Cycling for Education? Heterogeneous Preferences for Academic Tracks at Secondary School[†]

Sergio Parra-Cely

February 15, 2019

Using a unique dataset from the Netherlands, we estimate the parental demand for academic programs in a free secondary school system. Our specification, a multinomial discrete model in combination with an instrumental variables approach, controls for the endogeneity of commuting costs, school's attributes, and the characteristics of pupils attending education. By recovering consistent willingness to travel estimates, we find that parents with higher socio-economic background let their children commute 1.5 kilometers more to attend school (equivalent to a 0.5 standard deviation effect), and that their children enroll more often into schools with academic tracks that grant direct admission to university. These academic tracks are socially valuable, explaining an average welfare loss of 1 Euro (2.8 kilometers commuted) per day at school for those children without the opportunity to attend these programs. Our results support the notion that less educated parents sort into academic tracks with lower economic prospects, even in school systems where restrictions to choice are negligible.

Keywords: Education, School Choice, Early Tracking, Characteristics Model.

JEL Classification: I20, I21

[†]USFQ School of Economics, Diego de Robles y Vía Interoceánica, Quito, Ecuador. This paper served as part of my PhD dissertation in Economics at Maastricht University. I am indebted to Lex Borghans and Bart Golsteyn for their guidance, as well as to Roxanne Korthals, Stan Vermeulen, Omar Rodriguez, Clotilde Mahe, and participants at the DUHR Seminar at Maastricht University and the Lisbon 4th Workshop on Economics and Econometrics of Education for their valuable comments and suggestions. I am also grateful to Tyas Prevoo for his help clarifying some aspects of the database (email: sparrac@usfq.edu.ec).

1 Introduction

Identifying the degree to which vulnerable groups engage in suboptimal schooling decisions is crucial to evaluate the effectiveness of education as a vehicle to promote human development and intergenerational mobility (Manski, 1992; Burgess and Briggs, 2010). Large costs of attendance, administrative restrictions, congestion, and other market frictions, are arguments provided by a large body of literature on school choice to explain how children, with the same academic ability and motivation, end up with unequal educational outcomes (Hoxby, 2003). Furthermore, confounding factors influencing parental valuations for academic features make it difficult to isolate a "desire for education" effect. Thus, traditional hedonic regressions and multinomial discrete choice models used in the school-choice literature, while easily tractable, fail to recover decision makers' differences in tastes induced by variation on school's features (Nevo, 2000; Bayer and McMillan, 2005). An alternative to circumvent these problems is to implement econometric methods designed to identify endogenous preferences for schooling, as well as studying specific educational markets with a large variety in the alternatives offered but no restrictions to choice in place.

In this chapter we propose a framework to recover households' preferences for academic programs in the presence of educational sorting. By considering parents' revealed decisions in an unconstrained school system with early tracking and large product differentiation, the objective of this study is twofold. First, we aim to provide a consistent estimation of the existing enrollment gap into the higher academic tracks at secondary education. This enrollment gap privileges the schooling outcomes of students from higher socio-economic family backgrounds, relative to pupils with less fortunate upbringings but comparable academic performance in primary education. Hence, quantifying the extent to which non-academic household's attributes affect the likelihood of human capital investments is of paramount importance for the design of educational policies that account for such differences. Second, this study seeks to recover the heterogeneous preferences for schooling options that parents with diverse educational attainments might exhibit, as an attempt to identify the elements that determine the enrollment gap.

The secondary school system in the Netherlands has a large tradition encouraging participants to choose educational options freely. While regulated by the Inspectorate of Education, publicly and privately managed schools are characterized by being equally funded with central government's revenues, implying that almost no tuition fees are charged. Moreover, the small variation in geographic conditions and the easy anticipation of changes in climate privilege cycling as the main mode of transportation (Van Goeverden and De Boer, 2013). Without catchment areas, or other sophisticated mechanisms to deal with congested markets, the Dutch case provides a good opportunity to identify educational sorting incentives that are not determined by restrictions to choice, which is a novelty with respect to the previous literature on the economics of education.[1]

Students' performance at primary school plays a key role in determining future enrollment outcomes. The secondary education system in the Netherlands is characterized as an early tracking system where pupils, after finishing primary school, are required to attend one among three academic options to continue their education: A higher track that, after completion, grants direct admission to university programs, a lower track which gives admission to community college education, and a vocational track that accounts for rather technical, on-the-job training needs. Since the right to pursue tertiary education at the university level depends on whether the high track is selected, suboptimal schooling decisions occur when children are enrolled in lower tracks but prior academic performance would suggest they are suited to attend the higher track.

