School Choice and Post-Secondary Education: Evidence from Canada

Christopher W. Jones

August 2, 2010

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

August 2010

To my advisor, Professor Chris Ferrall, for helping me to cultivate a germ of an idea and grow it into something more. To Casey Warman, for keeping the data center open late. To Rachel Ferreira, for always answering emails. And of course to all the MAs at Queen's who made this year so unforgettable.

1 Introduction

When it comes to their children's education, Canadian parents have choices. In fact some 92 percent of Canadians have access to more than one publicly funded option for their children's primary or secondary education – more than most developed countries, and certainly more than their neighbors living south of the border (Hepburn, 2002).

Indeed, the variety is impressive. Each province operates their own English and French language school boards as mandated by the *Charter of Rights and Freedoms*. Ontario, Alberta, and Saskatchewan also operate Catholic school systems which are separate from the English and French public systems and have a clear religious affiliation[1] In these provinces there is, in essence, "one publicly funded system of education with two dimensions, the public schools and the separate schools" ("Public and separate schools," Alberta DOE). Like the public school systems, the separate school systems are comprised of both English and French school boards. In provinces where separate schools are not mandated by legislation, religious schools usually take the form of semi-public independent schools, or private schools, some of which receive a significant portion of their funding from the provincial government.[2]

Parents can enroll their children in either a public or separate school in provinces where separate schooling exists. They are also free to choose between English- and French-language education. The right to attend a French school is guaranteed by law for children of Francophone parents, or children of parents who themselves attended school in French. Anglophone parents may also enroll their child in French school but admission is not automatic. In order to be accepted, the child must provide the school with satisfactory evidence of their French language proficiency (Ontario MOE, 2009). Parents who prefer their child to remain in the English system, but also desire French language education, have the option of enrolling them child in a French immersion program. Administered within the English or Separate-English school systems, immersion students typically receive 50% or more of their core curriculum in French. Not all high schools offer these programs, however, and admission into one is typically competitive (Ontario MOE, 2009).

[1] The website of the Alberta Ministry of Education gives a good overview.

[2] Not to be confused by terminology, separate schools, while not part of the "public" system, are still fully publicly funded.

In sum, a Canadian family may have upwards of six publicly-funded options for their child's education – English, French, French immersion, separate English, separate French, separate immersion – with semi-publicly funded independent schools and private schools as two additional options. But with so many choices, Canadian parents must work hard to know which type of school will maximize their children's future outcomes. And it is not clear whether one school type is superior to the others in this regard.

It is well documented that some types of schools yield better outcomes than others. For example, numerous studies have uncovered evidence that Catholic schools in the United States tend to have higher post-secondary participation rates relative to public schools.[3] But these studies only look at two types of schools, where Catholic schooling exists as the lone alternative to the public school system. In Canada, school choice is more multidimensional than simply "free or fee." Here, English, French, and separate schools co-exist as public entities, all "free" options that are fully funded by taxpayer dollars.[4]

Given the uniqueness of the Canadian model, this paper asks the question: does school type really matter? More formally, does there exist a causal relationship between the type of secondary school that a student attends and that student's educational attainment?

The answer is not an easy one for it is complicated by the fact that parental choice can be explained by a variety of factors, many of which are not directly observable. Consider the case where more choice leads to increased social segregation, with the brightest, most advantaged students opting to attend one type of school and the lower, disadvantaged students opting for another. Consequently, a sizeable post-secondary achievement gap is observed between students in the two schools. Is this gap owing to disparities in instructional quality between the schools, or to differences in the ability of the students that populate them?

The example underscores the importance of controlling for the fact that school choice is not pre-determined. Suppose we hypothesize that enrollment in a French immersion program will have a positive impact on a person's educational attainment. We believe that immersion programs

[3]See Evans and Schwab (1995) and Neal (1997).

[4]And even some private schools receive funding from the government, either through direct per-student grants or through a tuition tax credit for parents who send their children to private schools. See Hepburn (2002).

are rigorous, and so do a better job than non-immersion programs at preparing students for the rigors of university coursework. In testing this hypothesis, we would would be remiss to ignore the possibility that immersion students tend to be more gifted anyway, and are thus more likely to have higher outcomes regardless of the language they learn in. With this possibility left unaccounted for, the naive researcher would attribute any observed attainment gap to differences in school type. But this would grossly overstate the causal effect of type on outcomes since the estimates would be picking up the fact that one school type – French immersion – is populated by high ability students.

An appropriate fix in these situations is to introduce an instrument that can explain school choice but not educational attainment directly. If high school type is still found to be a significant determinant of outcomes then we could infer that the source of the school's effectiveness lies not in the innate characteristics of the students that are enrolled there, but rather in some feature of the school itself. Di Pietro and Cutillo (2006) invoke similar reasoning in their study of Catholic secondary schools in Italy. If brighter students tend to enroll in Catholic schools, then Catholic school students may have better post-secondary outcomes, they surmise, because of peer effects and innate ability, and not because they attended a Catholic school. But even after controlling for peer effects and resource quality, they still find that Catholic secondary schools have higher university participation rates relative to public schools. They conclude that it is some feature of the Catholic school system – their rigorous adherence to discipline and the more conducive learning environment that accompanies it perhaps – that is driving their results (p. 222).

