and private schools in heterogeneous municipalities, presumably due to the inadequateness of the public school offer for well-educated families in the municipality.

3 Empirical estimation

Data The empirical strategy outlined requires two sets of exogenous variables: first, variables that appear in both the second-stage educational attainment estimation and the first-stage estimation of sector choice; and second, the instrument, $DPPR$, that is excluded from the second-stage regression but used to explain sector choice in the first stage. The set of variables that appears in both the first and second stages include data on the student's personal and family background, and neighborhood characteristics on the municipality level. The present paper uses a rich data set combining individual level data (a panel 10% random sample) from administrative registers for the young people themselves and their families with data for socioeconomic conditions at the municipal level[11]. We adopt a set of variables for inclusion as controls in both the first and second-stage regressions that is roughly comparable to the sets used in other recent studies,

[11]The data employed in this study are drawn from a data set originally collected for the analysis of Graversen and Heinesen (2000) on the effect of school resources on educational attainment in Danish (public) schools. In addition to the data set used in Graversen and Heinesen, we include data on private school students as well.

although perhaps more comprehensive than the typical set of controls. Three sets of explanatory variables are considered[12]:

1. Personal characteristics (age, gender, ethnicity)
2. Family background (parents' education, income, wealth, labor market status and unemployment, number of (younger) siblings, family structure and housing conditions (all measures taken at the student's age 15), teenage parents)
3. Socio-economic conditions in the municipality (general level of education and unemployment and the share of bilingual students in the municipality at the student's age 15)

We classify students as public or private school students based on the school they attended in 8th grade. This study focuses on two measures of educational attainment measured at age 25: (i) the probability of having passed upper secondary school or a vocational education and (ii) the probability of being enrolled in or having finished college. The data set used in the analyses contains data on more than 23.000 young people born between 1965 and 1970, i.e. they have started school between 1972-1977 and have finished comprehensive school between 1981-1987 (attendance of the 10th grade is voluntary). The share of private school students in the sample is 9.5 percent. We restrict the analysis to the municipalities belonging to the metropolitan areas of the three biggest cities (Copenhagen, Århus and Odense) and to other bigger cities. We do so, because preliminary work has shown the instrument not to be valid in the "whole country" sample. This is probably due to the low variation of the instrument variable in rural areas. Municipalities without observation for private school students are excluded from the analysis because it is reckoned that private school attendance might not be feasible in these municipalities. Missing data on parental education has been coded to 7 years of schooling, the minimum years of schooling for these generations.

We would note that we cannot analyze whether a positive private school effect on academic attainment can be attributed to difference in expenditure. Even though per capita expenditure in public schools is about 6% higher than in private schools, there are fewer students with special needs in private schools, too. As we do not have resource data on the individual private schools, we cannot pursue this issue any further.

The model The key empirical problem in implementing a two-stage model is in distinguishing the private school effect from unobservable factors that are correlated with

[12]The tabel of descriptive statistics are in the appendix.

educational attainment. We address this endogeneity problem by using an instrumental-variables strategy that uses variation in potential demand across municipalities due to heterogeneity in the educational needs of the service population of public schools.

As suggested above, ordinary least squares estimates of private school effects are biased if there is selection on unobservables into the private school sector. We estimate the instrument variable model as a linear probability model. As Angrist (1991) has shown, linear instrumental variables estimators perform nearly as well as the correctly specified maximum likelihood estimator, especially in large samples. Also in a recent study, Angrist (2001) argues that standard linear model techniques are generally applicable (also in the case of limited-dependent outcome variables), if the focus of the study are treatment effects rather than structural parameters.

Formally, we need to obtain a consistent estimate for the parameter β_3, the effect of private schooling ($P_{ij} = 1$) on post-secondary educational attainment, Y_{ij} (for student i in municipality j) in equation (1). The model that is estimated is given by

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 D_j + \beta_3 P_{ij} + \varepsilon_{ij} \tag{1}$$

where X_{ij} is a vector of individual and family specific controls, D_j are municipality characteristics and ε_{ij} is the residual.

To control for the potential endogeneity between P_{ij} and ε_{ij}, Y_{ij} is estimated using two-stage least squares where P_{ij} is treated as endogenous and an additional exogenous variable, namely $DPPR$, is used for identification. The educational attainment model thus becomes:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 D_j + \beta_3 \widehat{P}_{ij} + \varepsilon_{ij} \tag{2}$$

where $\widehat{P}_{ij}$ is the predicted value of a regression of P_{ij} on X_{ij}, D_j and $DPPR_j$.