This study starts by estimating a model of demand for differentiated products, applied to the education market (Berry et al. 1995, 2004; hereafter BLP). After classifying pupils by their municipality of residence and parental educational attainments, we exploit variation in both students characteristics and their choice sets, within and across markets, to recover unobserved mean valuations for all school-tracks observed. The specification considered here decomposes the utility of choosing an academic program into the satisfaction level that is common to all students in a market where that option is selected, and deviations from mean preferences which are determined by students' specific characteristics. In order to obtain these valuations, we implement the contraction mapping algorithm suggested in the literature of demand estimation to ensure that, for a given set of pupils' attributes, mean preferences equalize the observed market shares -understood as the fraction of children at each market who select a particular option- with the ones the empirical model predicts (Berry et al., 1995).

By estimating individual enrollment probabilities with the BLP model, we find that students with highly educated parents (i.e. at least one parent obtained a post-secondary degree) are 24% more likely to be in the high track, relative to equally capable pupils with low-educated parents (i.e. parents whose maximum educational attainment is high-school or below). We compare individual enrollment probabilities with estimates based on more simple, conditional logit demand models. Results from this exercise suggest that traditional models either underestimate or neglect the enrollment rate gap among pupils from different socio-economic background. The BLP approach is then more appropriate to estimate these relationships as traditional models do not account for substitution patterns in the demand for academic programs that depend on students' characteristics, inducing different schooling decisions for a given choice set.

The analysis continues by decomposing mean utility levels obtained from the BLP

[1] Some restrictions might exist in big cities such as Amsterdam, where due to capacity constraints other mechanisms such as waiting lists, random lotteries, and clearinghouses have been recently implemented. However, as we will elaborate later, the education system considered here is free from such interventions.

model on the different academic features per school-track combination to recover parental preferences for academic programs, school denomination, and observed quality, among other characteristics. In our model, we expect unobserved features to be correlated with the average distance students commute to attend secondary education, in the same way as the price of an item can be correlated with its unobserved quality. We deal with this endogeneity problem by exploiting the plausibly exogenous variation of rival options outside each school's "geographical area of influence" as an instrumental variable to provide consistent mean willingness to travel (WTT) estimates for a particular bundle of educational features.

Specifically, we construct different concentric rings (distance buffers) to calculate the number of rival schools located between say, a and b kilometers *away* from each school, in order to account for the fact that competition is spatially concentrated but uncorrelated with latent school-track covariates. Since both residential and school locations are not exogenous, in order for our identification strategy to work we need to assume that schools' locations in the arbitrarily-built areas outside the influential zone of each institution are not correlated with unobserved elements that influence the educational choices observed in the data. Moreover, we expect the larger the number of rival schools located around each option, the lower the final average distance traveled per school-track. We also include the cycling distance of each educative institution to the closest city hall as an additional instrument to account for each school-track's degree of accessibility from urban and rural areas (Noailly et al., 2012). Hence, we expect the closer the distance between the school and the city-hall is, the higher its accessibility and the lower the average cycling distance pupils will travel to receive education.

Finally, we conduct differences-in-means tests by parental educational attainments on WTT estimates to assess whether parents self-select their children into higher tracks. Results from this (second) estimation stage show that, considering the pure vocational track as a baseline option, parents value all remaining tracks positively. However, in high-educated families parents are willing to let their children commute 0.8 kilometers (0.3 standard deviations) more (less) to attend the higher (lower) track. This asymmetry in preferences is almost fifty percent smaller than the one obtained using traditional hedonic travel models, suggesting that the IV estimator accounts for the potential overestimation of parental WTT for education. All results combined provide a clear evidence of educational sorting based on parental socio-economic status, even after controlling for pupils' previous academic performance.

We also conduct extensions of the model to evaluate how variation of preferences for academic tracks depends on prior academic performance, as well as on primary school's peers. In particular, lower preferences for the high track are more likely to develop when the pupil's performance at the primary school exit test does not belong to the group of top 10% performers. Surprisingly, neither receiving positive recommendations to attend the high track, nor having more primary school peers choosing the same alternative, modify preferences of low educated parents. We claim that these results can partially be attributed to the fact that pupils from highly educated parents have better opportunities to build academic profiles that make them more attractive for secondary schools.