In the empirical work that follows, I account for the fact that school choice is endogenous by constructing instruments based on distance. Consider a parent with a choice of two schools, A and B. School A is within walking distance while school B is an hour's drive away. One might assume the parent will choose A over B because the transportation costs associated with B are prohibitively high. Even if B is the better school, the parent may very well be forced to settle for A. Consequently, distance helps explain why the child went to school A but conditional on that choice the location of school B has no direct impact on the child's educational attainment.

My study makes two very original contributions to the existing literature on school choice. First, while the successful use of geographical features as instruments is well-documented in the literature, I know of no other study that employs distance as an instrument for school type.[5] Second, my study makes use of a new Canadian data source that so far has been underutilized by the literature. The remainder of the paper is organized as follows. Section 2 develops a model of educational attainment, where attainment is measured in years of schooling and is dependent upon secondary school choice, which is endogenous. Section 3 contains a discussion of the estimates themselves. Finally, Section 4 concludes with suggestions for future research.

2 Data and Methodology

2.1 Theoretical Model

Often when dealing with educational outcomes it is appropriate to employ a discrete choice model along the lines of Evans and Schwab (1995) and Neal (1997). Here, the authors define outcomes as a dichotomous dependent variable that takes on a value of one if a student enrolls in university, and zero if not. They then employ a bivariate probit model to estimate the effect of Catholic schooling on the likelihood of enrolling in university.

By definition, a binary probit is appropriate for modeling outcomes when outcomes take on one of two unique values (e.g. "University" and "No University"). But what of discrete dependent variables that take on three or more values? A binary probit, for example, does not distinguish between students who dropped out of high school and high school graduates who did not attend university. Both are lumped into the "No University" category, even though the likelihood estimates for each category may be of interest to the researcher. When dealing with more than one ordered category, an ordered probit would be the appropriate choice of model.

However, the problem with the ordered probit model is that it is ill-equipped to handle instrumental variables, which makes its use in the present study hard to justify. The problem is

[5] Bedard (2001) finds that areas with a higher concentration of universities have higher post-secondary participation rates. That is, people who live closer to a university are more likely to pursue higher education relative to people who live farther away from one. Similarly, Neal (1997) uses the number of Catholic schools per square mile as an instrument for Catholic school enrollment and finds that the Catholic school students tend to have higher test scores. Hoxby (2000) finds that schools in cities with more streams also tend to have higher test scores. The idea is that competition among school districts promotes academic achievement, and that schools in cities with more districts tend to be more competitive. Hoxby corrects for endogeneity by using the number of streams as an instrument for number of school districts, since streams are natural boundaries around which school districts tend to form.

compounded by the fact that the model we seek to build will have more than one endogenous variable. In the studies cited above, the authors are only interested in the effect of one endogenous variable – Catholic school enrollment – on post-secondary outcomes. But this paper seeks to explain the effect of school type on outcomes when, as in Canada, parents have more than one type to choose from.

An alternative approach is to model a system of simultaneous equations and then estimate them jointly using three-stage least squares (3SLS). The benefit of 3SLS is that it allows us to obtain "efficient estimates of the full set of parameters that appear in all of the simultaneous equations" (Mackinnon 2004, p. 522). In the paragraphs that follow, I develop the theoretical underpinnings for a model of educational attainment and the 3SLS procedure that will be used to estimate it.

I begin by writing a system of equations in which some parameters are jointly determined.[6] Let $y_{i,n}$ define the i^{th} equation of a linear simultaneous system of the form

$$y_{i,n} = Z_{i,n}\beta_{1i} + Y_{-i,n}\beta_{2i} + u_{i,n} \quad , \qquad n = 1, \ldots, N \tag{1}$$

where $Z_{i,n}$ is an n x k_{1i} matrix of exogenous variables, $Y_{-i,n}$ is an n x k_{2i} matrix of endogenous variables, n is an index of observations, and $k_{1i} + k_{2i} = k_i$. The notation "$-i$" indicates that the k_{2i} columns of $Y_{-i,n}$ may contain any of the endogenous y's except for $y_{i,n}$ since it is already on the left-hand side of the equation. There are g endogenous variables in the system, y_1 through y_g, all of which are assumed to be jointly generated by g equations of the form (1). Let one of them, y_g, be the total years of schooling attained by person n, and let the rest y_{-g} be n's high school type. Define $Z_{i,n}$ as a vector of exogenous controls for person n's individual, family, and school characteristics. In this model, y_g ("years") and y_{-g} ("type") are jointly determined by $Z_{i,n}$. That is, both educational attainment and school type are dependent upon the set of exogenous characteristics contained within the matrix $Z_{i,n}$.