One complication comes from the fact that our instruments vary at the level of the municipality rather than the individual level, which leads to downward biased standard errors for the second-stage point estimates. We address this problem by estimating Huber-White standard errors. [Not done, yet.] Because my instrument relies on the municipality as the geographical unit of variation, the results might be sensitive to the potential for endogenous household location choice within a multicounty metropolitan area. Figlio and Stone 1999 try to aggregate over the problem. Their results are largely unchanged.

Results Before discussing results, we will provide evidence of the validity of the instrument: (i) significant incremental power in explaining sector selection in the sector choice equation, but (ii) no significant independent explanatory power in the educational attainment function that controls for private school selection.

(i) Predictive power for sector selection

Results from the first stage regression are shown in Table 2. As expected, the potential of parents dissatisfied with the local public school ($DPPR$) increases the probability of attending private school. The instrument has strong explanatory power in the first-stage regression for the decision to attend private rather than public school. The t-statistic of the instrument is equal to 4.89 ($p<0.0001$), corresponding to a F-statistic of 23.9 ($F[1, n-K] = t^2\,[n-K]$), which exceeds the cutoff value of 10 suggested by Staiger and Stock (1997) by far. This means that the probability of parents choosing private schools increases, the more heterogeneous a municipality is with respect to education.

(ii) Validity of the instrument

We saw that the instrument satisfies the requirements to predictive power in the selection equation. Thus, the credibility of the instrument variable results turns on the assumption that students from heterogeneous municipalities are not more or less likely to graduate/enroll in college once the school choice decision is taken into account than otherwise identical students who live in homogeneous municipalities. Simple regressions of the educational outcome measures on the instrument suggest that, while the probability for finishing upper secondary school or getting an educational degree on average is equal in heterogeneous and homogeneous municipalities, the probability to go to college is significantly higher in heterogeneous municipalities. This will pose a problem for the estimation if, after controlling for all observable characteristics, the instrument is correlated with a student's unobserved propensity to enter college. Although we cannot formally prove that this is not the case in my just-identified model, we would note that when adjustments are made for background variables and type of school attended, the instrument, $DPPR$, is not significantly related to educational attainment. The estimated coefficients are positive, but insignificant with p-values as high as 40% and 56%, respectively). Although this is not a direct test of whether our instrument is valid, it does indicate that, as a group, families who choose to live in homogeneous municipalities are not different (with respect to unobservables that might influence outcomes) from those living in heterogeneous municipalities. More on instrument validity later in this paragraph.

Estimation results

[Preliminary.] As can be seen from Table 2, the instrument variable estimation does

not provide evidence that private schools on average do any better or worse than public schools. By comparison, in an ordinary least squares model, the private school effect is insignificant, too [results not reported]. Moreover, the fact that controlling for unobserved characteristics in the instrument variable estimation does not alter the results compared with OLS is an indicator of the absence of bias due to unobservables.

Effects of background controls

[To be written.]

More on instrument validity

Intuitively, what can explain the validity of the instrument variable? Basically, our hypothesis is that, no matter what, students are always sorted into similar peer groups. Either, they self-sort by residential location, thereby forming groups that are relatively homogeneous in terms of their preferences with regard to schooling. In these municipalities, public schools can be targeted to this group. Or, in heterogeneous municipalities, students self-sort across public and private schools.

In this paper, as in most of the existing empirical literature, we assume that we can treat the location decision of the household as exogenous, i.e. parents choosing to live in a homogeneous municipality are not unobservationally different (with respect to characteristics affecting educational outcome) from parents choosing to live in a heterogeneous municipality.[13] We can argue that this is so because parents, no matter where they live, seek a good peer group fit for their child. Whether they can find the adequate peer group within the private or public school sector depends on the type of municipality they live in. We argue that once students have been selected into their individually adequate type of school (private or public), i.e. the type that matches their peer group, there is no effect of the residential composition on educational attainment. This result is most probably the result of sorting by peers. Thus, the variation we use to identify the model is the exogenous variation in the population composition of communities, which is exogenous to parents, when we treat the location decision of the household as fixed. For this factor of determination of private school enrollment we use, private school choice is merely a sorting device. The variation in the population composition of communities that identifies my model is mere peer sorting: in those communities, where "high-ability" peers cannot meet in public school because of low academic standards, they select into private schools.