Our model also allows us to conduct a policy simulation in order to assess the value of the different academic tracks offered after primary school is finished. The policy counterfactual exercise implemented here is to calculate the utility loss decision makers would face as all alternatives pertaining to a particular track are removed from the original choice set, forcing them either to commute longer distances to attend the desired track, or choosing a lower track available at reach. Using a compensated variation approach we find that, in absence of the higher track, commuting costs for the average student increase by 2.8 kilometers, implying a welfare loss of 1 Euro per day attended at school. We observe qualitatively similar (but quantitatively lower) results for the lower track, suggesting that the high track is the academic option that socially is the most valuable. To our knowledge, these are the first of such kind of estimates for the Dutch education system.

This chapter contributes to the literature on parental preferences for schools. From a demand perspective, many papers address the extent to which education systems with barriers to free selection have a detrimental effect on both students' and schools' educational outcomes (Hoxby, 2003). Because school choice is an endogenous phenomenon, many of these studies try to exploit random variation in school competition by the use of natural experiments on admission procedures (Rincke, 2006; Cullen et al., 2006), introduction of vouchers (Hsieh and Urquiola, 2006; Figlio and Hart, 2014; Chumacero et al., 2011), variations on housing prices along school district boundaries (Black, 1999; McMillan, 2005), and availability of information about school quality (Hastings and Weinstein, 2008; Hastings et al., 2009), in order to consistently measure preferences for academic attributes. This study brings additional evidence about the presence of incentives to self-select into suboptimal academic options, in a context where the role of latent cofactors that determine final schooling decisions is negligible because of the absence of transaction costs, market frictions, and administrative restrictions that will impair parents and/or children to reveal their true preferences.

Moreover, this chapter also intends to provide updated estimates on the heterogeneity of parental preferences for education in the Dutch secondary school system. Koning and Van der Wiel (2013) study the effect of school-quality information on school choice using individual and school level information of first year students at secondary schools. Despite the fact that schools' quality scores have a positive relation with the number of new students enrolled in the first year, the authors recognize that no causal inference can be extracted from individual choice data. Their paper presents commuting distance to each school as the leading determinant of choice, with pupils willing to travel 220 meters on average to attend an institution with higher quality. In a related article, Cabus and Cornelisz (2014) use kernel estimation techniques to match students in different education markets, in order to assess the effect of market competition on both academic sorting and students' performance. After controlling for pupil and household level information, these authors find that schools use supply-driven product differentiation in order to attract high-achieving students. Nonetheless, higher ability schools converge, in terms of performance, to lower ability schools by accepting "students at the margin" in order to complete their enrollment requirements. Using administrative data from the city of Amsterdam, Ruijs and Oosterbeek (2014) analyze the determinants of secondary school choice. Their results contradict previous findings by suggesting that school quality is an inconsistent predictor of revealed parental choices. Moreover, they find a significant, positive primary school peers' effect on final enrollment decisions. Close to the spirit of this research, Borghans et al. (2015) estimate mixed logit models to recover parental preferences for primary schools. They find that lower educated parents tend to choose underperforming schools, and overall, aspects such as distance, religious affiliation, and the implementation of an alternative teaching philosophy are key determinants of school choice. With the notable exception of Cabus and Cornelisz (2014), the previous papers do not address the potential endogeneity of school choice. Hence, this research seeks to fill this gap by explicitly addressing the potential correlation between traveling costs and unobserved attributes at the school-track level.