The assumption is important, for it implies that the error terms are correlated across the equations of the system. As a simple example, consider the fact that parental wealth is an exogenous

[6]The following is a summary presentation of the theory presented in *Econometric Theory and Methods* by Davidson and MacKinnon, 2004. I have retained the author's original notation wherever possible.

determinant of educational attainment *and* school choice. It is well known that children from wealthier families tend to have better opportunities than children from poorer ones. Consequently, wealthier children may have higher innate ability that allows them to go further in school. But high-income parents are also more likely to opt for expensive private education for their children since they can afford to make the tuition payments. This means that the error term in the equation for private school (one of the "type" equations) will be correlated with the error term in the "years" equation. In other words, the error terms are cross-correlated for a given n. More formally,

$$E(u_{i,n}u_{j,n}) = \sigma_{ij} \text{ for all } n, \quad E(u_{i,n}u_{j,m}) = 0 \text{ for all } n \neq m \tag{2}$$

where σ_{ij} is the ij^{th} element of the g x g positive definite matrix Σ. The second part of (2) says that although the $u_{i,n}$ are allowed to be correlated for a given n, they must still be homoskedastic and independent across n.

By stacking the y_1 through y_g equations we can express the full system as

$$y_\bullet = X_\bullet \beta_\bullet + u_\bullet, \qquad E(u_\bullet u_\bullet^T) = \Sigma \otimes I_n \tag{3}$$

where $X_\bullet$ is a gn x k diagonal block matrix with typical diagonal element X_i that can be partitioned as $X_i = \begin{bmatrix} Z_i & Y_i \end{bmatrix}$. Note that estimating (3) in its current form will generate biased and inconsistent estimates if we believe that school type is correlated with the error terms $u_{i,n}$. Indeed, it is likely that a parent's choice of school is related to some unobserved factor pertaining to their children's ability, in which case $E[u_i|X_i] \neq 0$. Take an example where a mother must decide on the best schooling option for each of her two children. One child excels in academics, while the other struggles a bit more. Given their different abilities, she decides to enroll the first child in a French immersion program but not the second. Even if she prefers bilingual education to all other options, she knows that her second child is less equipped to cope with learning in two languages. Ten years later the first child may be seen with higher educational attainment, which our model would attribute to differences in school types.[7]

[7] If across all observations high ability students tended to be placed in French immersion programs, then this would be an example of selection bias.

Instrumental variables can be used to obtain efficient estimates of the dependent variables when the independent variables have non-zero correlation with the error terms. A valid instrument would be one that has positive correlation with the endogenous independent variables but is uncorrelated with the error terms. In this case, we need an instrument that explains school choice but not ability directly.

As explained earlier, distance is an appropriate choice of instrument for school type. Most children have the option of riding a school bus when they attend school in their zoning district. But when a parent opts to send their children elsewhere, public transportation may no longer be guaranteed. Parents who work full time, for example, may have schedules that are not conducive to their children's early morning drop-offs and mid-afternoon pickups. These parents may send their children to the "neighborhood" school, even if they prefer a school that is farther away, if only because they need the school bus to get them there. Thus, distance can help explain a parent's choice of school for their child, but it does not explain the child's ability nor does it explain the child's future educational attainment. It should not be the case that higher ability children live closer to a given school type. Nor should the fact that a child lives near his school say anything about his post-secondary outcomes.

With potential instruments identified, I proceed by writing the estimating equation for the efficient GMM estimator for the system

$$X_\bullet^T(\Sigma^{-1} \otimes I_n)(y_\bullet - X_\bullet \beta_\bullet) = 0 \tag{4}$$

where Σ is the true contemporaneous covariance matrix. While it would be nice to obtain efficient estimates of (3) by solving equation (4) directly, it would require us to know the true value of Σ. If we assume the true Σ is diagonal, then (4) simplifies to

$$X_\bullet^T(I_g \otimes P_W)(y_\bullet - X_\bullet \beta_\bullet) = 0 \tag{5}$$

where P_W is the orthogonal projection matrix of W, an n x l matrix of instrumental variables whose columns also include the exogenous regressors contained in X_i. Although we can now estimate the system directly through (5), we are still forced to make some rather strong assumptions as to the true value of Σ, since (5) only produces an efficient estimator if we assume Σ is diagonal.

The nice thing about the 3SLS method is that it allows us to obtain efficient estimates for the system when the true Σ is unknown. The procedure is as follows. First, we calculate the 2SLS estimates for each equation in (3) by solving equation (5). Next, we use the 2SLS residuals to estimate the variance-covariance matrix of the structural errors, $\hat{\Sigma}_{2SLS}$, defined as

$$\hat{\Sigma}_{2SLS} \equiv \frac{1}{n}\hat{U}_{2SLS}^T\hat{U}_{2SLS} \tag{6}$$

where $\hat{U}_{2SLS}$ is a n x g matrix with i^{th} column $\hat{u}_i$. Finally, we substitute $\hat{\Sigma}_{2SLS}$ for Σ^{-1} in equation (4) and compute. This gives us the 3SLS estimator for our model

$$\hat{\beta}_{\bullet}^{3SLS} = (X_{\bullet}^T(\hat{\Sigma}_{2SLS}^{-1} \otimes P_W)X_{\bullet})^{-1}X_{\bullet}^T(\hat{\Sigma}_{2SLS}^{-1} \otimes P_W)y_{\bullet} \quad . \tag{7}$$