The hypothesis is confirmed by the finding of Lankford and Wyckoff (2000), who find that students are equally segregated by school in metropolitan areas with greater and lesser

[13]There is suggestive evidence on this; e.g. in inner-city Copenhagen (heterogenous), one out of two students from high-educated families is attending private school, while this, for example, is only the case for one out of five students in the well-off suburb of Gentofte (homogenous).

degrees of school choice among districts. They argue that the degree of inter-district school choice does affect the way households are sorted across districts, but that the resulting sorting appears to have little effect on the average level of achievement. This may be so, because students may already be self-sorted into peer groups across schools (within the district). Therefore, greater inter-district choice will not alter the peer pattern and thus achievement. Also, Lankford and Wyckoff show that when inner-city schools in the US were desegregated (for example by busing), whites began to move from the urban to the white suburban areas. This underlines the strength of peer sorting: when peer matching across schools within the district is disrupted by desegregation (or other) policies, people start to sort to a greater extent by residential location and they thereby re-establish the former peer groups in new locations.

The intuitive explanation for the validity of the instrument, confirmed by the results from (informal) testing, is that because students will always be sorted by peer groups (either by location or by type of school), the degree of heterogeneity in the municipality does not influence student performance, except for a potential private school effect.

4 Concluding remarks

In this paper, we estimate the effect of attending private rather than public school on educational attainment measures. We suggest a new, more targeted instrument variable, especially designed to capture a part of the variation in demand for private schooling that is not related to educational outcomes. We find the instrument to be both particularly strong and (informal) tests show that it can be validly excluded from the outcome equation. Intuitively, this is probably due to peer sorting across either school types or residential location. We estimate a two-stage least squares model (as linear probability model) and find positive, but insignificant private school effects for both outcome measures.

References

[Angrist, Joshua (1991)] Instrumental Variables Estimation of Average Treatment Effects in Econometrics and Epidemiology; NBER, Technical Working Paper no. 115, 1991

[Angrist, Joshua (2001)] Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors: Simple Strategies for Empirical Practice; Journal of Business&Economic Statistics, 19, 1, pp. 2-16

[Benabou, Roland (1993)] Working of a City: Location, Education and Production; The Quarterly Journal of Economics, August 1993; pp.619-652

[Betts, Julian R. and D. Morell (1999)] The Determinants of Undergraduate Grade Point Average - The Relative Importance of Family Background, High School Resources, and Peer Group Effects; The Journal of Human Resources, vol. 34, pp.268-293

[Dee, Thomas S. (1998)] Competition and the Quality of Public Schools; Economics of Education Review; 17, 4, pp. 419-427

[Downes, Thomas A. and Shane M. Greenstein (1996)] Understanding the Supply Decisions of Nonprofits: Modelling the Location of Private Schools; RAND Journal of Economics; 27, 2, pp.365-390

[Epple, Dennis and R. Romano (1998)] Competition Between Private and Public Schools, Vouchers and Peer Group Effects;American Economic Review , March 1998

[Epple, Dennis and R. Romano (2001)] Neighborhood Schools, Choice, and the Distribution of Educational Benefits; forthcoming in The Economics of School Choice (C. Hoxby, ed), National Bureau of Economic Research.

[Evans, W. and Schwab, R. (1995)] Finish High School and Starting College: Do Catholic Schools make a Difference? Quarterly Journal of Economics, 110, 947-74

[Figlio, David N. and Joe A. Stone (1999)] School Choice and Student Performance: Are Private Schools Really Better?; Solomon Polachek, ed., Research in Labor Economics, 1999, pp.115-140

[Goldhaber, Dan (1996)] Public and Private High Schools: Is School Choice an Answer to the Productivity Problem? Economics of Education Review, 15, 2, pp.93-109

[Goldhaber, Dan and Dominic Brewer (1997)] Why Don-'t Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity; The Journal of Human Resources, vol. 23, 3, pp. 503-523

[Graversen, Brian; Eskil Heinesen (2000)] Effects of School Resources on Students' Educational Attainment; mimeo; 2000

[Hanushek, E. A. (1986)] The Economics of Schooling: Production and Efficiency in the Public Schools; Journal of Economic Literature. XXIV(3): 1141-1178.