Finally, this chapter aims to discuss the effects of traveling costs on educational choices. Alderman et al. (2001) find a positive effect of reducing commuting distances on enrollment rates in private schools for poor children in Pakistan. Examining the Chilean case, some evidence has been found of the impacts of traveling costs on schooling decisions, showing that students are willing to commute more to avoid paying larger fees (Chumacero et al., 2011). Comparable results have been found for primary education in the Netherlands, where parents are willing to let their child commute 300 meters more to avoid attending education at lower quality schools (Borghans et al., 2015). However, none of these studies addresses the potential correlation between unobserved academic attributes and pupils' commuting costs. The present study is closely related, in spirit and methodology, with the research developed by Carneiro et al. (2016). After instrumenting commuting distances with the distance from each household to other relevant facilities (Hospitals), they find that students not only prefer to attend education at closer options, but also are willing to pay at least three quarters of the average annual fees at private schools to reduce their commuting distance by 500 meters. The WTT estimates we obtain are also higher than the ones suggested by previous research. We claim related papers underestimate parental WTT because of i) the downward bias in such coefficients when the correlation between the student's commuting distance and other latent cofactors at the school-track level is not addressed, and ii) the heterogeneity in transportation modes that reduces the parental sensitivity to changes in the distance their children travel to attend school.

This chapter proceeds by first offering a brief overview of the Dutch education system. Section 3 describes the empirical strategy implemented, emphasizing on the different roles of both individual and school-track information. Section 4 presents the data overview. In Section 5 we discuss the parameter estimates, as well as results on WTT calculations. Finally, Section 6 concludes.

2 The Dutch Education System: An Overview

Since 1917, the Dutch government has guaranteed freedom of education consisting in (1) giving citizens the right to establish schools that embrace their own educational, religious, and social beliefs; and (2) allowing parents to freely decide, at any academic level, to which educational institutions their children should register. Education is publicly funded. Hence, both privately and publicly managed schools receive the same resources, but they are subject to quality monitoring from the Inspectorate of Education. Schools that fail to comply the Inspectorate's quality standards will not receive funding, but this occurs seldomly. In practice, this implies that education is almost free, with no administrative restrictions placed in order to influence school choices. Moreover, the fairly accurate weather forecasts, and the little variation in geography, induce parents to select schools that are close to their residence, with very low incentives to change their home location in order to live close to preferred schools (Koning and Van der Wiel, 2013).

Education is mandatory at the primary and secondary education levels until age sixteen. From age 4 to 12, students attend primary school, where they receive a unified curriculum of courses and preparatory activities. In their last year of elementary education, pupils are required to present a nationwide, standardized exam (in Dutch: CITO), that serves to build an objective performance measure for pupils and schools. In addition, each student receives a (subjective) recommendation from the primary school teacher about the academic track they are most suited to follow afterwards. From age 12 to age 14, most students attend general secondary education. By the end of the second year, schools will consider current performance and previous performance indicators from elementary school, to suggest students a placement in one among the three different academic tracks the system offers. The pre-university track, or high track (in Dutch: VWO), is designed to prepare students interested in post-secondary education at the university level. The higher vocational education, or lower track (in Dutch: HAVO) provides sufficient instruction to attend higher professional education at the community college level. Finally, the system offers a pure lower vocational track (in Dutch: VMBO) which trains pupils to receive further non-academic, technical training after finishing high school. Mobility between tracks is possible but costly. For instance, upper mobility from the low track implies taking one or two additional years of instruction, relative to the time will take to complete the high track from the beginning. This implies that parents concerned to facilitate their children with opportunities to attend university will be better off by choosing the pre-university track.

Starting the third year of secondary education, the school proposes a tracking allocation to the parents, incorporating information about the pupil's previous academic performance indicated by both the CITO tests scores and the advice given by the primary school. While secondary schools have enough discretion to abide the tracking recommendation from the pupil's primary school, parents can indirectly influence the final tracking result as well. They can exert pressure to elementary school's teachers seeking to obtain a better recommendation for their child, or bargain with the secondary school to guarantee enrollment in the academic track that fits their preferences the best (Korthals, 2015). In practice, this implies that parents can decide whether to comply or not with the final tracking outcome by approaching many secondary schools until the child is accepted in the desired track.[2] At the end of secondary education, students present a national standardized exam that accounts for fifty percent of final grades. The remainder percentage is determined by school-specific exams that are held between the last two or three years of secondary education (Ruijs and Oosterbeek, 2014). In order to assess the quality at the secondary education level, the Inspectorate of Education monitors results at the national standardized exam, as well as the percentage of pupils that failed to graduate in time. This accountability policy gives schools enough incentives to refine their tracking placements, either through being more selective, or by assessing information from primary schools more carefully.