2.2 Data

2.2.1 Youth In Transition Survey

Data for the model's parameters, with the exception of distances, is taken from the *Youth in Transition Survey* (YITS), a longitudinal survey that began in 2000 to track the educational progress of young people in Canada. The YITS is unique in being the only Canadian survey to link information on respondents' high school and post-secondary experiences with a rich set of individual and family characteristics, making it well-suited for use in this study. And as an added appeal, the YITS has not seen much use in the literature, possibly because of access restrictions put in place by Statistics Canada.

The survey began with two target populations. The first (Cohort A) comprised a national sample of Canadian 15-year-olds who had also participated in the Programme for International Student Assessment (PISA) exam, an OECD-ccordinated literacy test. The second (Cohort B) comprised a slightly older group of young people, aged 18 through 20. Both groups have been followed up longitudinally every two years. The latest followup was carried out in 2008 when the respondents were aged 23 and 26 through 28, respectively. For Cohort A, the initial survey (cycle one) consisted of four components: the reading portion of the PISA exam, a student survey, a parent survey, and a school survey. Students were asked a variety of questions about their academic habits, extracurricular activities, and educational aspirations. Parents supplemented the student survey with information on family circumstances such as living arrangements, household income, and their own academic qualifications. Finally, secondary school principals provided information on their school's academic programs and resources. Parent and school questionnaires were not administered to Cohort B since most of the respondents had already finished high school by the time the first cycle was undertaken. For both cohorts, subsequent survey cycles consisted of only a student survey in which participants provided information on their post-secondary enrollment status, among other things.[8]

I begin by restricting my sample to include only those who had completed the initial survey in 2000 and who were successfully followed up with in 2008. I further restrict the sample to members of the younger Cohort A, since the data on school type is derived exclusively from the parent and school questionnaires which the older cohort did not fill out. For each observation, I merge the information on personal, family, and school characteristics from the cycle one survey with information on educational attainment from the latest followup survey. The personal characteristics to be included in the model are indicator variables for gender, citizenship, and PISA reading score, which serves as a proxy for academic ability. Family-specific variables are parental income, mother's level of education, and indicator variables for parents' native language (English or French), intact family, and whether the family lives in a rural area. Finally, student-teacher ratio serves as the lone school-specific control variable.

There are a total of seven endogenous indicator variables, one for each school type. In general, respondents are observed to be enrolled in one (and only one) school of the following types: English, French, French immersion, English separate, French separate, French immersion separate, and private. A student is said to be enrolled in an English school if English is the school's language of instruction and the parents have indicated that the student is not enrolled in an immersion

[8]See Motte, et al. (2008) and Finnie and Mueller (2008) for a more comprehensive discussion of the YITS.

program. Similarly, a student is considered enrolled in a French school if French is the language of instruction, and the student is not enrolled in an immersion program. An English school, French school, or French immersion program is considered part of a separate school system if, in the school survey, the student's principal has indicated that the school is religiously affiliated, is publicly funded, and is located in one of the three provinces that have a separate public system in place.[9] Any school that is not fully public is categorized as private, regardless of language instruction and religious affiliation.

Turning now to the data on educational attainment, I note that it is unusable in its current form. In the followup surveys, participants are asked to give the highest level of education attained to date, and the responses are coded as ordered levels of degree qualifications (i.e. diploma, bachelors, masters, etc). Since our model requires that the attainment variable be a continuous one, I am forced to transform the data. I allow each grade level up to grade twelve to count as one year of schooling. Respondents with an associates degree, bachelors degree, or masters degree are assigned 14, 16, and 17 years of schooling respectively. Respondents who indicated some post-secondary education below an associates degree get 13 years. Those who indicated some post-secondary education beyond an associates degree but below a bachelors degree get 14 years.[10] In some instances, respondents indicated that they were enrolled in a post-secondary program but gave no information on highest attainment to date. I give these types 13 years. Finally, for respondents who failed to complete high school but did not indicate at which grade they dropped out, I assume left school after grade ten and so give them ten years.[11]

The transformations may seem arbitrary – and to some extent they are. While a more accurate measure of attainment could be constructed by weeding through the information contained in all of the survey cycles and not just the first and last, time constraints rendered this infeasible for the present study. Nevertheless, my rough and ready approximation of attainment should still allow me to say something of interest about the relationship between school choice and educational outcomes.

[9] As noted before, they are Ontario, Alberta, and Saskatchewan

[10] A handful of respondents indicated that they had completed their PhDs and so received 21 years of education.

[11] In some instances, attainment data was not supplied. These observations were subsequently dropped from the sample.