[Hanushek, E.A.; Kain, John F.; Markman, Jacob M. and Steven G. Rivkin (2001)] Does Peer Ability Affect Student Achievement? NBER Working Paper 8502

[Hoxby, Caroline Minter (1994)] Do Private Schools Provide Competition for Public Schools?; NBER Working Paper no. 4978

[Hoxby, Caroline Minter (2000)] Peer Effects in the Classroom: Learning from Gender and Race Variation; NBER Working Paper no. 7867

[Lankford, Hamilton and James Wyckoff (1992)] Primary and Secondary School Choice Among Public and Religious Alternatives; Economics of Education Review; vol.11, no. 4, pp. 317-337

[Lankford, Hamilton and James Wyckoff (2000)] The Effect of School Choice and Residential Location on the Racial Segregation of Students; mimeo, State University of New York-Albany

[Lankford, Hamilton; Lee, E.S. and James Wyckoff (1995)] An Analysis of Elementary and Secondary School Choice; Journal of Urban Economics, 38, pp. 236-251

[Neal, Derek (1997)] The Effects of Catholic Secondary Schooling on Educational Achievement; Journal of Labour Economics, 15, 1, pp.98-123

[Rouse, Ceclia Elena (1997)] Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program; NBER Working Paper No. 5964

[Sander, William (1996)] Catholic Grade Schools and Academic Achievement; The Journal of Human Resources, 31,3, pp. 540-548

[Sander, William (1999)] Private Schools and Public School Achievement; The Journal of Human Resources, 34, 4, 697-709

[Sander, William (2000)] Catholic High School and Homework; Educational Evaluation and Policy Analysis; 22, 3, pp.299-311

[Sander, William and Anthony Krautmann (1995)] Catholic Schools, Dropout Rates and Educational Attainment. Economic Inquiry, 33, 2, 217-233

[Staiger, Douglas and James H. Stock (1997)] Instrumental Variables Regression With Weak Instruments; Econometrica; vol. 65 (3), pp. 557-586

[West, Edwin G. and H. Palsson (1988)] Parental Choice of School Characteristics: Estimation Using State-Wide Data; Economic Inquiry; vol. 26, pp.725-740

[Wright, Robert (1999)] The Rate of Return to Private Schooling; IZA Discussion Paper no. 92

Table 1: Descriptive Statistics

Variable	Mean		
	Public	Private	
Number of observations	21582	2050	
Personal characteristics			
Born in 65	0,12	0,09	***
Born in 66	0,13	0,09	***
Born in 67	0,12	0,10	***
Born in 68	0,11	0,10	
Born in 69	0,10	0,11	***
Born in 70	0,10	0,12	***
Female	0,49	0,49	
Of Danish/Western origin	0,994	0,993	
Immigrant	0,004	0,005	
Second generation immigrant	0,002	0,002	
Family background			
No. of siblings aged 0-17	0,76	0,71	***
Has younger siblings	0,49	0,45	***
Lives with both parents	0,77	0,72	***
Lives with single mother	0,12	0,16	***
Lives with mother and stepfather	0,07	0,07	
Lives with single father	0,018	0,022	*
Lives with father and stepmother	0,01	0,01	
Does not live with father or mother	0,02	0,02	
Father and mother live together, but child does not live with parents	0,004	0,004	
Mother not in the register	0,01	0,01	
Father not in the register	0,07	0,09	***
Teenage mother	0,12	0,08	***
Teenage father	0,03	0,02	***
Parental education	11,2	12,3	***

*(**,***) indicates significance above the 0.01 (0.001, 0.0001) level

(continued)

Table 1, continued

Variable	Mean		
	Public	Private	
Family background, continued			
Mother wage earner	0,74	0,73	
Mother self-employed	0,11	0,14	***
Mother student	0,004	0,11	***
Mother receives social assistance	0,01	0,01	
Mother not active in labor market	0,13	0,11	**
Father wage earner	0,69	0,64	***
Father self-employed	0,19	0,22	***
Father student	0,001	0,002	
Father receives social assistance	0,005	0,005	
Father not active in labor market	0,04	0,04	
Mother's labor income (in 100,000 DKK)	1,23	1,41	***
Father's labor income (in 100,000 DKK)	2,05	2,30	***
(Mother's income)*(child does not live with mother)	0,49	0,66	***
(Father's income)*(child does not live with father)	2,66	3,70	***
Mother's degree of unemployment	6,08	5,23	*
Father's degree of unemployment	4,37	3,53	**
(Mother's degree of unempl.)*(not with mother)	0,49	0,38	
(Father's degree of unempl.)*(not with father)	1,33	1,34	
Mother's financial wealth	0,42	0,88	***
Father's financial wealth	2,85	3,53	**
(Mother's financial wealth)*(not with mother)	0,01	0,02	
(Father's financial wealth)*(not with father)	0,10	0,10	