3 Empirical Strategy

In this section we present an adaptation of the methodology proposed by Berry et al. (2004) to recover taste parameters from a model of choice with customer heterogeneity and product differentiation. The key aspect of this estimation strategy is to incorporate micro-level and market level information, in contrast with previous contributions that use market level aggregated data (Berry et al., 1995; Nevo, 2000) or individual data alone (McFadden, 1974). We model the decisions to choose a particular school-track at the end of the second year of secondary education. In this sense, as most of the tracking decisions are made at the school level, this model is coherent with the parental decision to comply or not with this final track allocation.

3.1 A Model of School-Track Choice

Consider a secondary education system where T different academic markets are observed,

[2]To our knowledge, there is no legal restriction that impairs parents to implement such strategy.

with $t = 1, ..., T$ and $i = 1, ..., N_t$ indexing each market, and the number of pupils per market, respectively.[3] By using this classification we are accounting for geographic and socio-demographic variation that might affect both, the composition of students in each selected option, and the set of alternatives that all pupils from a particular market can choose from. Hence, a choice set is composed by $1+J_t$ alternatives, where $j = 0$ represents an outside option common to all markets, and $j = 1, ..., J_t$ indexes the remaining available alternatives at market t.

We define an alternative as a school-track pair. The main purpose of this demand model is to account for the selection process made by parents on the academic track their child should follow. These academic tracks, after completion, grant the student the right to attend post-secondary education at the university level. Thus, we model school-track choices between a "High Track", that allows the pupil to enroll in a scientific university program, and a "Low Track", which leads the student to register at a mixed vocational-academic program offered by a university of applied sciences[4]. We set the outside option as all school-track combinations leading to a pure vocational high-school track degree. Hence, this outside option has a non-zero market share across all markets, and market heterogeneity is largely explained by the availability of remaining alternatives.[5]

We suggest the following specification of the indirect utility obtained by student i from attending school-track j at market t :

$$u_{ijt} = \gamma x_{jt} + \xi_{jt} - \beta_i \overline{d_{jt}} + \varepsilon_{ijt}, \tag{1}$$

where $x_{jt} = \{x_{jt1}, ..., x_{jtk}\}$ is a vector of k observed attributes of school-track j, e.g. school-denomination, type of track, and quality aspects such as average standardized test scores and percentage of pupils who fail to graduate in time. ξ_{jt} denotes all attributes of school-track j at market t that are unobserved by the econometrician, but observed by the decision makers. With respect to the cost of attending classes at a particular school-track, let $\overline{d_{jt}}$ denote the average cycling distance a student from market t has to travel in order to attend education at option j.[6] Finally, ε_{ijt} is an individual mean-zero

[3] We recognize that schooling decisions are strongly influenced by parents' preferences. In this approach we are modeling the final choice made as a result of a bargaining process between all members of the household. Hence, we will use "Parents', "Pupils", "Students", "Households", and related terms interchangeably.

[4] Universities of applied sciences (hogeschool) offer education at the level of community colleges in the US education system

[5] This definition of the outside option is possible since education in the Netherlands is mandatory until the age sixteen. Hence, all pupils should attend education at one of the educational tracks offered by the Dutch educational system.

[6] We focus on this particular notion of distance as a way to capture preferences for commuting that do not depend on pupils' residence location alone. For instance, we can consider the hypothetical case where all pupils in a market face the same distance for each option and their covariates determine the propensity to travel to a certain destination. An alternative approach would be to work with individual commuting distances, but this is an individual characteristic rather than a school-track attribute to control for. Namely, working with average distance per alternative implies that commuting costs have a direct role on determining demand for the different options available in a market.

stochastic term with a type-I extreme value distribution.

The utility specification in (1) implies that the heterogeneity of preferences for academic tracks and other school attributes is driven only by how much parents value the average commuting costs their children face to attend school-track j. We denote this valuation by β_i. We assume that the heterogeneity in such valuation is explained by a set of observed individual covariates D_i, which include demographic attributes and all relevant academic information that influences the final school-track outcome.[7] In order to maintain simplicity, we adopt a linear relationship in parameters:

$$\beta_i = \beta^m + \beta^o D_i, \tag{2}$$

where β^o indicates the weights of observed individual characteristics on preferences for cycling. It is important to point out that, because it is possible to add an individual specific constant without changing the preference order over all alternatives, we can normalize utility levels by fixing the indirect utility value of the outside option to zero. This implies that $u_{i0t} = 0 \; \forall i \in n_t, \; \forall t \in T$. Therefore, utility levels now should be interpreted as the differences in utility among all options, relative to the utility level perceived by attending a vocational track.