2.2.2 School Directories

Data used to construct the instrumental variables comes from 2009-2010 school directories published by the provincial Ministries of Education. Each directory contains mailing addresses for all schools in a given province as well as information pertaining to school type. I began by merging information from these directories to create a master list of Canadian high schools grouped into seven types.[12] I then inputed the schools' postal codes into an online mapping tool to generate latitude and longitude coordinates for each through a process called "geocoding."[13] Next, I turned to pinpointing the location of the YITS respondents themselves. While the YITS does not list respondents' physical address, it does give their census enumeration area (EA). EA is one of the smallest census geographies, often encompassing an area no bigger than a few city blocks.[14] Consequently, they are a close approximation for location. Using the PCensus software, I generated coordinates for the EAs' geographical centers and merged them with the school coordinates in the master list.

With EA and school coordinates in hand, calculating distances is fairly straightforward. The distance in kilometers from the centroid of EA i to the n^{th} school of type g is given by

$$d_{g_n}^{\,i} = \sqrt{(x_{g_n} - x^i)^2 + (y_{g_n} - y^i)^2} \tag{8}$$

where x and y are latitudes and longitudes measured in kilometers, respectively.[15,16] Distance from i to the *nearest* school of type g is

$$D_g^i = min(d_{g_1}^{\,i}, d_{g_2}^{\,i}, \ldots, d_{g_n}^{\,i}, 100) \tag{9}$$

By imposing the constraint $D_g^i \leq 100$, I am assuming that no child will attend a school that is over

[12] Again these types are English, French, French immersion, separate English, separate French, separate immersion, and private.

[13] The website batchgeo.com runs off Google maps technology. Although I was hesitant to use it at first, I found the geocoding to be quite accurate, more so in fact then the geocoding available through PCensus.

[14] The EAs in YITS cycle one of are from 1996 Census geography. No effort was made to account for the fact that some of the newer schools in the 2009-2010 directories were not around in 1996.

[15] This is simply the distance formula for Cartesian distances. The formula does not account for the curvature of the Earth, which is often important when dealing with geographical distances. However, since the distances being calculated are relatively small – less than 100 kilometers – this formula is accurate.

[16] For each EA, I converted the coordinates' unit of measure from decimal degrees to kilometers. The conversion is $111.\overline{11}$ kilometer per degree.

100 kilometers away. This is mainly to address instances where a school type is unavailable in a province. For example, for a person in British Columbia, the nearest Catholic school would be in Alberta, since B.C. does not have a separate school system.

As a final step, I apply a logarithmic transformation to the distances in (9) and redefine them as *relative* distances, using English public schools as the reference category. The final set of instruments looks like

$$r_g^i = ln(D_g^i) - ln(D_{English}^i) \qquad \text{where } g \neq English \tag{10}$$

and r is the shortest relative distance. Indeed, relative distances are really all we care about. A rural family may have the choice between two schools that are both twenty kilometers away, while an urban family faces the same choice, but each school is only one kilometer away. In relative terms the urban and rural families face identical choice sets, and we want our model to reflect that.

2.3 The Case of Quebec

Few studies on a Canadian topic could be considered complete without the author addressing, at some point, the uniqueness of French-speaking Quebec. Until now, I have glossed over the fact that Quebec's educational institutions, like its language, are quite distinct from the rest of Canada. For one, French education is the default. Youth in Quebec are required to attend school in French unless they have a parent who had attended school in English. But perhaps the most striking difference, and the one most relevant for our purposes, is the fact that unlike in the rest of the country, after grade eleven Quebecers enroll in a publicly-funded 2- or 3- year college prepatory program known as CEGEP.[17] A *diplôme* from CEGEP is required for admission to Quebec's universities, however Quebecers spend only three years in university before attaining their bachelors. The incentive to continue with formal education after high school is higher in Quebec because the first year of post-secondary studies is, to put it simply, completely paid for by the provincial government.[18]

The fact makes it difficult to compare educational attainment in Quebec with the other provinces as the provincial subsidy *à la* CEGEP means that attainment will be skewed in the Quebec sample.

[17] In French, *Collège d'enseignement général et professionnel.*

[18] According to one publication by Statistics Canada, 64% of young Quebecers attend CEGEP. This is well over twice the college participation rate for most other provinces. See Shaienks and Gluzynski (2008).

To avoid dealing with these difficulties, I have elected to drop the Quebec observations from the sample. I leave the task of working them back in to future research.

3 Empirical Findings

3.1 Preliminary Results

To verify that my model is correctly identified, I begin by running an OLS regression of educational attainment on the set of exogenous controls alone. The goal is to see whether the individual and family characteristics that I have included in the model really do help explain some of the variation in outcomes.

The results are presented in column (1) of Table 1. In general the estimates are of the correct sign and significance. Parental income and education level are both positive and significant determinants of outcomes. Females, native-born Canadians, and children of French-speaking parents all tend to stay in school longer, although the effect for the latter group is not statistically significant. As expected, PISA scores are positively associated with attainment, while student-teacher ratio is negatively associated. Interestingly, the rural school school dummy is not a significant determinant of outcomes.