*(**,***) indicates significance above the 0.01 (0.001, 0.0001) level

(continued)

Table 1, continued

Variable	Mean		
	Public	Private	
Family background, continued			
Lives in owner-occupied dwelling	0,73	0,74	
Lives in rented dwelling	0,23	0,21	*
Lives in not-categorized dwelling	0,02	0,03	***
Type of dwelling unknown	0,01	0,01	
Number of rooms per person	1,21	1,34	***
Number of rooms per person unknown	0,006	0,006	
Lives in socially deprived area	0,037	0,026	***
Socioeconomic conditions in the municipality			
Proportion of bilingual children in the school municipality (%)	1,34	1,99	***
Rate of unemployment in the school municipality (%)	8,68	8,81	**
Proportion with only compulsory school education in municipality (%)	56,28	54,83	***
Proportion with a vocational education in the municipality (%)	30,75	30,58	
Proportion with a further or higher education in the municipality (%)	12,97	14,59	***
Outcomes			
Passing upper secondary school or getting a vocational degree	0,76	0,81	***
Being enrolled in/having finished college	0,29	0,38	***

*(**,***) indicates significance above the 0.01 (0.001, 0.0001) level

Table 2: Instrument variable estimates - first and second stage regressions

Variable	School choice		Educational attainment			
			Upper secondary/ vocational educ.		Enrolled in/ finished college	
DPPR	0,016 **	0,004	-		-	
Private school	-		0,031	0,358	-0,038	0,367
Personal characteristics						
Born in 65	-0,040 **	0,007	-0,002	0,019	-0,032	0,019
Born in 66	-0,042 **	0,007	0,018	0,016	-0,031	0,019
Born in 67	-0,020 *	0,007	0,026	0,013	-0,005	0,012
Born in 68	-0,012	0,007	0,003	0,010	-0,007	0,010
Born in 69	-0,016	0,007	0,003	0,011	0,005	0,011
Born in 70			(reference)			
Male			(reference)			
Female	-0,001	0,004	-0,004	0,005	0,030 **	0,007
Of Danish/Western origin			(reference)			
Immigrant	0,027	0,023	-0,086	0,034	0,147	0,029
Second generation immigrant	0,010	0,071	0,076	0,074	0,186	0,077
Family background						
No. of siblings aged 0-17	0,015 **	0,003	-0,026 *	0,008	0,012	0,008
Has younger siblings	-0,019 *	0,006	0,050 **	0,011	0,027	0,011
Lives with both parents			(reference)			
Lives with single mother	-0,008	0,010	-0,115 **	0,015	-0,024	0,016
Lives with mother and stepfather	-0,002	0,011	-0,100 **	0,018	-0,052 *	0,017
Lives with single father	-0,022	0,019	-0,089 *	0,031	-0,059	0,028
Lives with father and stepmother	-0,043	0,023	-0,106 *	0,037	-0,089 *	0,033
Does not live with father or mother	0,064 *	0,021	-0,301 **	0,039	-0,046	0,044
Father and mother live together, but child does not live with parents	-0,013	0,030	0,027	0,061	0,015	0,033
Mother not in register	0,039	0,027	0,0003	0,041	0,084	0,044
Father not in register	0,071 **	0,011	0,079	0,034	0,128 **	0,031
Teenage mother	-0,018 *	0,006	-0,054 **	0,010	-0,066 **	0,010
Teenage father	-0,004	0,010	-0,020	0,013	-0,033	0,014

For each estimation, the figure on the left is the regression coefficient, the figure to the right is the standard error (robust and corrected for clustering at the municipality level) of the estimate. *(**) indicates significance above the 0.01 (0.001) level..