Substituting equation (2) into (1), the student-level choice model is obtained:

$$\begin{aligned} u_{ijt} &= \delta_{jt} - \beta^o \overline{d_{jt}} D_i + \varepsilon_{ijt} & (3) \\ \delta_{jt} &= \gamma x_{jt} - \beta^m \overline{d_{jt}} + \xi_{jt}, & (4) \end{aligned}$$

where δ_{jt} captures mean-valuations of school-tracks characteristics that are invariant across students. Equations (3) and (4) also show that is possible to decompose the indirect utility in a linear component (i.e. the mean utility level), plus a non-linear term that measures deviations from mean preferences that depend on the student's covariates.

Regarding individual's heterogeneity, we define a pupil as a duple (D_i, ε_{it}) with $\varepsilon_{it} = (\varepsilon_{i0t}, ..., \varepsilon_{iJ_t t})$. Then, the set of students at market t that choose to attend school-track j is defined as the group whose attributes lead them to select such option against all other available alternatives:

$$A_{jt}\left(x_{jt}, \overline{d_{jt}}, D_i; \beta^o\right) = \{(D_i, \varepsilon_{it}) \mid u_{ijt} \geq u_{ilt}, \; \forall l = 0, 1, ..., J_t\}.$$

Let us denote $\delta_t = \{\delta_{1t}, ..., \delta_{J_t t}\}$, $x_t = \{x_{1t}, ..., x_{J_t t}\}$, and $\overline{d_t} = \left\{\overline{d_{1t}}, ..., \overline{d_{J_t t}}\right\}$. Including the

[7] To understand how variation on households' characteristics imply different preferences for academic programs consider the following example: Assume that there are only two families that share the same choice set. As these families differ in many dimensions, their preferences for commuting to the same alternative will vary as well. In specification (1) the main implicit assumption we impose is that there are no heterogeneity in preferences for academic tracks that is explained by any other reason than differences in average commuting costs' valuations that, in an hedonic model perspective, will imply different willingness to travel estimates. By restricting our model to account for this transmission mechanism alone we make sure that heterogeneous preferences for academic tracks are revealed by the cycling distance pupils commute.

assumption that ties do not occur, and recalling that ε_{it} follows a type-I extreme value distribution, the probability that a pupil from market t with characteristics D_i prefers alternative j takes the logit form:

$$\hat{P}_{ijt}(x_t, \delta_t, \overline{d_t}, D_i; \beta^o) = \frac{exp(\delta_{jt} - \beta^o \overline{d_{jt}} D_i)}{1 + \sum_{l=1}^{J_t} exp(\delta_{lt} - \beta \overline{d_{jt}} D_i)} \tag{5}$$

where $P(.)$ denotes population distribution functions. Since all pupils from a particular market select their preferred alternatives from the same choice set, aggregate (predicted) market shares are recovered by calculating the average of all individual probabilities of selecting alternative j at market t.

$$\hat{P}_{jt}(x_t, \delta_t, d_t; \beta^o) = N_t^{-1} \sum_{i=1}^{N_t} \hat{P}_{ijt}(x_t, \delta_t, \overline{d_t}, D_i; \beta^o). \tag{6}$$