To continue, column (2) presents the OLS estimates once dummies for school type are included. For the most part, the coefficients on the control variables remain unchanged. However, the results suggest that variation in educational attainment is not very well explained by school type: only the coefficient on the English separate school dummy is significant. Oddly, immersion programs in the public system have a *negative* effect on years, while immersion programs in the separate systems have a positive effect on years. French schools in both the public and separate systems, as well as private schools, are all positively associated with the dependent variable but again the relationship is not statistically significant.

We should not get too hung up on these results for, as explained earlier, a strong case can be made for an alternative estimation procedure, as explained earlier. If school type is endogenous as we suspect, then the OLS estimators are biased and inefficient. They are, in a word, meaningless.

$N = 8,316$	(1)	(2)
PISA Reading	0.004^{***}	0.004^{***}
	(24.50)	(24.25)
Parent Income ($000s)	0.002^{***}	0.002^{***}
	(4.72)	(4.77)
Intact Family	0.392^{***}	0.394^{***}
	(9.22)	(9.25)
Foreign Born	-0.118^{*}	-0.127^{*}
	(-1.80)	(-1.92)
Female	0.390^{***}	0.393^{***}
	(11.84)	(11.90)
Rural School	0.003	-0.011
	(0.09)	(-0.28)
Student-teacher Ratio	-0.019^{***}	-0.017^{***}
	(-3.04)	(-2.72)
Mother Education		
Some Post-Secondary	0.145^{***}	0.145^{***}
	(3.95)	(3.93)
Bachelors Degree	0.517^{***}	0.517^{***}
	(10.32)	(10.28)
Graduate Degree	0.588^{***}	0.578^{***}
	(7.32)	(7.18)
Language - Mother		
English	-0.203^{***}	-0.214^{***}
	(-3.67)	(-3.83)
French	0.145	0.097
	(1.48)	(0.89)
Language - Father		
English	-0.082^{*}	-0.081^{*}
	(-1.71)	(-1.69)
French	0.012	-0.038
	(0.12)	(-0.35)
French	-	0.102
		(0.69)
Immersion	-	-0.070
		(-0.96)
Private	-	0.095
		(0.96)
English - Separate	-	-0.095^{**}
		(-2.06)
French - Separate	-	0.152
		(0.99)
Immersion - Separate	-	0.145
		(0.97)
Adjusted-R^2	0.136	0.136

Population weights used. t-values in parenthesis. Dummy variable for missing student-teacher ratio observations not reported. Reference Category: English public schools.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1: Years of Schooling: OLS

3.2 3SLS Estimates

I explore the issue of endogeneity by estimating the model using 3SLS, where school type is instrumented by distance. The results are presented in the first column of Table 2. A few points of note before discussion. First, not all exogenous variables from the "years" equation enter into the "type" equations. In particular, I have left out student-teacher ratio because it does not tell us anything about the *family* that is making the schooling decisions. Second, only one distance variable will enter the right-hand side of any given "type" equation. This would be impossible in a regular IV regression where all the instruments are thrown in together. For example, the distance to the nearest private school (relative to the nearest English school) can help explain choice of private school. But that same choice cannot be independently explained by distance to, say, the nearest French school (relative to the nearest English school).

The estimates have some interesting properties. In general, the controls for individual characteristics remain unchanged with the same signs and significance level as before. Notably, the sign on the rural school dummy has been reversed and it is now significant. Students in rural communities get about eight months less schooling over the course of their formal education relative to students in non-rural areas. Although not reported, the coefficients on distance on the right-hand side of the "type" equations are all negative and significant at the 99% level. This is exactly as we would expect, for it means that a parent is less likely to choose a particular school the farther away that school is relative to her other options.

Turning to the school coefficients themselves, we see that enrollment in a French immersion program or French school is positively associated with attainment, but only the latter is significant. Relative to English school students, French school students attain an average of 3.83 more years of education. This is a big number, but perhaps not outrageous. It says that the attainment gap between French high school students and English public school students is about the difference between a high school diploma and a bachelors degree, in terms of years of schooling. This is consistent, to some degree, with 2001 census statistics that show that French speakers in every province other than Quebec are more likely than English speakers to attain post-secondary credentials (Corbeil, 2003).