(continued)

Table 2, continued

Variable	School choice			Educational attainment					
				Upper sec./ vocational educ.			Enrolled in/ finished college		
Family background, continued									
Parental education	0,011	**	0,001	0,022	**	0,005	0,046	**	0,004
Mother wage earner				(reference)					
Mother self-employed	0.062	**	0,008	-0,018		0,021	0,009		0,022
Mother student	0.109	**	0,023	-0,050		0,047	0,094		0,055
Mother receives social assistance	0.035		0,016	-0,157	**	0,027	-0,019		0,023
Mother not active in labor market	0.020	*	0,007	-0,056	**	0,011	-0,005		0,012
Father wage earner				(reference)					
Father self-employed	0.062	**	0,008	0,020		0,025	0,054		0,025
Father student	0.097		0,044	0,062		0,076	0,105		0,074
Father receives social assistance	0.047		0,025	-0,113		0,045	-0,009		0,032
Father not active in labor market	0.038	**	0,010	-0,033		0,019	0,037		0,020
Mother's labor income (100,000 DKK)	0.001	**	0,0003	-0,0004		0,001	0,002		0,001
Father's labor income (100,000 DKK)	0.001	**	0,0002	0,001		0,0004	0,002	**	0,0004
Mother's income*not live with mother	0.002		0,001	-0,00002		0,001	-0,0005		0,001
Father's income*not live with father	0.0003		0,0003	0,001	*	0,0004	-0,0003		0,0005
Mother's degree of unemployment	0.0004	*	0,0001	-0,001	**	0,0003	-0,0005		0,0002
Father's degree of unemployment	0.0001		0,0002	-0,002	**	0,0003	-0,0003		0,0002
(Mother's degree unempl.)									
*(child does not live with mother)	-0.0002		0,0004	0,0008		0,0007	0,0004		0,0005
(Father's degree unempl.)									
*(child does not live with father)	0.0003		0,0003	0,002	**	0,0005	0,0005		0,0004
Mother's financial wealth	0.0016		0,0007	0,003		0,001	0,003		0,002
Father's financial wealth	0.0004		0,0002	0,001		0,0003	0,01	*	0,0005
Mother's financ. wealth*not with mother	-0.0012		0,004	0,001		0,005	0,003		0,006
Father's financ. wealth*not with father	-0.0006		0,0003	-0,001		0,0003	-0,002	*	0,0006

For each estimation, the figure on the left is the regression coefficient, the figure to the right is the standard error of the estimate. Standard errors are robust and corrected for clustering at the municipality level. *(**) indicates significance above the 0.01 (0.001) level

(continued)

Table 2, continued

Variable	School choice		Educational attainment					
			Upper secondary/ vocational educ.			Enrolled in/ finished college		
Family background, continued								
Lives in owner-occupied dwelling	(reference category)							
Lives in rented dwelling	0.005	0,005	-0,083	**	0,009	-0,058	**	0,008
Lives in not-categorized dwelling	0.022	0,011	-0,022		0,014	-0,024		0,017
Type of dwelling unknown	0.008	0,025	-0,084		0,042	-0,102		0,039
Number of rooms per person	0.051	** 0,005	0,044		0,019	0,075	*	0,021
Number of rooms p.p. unknown	0.038	0,031	0,054		0,043	0,119		0,045
Lives in socially deprived area	-0.021	0,008	-0,048	*	0,018	-0,024		0,013
Socioecon. conditions in municipality								
Proportion of bilingual children in the school municipality (%)	0.006	** 0,001	-0,004		0,003	-0,004		0,003
Rate of unemployment in the school municipality (%)	0.004	** 0,001	-0,004		0,003	-0,001		0,003
Proportion with only compulsory school education in municipality (%)	(reference category)							
Proportion with a vocational education in the municipality (%)	0.0006	0,0006	-0,001		0,001	-0,003		0,001
Proportion with a further or higher education in the municipality (%)	-0.0002	0,001	0,0003		0,001	0,002		0,001
Constant	-0,265	** 0,038	-0,423	**	0,114	-1,319	**	0,101
Number of observations	23632		23632			23632		
R^2	0,045		0,119			0,147		

For each estimation, the figure on the left is the regression coefficient, the figure to the right is the standard error of the estimate. Standard errors are robust and corrected for clustering at the municipality level. $^{*}\left(^{**}\right)$ indicates significance above the 0.01 (0.001) level