	Full Sample (N=9,616)	Non-French (N=8,317)	French (N=1,299)
PISA Reading	0.006***	0.005***	0.005***
	(8.15)	(7.82)	(6.60)
Parent Income ($000s)	0.004***	0.002***	-0.006
	(3.67)	(2.86)	(-1.56)
Intact Family	0.629***	0.502***	1.145***
	(5.42)	(4.31)	(2.97)
Foreign Born	-0.350***	-0.201**	1.367
	(-2.96)	(-2.17)	(0.95)
Female	0.231***	0.244***	0.852***
	(4.25)	(4.41)	(2.96)
Rural School	-0.673***	-0.191	-0.724
	(-2.56)	(-1.04)	(-1.75)
Student-teacher Ratio	-0.667***	-0.075*	-0.044
	(-3.10)	(-1.85)	(-1.12)
Mother Education			
Some Post-Secondary	0.249***	0.159***	-0.154
	(3.99)	(2.72)	(-0.82)
Bachelors Degree	0.562***	0.449***	-0.156
	(6.61)	(5.87)	(-0.40)
Graduate Degree	0.816***	0.680***	-0.305
	(4.53)	(3.57)	(-0.39)
Language - Mother			
English	-0.573***	-0.325***	0.323
	(-4.14)	(-2.88)	(0.81)
French	-1.133***	-	0.310
	(-2.51)		(0.51)
Language - Father			
English	-0.276***	-0.214**	0.430
	(-2.61)	(-1.97)	(1.47)
French	-0.233	-	-0.420*
	(-0.88)		(-1.71)
French	3.829**	-	-0.844
	(1.89)		(-0.89)
Immersion	0.402	1.707*	-2.031
	(0.54)	(1.75)	(-0.60)
Private	-13.180***	-7.317**	19.493***
	(-2.78)	(-2.02)	(3.91)
English - Separate	-1.913***	-0.777**	-2.994*
	(-4.08)	(-2.13)	(-1.78)
French - Separate	-1.671	-	0.460
	(-0.75)		(0.44)
Immersion - Separate	-3.257	5.063	4.268
	(-0.62)	(0.89)	(1.28)
Root MSE	2.809	2.139	2.950

Population weights used. z-values in parenthesis. Dummy variable for missing student-teacher ratio observations not reported.
Results for "type" equations not reported.
Reference Category: English public schools.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2: Years of Schooling: 3SLS

The separate school system tends to perform particularly poorly relative to the English public system. Attending a separate school of any type has a negative effect on attainment. Enrollment in an English or French separate school is associated with a 1.91- and 1.67-year decline in total schooling, respectively. Students in separate immersion programs see an even greater decline in their outcomes.

The coefficient on the private school dummy is perplexing. Contrary to what we might expect, private school enrollment is associated with a 13.18-year decline in average total schooling. This makes little sense. Unfortunately, the time I was able to spend working with the data was quite limited, and so was unable to address this anomaly.

Until now I have assumed that French and English speakers face the same set of choices when it comes to education. But as discussed in this paper's introduction, Anglophones who do not have at least one French-speaking parent are not automatically entitled to French-language education. Therefore, I am interested to see how the model performs for the sample of respondents who do not have at least one Francophone parent.

Column two of Table 2 gives the relevant estimates of the restricted model with French and French separate schools left out. Perhaps the most exciting finding is that French immersion is now a positive and significant determinant of attainment. Immersion students with Anglophone parents, on average, attain 1.7 more years of education relative to their English public school peers. This could be because bilingualism enhances immersion students' cognitive ability, or results in some other "hidden" personal characteristic that makes them more likely to continue with their formal education. [19]

As before, English separate schools underperform the public schools. Whereas in the combined sample English separate education is associated with a 1.91-year decline in total attainment, that number drops to 0.78 years in the non-French sample. Thus, there is evidence to suggest that some institutional feature of the separate school system is contributing to their students' lower educational outcomes.

How do the estimates change for the group of respondents with at least one French-speaking

[19] Students in separate immersion programs also do better, but the finding is insignificant.

parent? The estimates are presented in column three of Table 2. Three important results stand out. First, some of the control variables which so nicely predicted educational attainment in the non-French and combined samples are now insignificant. Parental income, education, and mother tongue are now especially poor predictors of outcomes. Second, I find once again that English separate school students have lower attainment relative to English public school students. Indeed, in all three samples I find that English separate schooling is negatively associated with years and is statistically significant throughout, although the magnitude of the effect is different for each. Finally, I find that French school is negatively associated with years but that the effect is insignificant. This stands in direct contrast to the combined sample where French schooling was shown to be a positive and significant determinant of outcomes.

4 Conclusion

Canada's education system is unique in that parents can choose from a variety of different secondary schooling options for their children. But until now, little work has been done to discern how these options relate to a child's educational attainment. This paper asks the question: does type of secondary school really matter for post-secondary outcomes?

The answer, in short, is yes, but only after we account for the fact that school choice is endogenous. In general, I find that students in English separate schools have lower attainment relative to students in English public schools. I also find that students from French schools attain nearly four more years of education relative to the reference group. For the sample of children of Anglophone parents, I find that immersion programs are positively associated with attainment. This is perhaps the paper's most exciting finding, and one that should guide future research. The results for the French sample are odd, and the estimates on the control variables do not fit the trend established in my earlier results. Still, I find that English separate school students in this group have lower outcomes, which is consistent with my other results.

While this paper presents some solid preliminary findings, much work remains. For one, many of the anomalies that show up in the estimates – the unexplainably large coefficients on the private school dummies, for example – could no doubt be tamed with more time spent working with the data. Thanks to the notoriously slow-moving research approval process at Statistics Canada, some critical pieces of the YITS survey were only made available to me a few weeks before the project deadline. With time, I could have developed a more finely tuned specification of the model, as well as done a lot more hypothesis testing.

A number of avenues persist for future research. First, the results clearly indicate that something must be done about the private school dummy. One thing might be to develop a two stage model of choice, where the parent decides between public and private schooling in the first stage, and then between the various types of public schools in the second stage. The two instruments in the first stage would be distance to the nearest private school and distance to the nearest public school of *any* type. The instruments in the second stage would remain unchanged. This two-stage model is probably a closer approximation to the way such decisions are made in real life. Second, I need to find a way to better define the "years" variable. Many of the respondents in the most recent cycle of the YITS were "continuers," meaning that they were currently enrolled in, but had not yet completed, a postsecondary program. In my model, this extra information was ignored. Finally, future studies should develop an appropriate way to reinsert Quebec into the sample. While dropping Quebec had the advantage of making this study more tractible in the short term, it also meant that a lot of information was excluded from the analysis – information that, if included, could have drastically altered the results.

References

[1] *Canadian Charter of Rights and Freedoms*, s. 23, Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (U.K.), 1982, c.11.

[2] *Policy statement and guidelines on the admission, welcoming, and support of students in French-language schools in Ontario*. Ontario Ministry of Education, Toronto, 2009.

[3] Alberta Department of Education. *Public and separate schools*. Available online at http://education.alberta.ca/parents/choice/public.aspx. Last accessed 1 August 2010.

[4] Alberta Department of Education. *Public, separate, francophone, and charter authorities*. Available online. Last accessed 1 August 2010.

[5] Kelly Bedard. Human capital versus signaling models: university access and high school dropouts. *Journal of Political Economy*, 190(4):749–775, 2001.

[6] British Columbia Ministry of Education. *School and district contacts*. Available online. Last accessed 1 August 2010.

[7] Russell Davidson and James G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2004.

[8] Giorgio Di Pietro and Andrea Cutillo. Does attending a Catholic school make a difference? Evidence from Italy. *Bulletin of Economic Research*, 58(3):193–234, 2006.

[9] Dissemination and Reporting Unit. *List of publicly funded schools*. Prepared on 23 June 2010. Ontario Ministry of Education.

[10] Education and Finance Facilities. *Écoles immersions 09-10*. Prepared on 18 June 2010. Saskatchewan Ministry of Education.

[11] Education Finance and Facilities. *2009 active list of Saskatchewan schools*. Prepared on 18 June 2010. Saskatchewan Ministry of Education.

[12] William N. Evans and Robert M. Schwab. Finishing high school and starting college: do Catholic schools make a difference? *The Quarterly Journal of Economics*, 110(4):941–974, 1995.

[13] Ross Finnie and Richard E. Mueller. The backgrounds of Canadian youth and access to post-secondary education: New evidence from the Youth in Transistion Survey. In Ross Finnie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters?*, pages 79–107. McGill-Queen's University Press, Montreal and Kingston, 2008.

[14] French Programs and Language Initiatives. *Liste écoles immersion 2009-2010.* Prepared on 7 June 2010. British Columbia Ministry of Education.

[15] Claudia R. Hepburn. Learning from success: what Americans can learn from school choice in Canada. *Mackinac Center for Public Policy*, 17 November 2002.

[16] Caroline M. Hoxby. Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238, 2000.

[17] *List of publicly-funded schools in Ontario offering French immersion and/or extended French.* Prepared on 21 June 2010. Canadian Parents for French (Ontario).

[18] Manitoba Textbook Bureau. *Schools in Manitoba - Écoles du Manitoba.* Catalogue #83855. Manitoba Education, Citizenship and Youth. Available Online.

[19] Anne Motte, Hanqing Qiu, Yan Zhang, and Patrick Bussière. The Youth in Transition Survey: following Canadian youth through time. In Ross Finnie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters?*, pages 63–75. McGill-Queen's University Press, 2008.

[20] Derek Neal. The effect of Catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15(1):98–123, 1997.

[21] New Brunswick Department of Education. *K-12 school directory.* Available online. Last accessed 1 August 2010.

[22] Newfoundland and Labrador Department of Education. *Database of schools.* Available online. Last accessed 1 August 2010.

[23] Nova Scotia Department of Education. *2009-2010 Nova Scotia directory of public schools.* Available online. Last accessed 1 August 2010.

[24] Prince Edward Island Department of Education. *Directory of schools and school board offices in Prince Edward Island 2009-2010.* 1 September 2009. Available online. Last accessed 1 August 2010.

[25] Danielle Shaienks and Tomasz Gluszynski. Postsecondary education - participation and dropping out: Differences across university, college and other types of postsecondary institutions. *Culture, Tourism and the Centre for Education Statistics Research Papers*, Ottawa, Statistics Canada, 2008.

[26] Statistics Canada. 2010. *Youth in Transition Survey.* Government of Canada.

[27] Statistics Canada. *1996 Census of Canada: Enumeration Areas.* Distributed by Tetrad Computer Applications. 1996. Available from PCensus for Mappoint, Queen's University Library.