3.2 Estimation

In our framework we follow the demand estimation problem using both market-level and individual-level data (Berry et al., 2004). Equation (5) implies that the model could be estimated using a conditional logit with alternative-specific fixed effects δ_{jt}. This model would yield consistent estimators of (δ_t, β^o) without resorting to additional restrictions on ξ_{jt}. Indeed, estimating equation (3) using a conditional logit specification is enough to recover individual choice probabilities, as well as market share predictions $\hat{P}_{jt}$.
However, in the above setting it is not possible to identify parameters (γ, β^m). Since δ_{jt} is a function of school-track $j's$ characteristics, decomposing δ_{jt} in observed attributes is not enough to capture consistent estimators of γ and β^m, unless there is no correlation between those attributes and ξ_{jt}. Namely, while estimating γ and β^m we face the classic endogeneity problem in econometrics. This issue is especially challenging when we consider the potential relationship between ξ_{jt} and $\overline{d_{jt}}$. Unobserved school-track quality, strategic school location, and residential sorting, among other factors, may influence the average distance pupils need to travel, $\left(\overline{d_{jt}}\right)$. As a consequence, any estimation of these structural parameters, as well as individual elasticities and WTT, will be inconsistent.
In order to identify all coefficients from the original model, we need to resort to market level data. By recovering mean utility values alone we can use aggregated information on school-track's attributes to consistently estimate γ and β^m. To perform this estimation process, we rely on an influential result from the empirical industrial organization literature, which states that there exists a contraction mapping between the observed market shares, and the market shares implied by the model, on δ_{jt} (Berry et al., 1995). Namely, given a set of parameters β^o, there is a unique vector of mean valuations per alternative $\delta = (\delta_1, ..., \delta_T)$ such that the observed market shares and the ones the model predicts are (asymptotically) equivalent. Once parameters (δ, β^o) are identified, it is possible to estimate (γ, β^m) as a function of school-track observed attributes by any econometric method that corrects the potential endogeneity problem stated above. The particularities of this

estimation process are explained in the next subsection.

3.3 The BLP Algorithm

In this setting, we need two different databases. On the one hand, we construct a market-level database, which includes all relevant market information, market shares per alternative, and observable characteristics of all options available within each market. In this database the unit of observation is a market-school-track combination. On the other hand, we build an individual-level database, with information about choice sets, school-track alternatives, demographics, and academic attributes per pupil. In this database the unit of observation is a student-school-track combination.

Using these sources of information, the estimation algorithm can be implemented by following the next steps:[8]

Step I: Given some initial values for δ, labeled δ_0, we use the individual-level dataset to estimate a conditional logit by maximum likelihood to recover β^o parameters, as well as individual choice probabilities denoted by $\hat{P}_{ijt}$.

Step II: Using results from Step I, recover δ_t by solving for each market an implicit system of equations:

$$\hat{P}_{jt}(\delta_t, \beta^o) = s_{jt}, \quad j = 1, .., J_t$$

where s_{jt} is the observed market share for alternative j at market t in the data. Using the contraction mapping suggested by Berry et al. (1995), the system of equations for all markets has a unique solution characterized by the following equation:

$$\delta^{h+1} = \delta^h + ln\,(S) - ln\left(\widehat{P}\left(\delta^h; \beta^o\right)\right).$$

Step III: We repeat Step I and II until the distance between mean utility vectors at the last two iterations is lower enough. That is, when $||\delta^H - \delta^{H-1}|| < \varepsilon$, with ε denoting the convergence criterion implemented, and H being the minimum number of iterations to reach convergence. After recovering mean utility values, it is possible to invert the system of equations to express the vector δ as a function of observed market shares:

$$\delta^*(S, \beta^{o*}) = \widehat{P}^{-1}\left(S, \beta^{o*}\right),$$

where $(*)$ denotes the parameters when they have been identified.

Step IV: We use results from Step III to define the structural error term as $\xi_{jt}(\delta^*_{jt}, \gamma, \beta^m) \equiv \delta^*_{jt}(s_{jt}, \beta^{o*}) - \left[\gamma x_{jt} - \beta^m \overline{d_{jt}}\right]$. Then for a set of $M + k + 1$ instruments (where M is the number of excluded instruments) $Z = \{x, d, z_1, ..., z_M\}$, we compute (γ, β^m) by using an estimator that will fit the set of moments:

$$E\left[Z'\xi(\delta^*, \gamma^*, \beta^{m*})\right] = 0.$$

This can be done either by using a minimum distance estimator, such as Generalized

[8] See Train (2009) for a detailed exposition of the BLP algorithm.

Method of Moments (GMM), or by using the instrumental variables estimator (IV). In summary, the whole estimation algorithm can be described in two simple steps: i) Using individual level data to recover unobserved mean utility levels, and ii) using mean utility levels as a dependent variable in a regression on school-track's features to recover mean preferences for the different academic programs available.
Since we are concerned to recover mean valuations of non-vocational academic tracks, we modify equation (4) as follows:

$$\delta_{jt} = \gamma_h High_{jt} + \gamma_l Low_{jt} + \alpha x_{jt} - \beta^m \overline{d_{jt}} + \phi_s + \psi_m + \xi_{jt}, \tag{7}$$

Where parameters γ_h and γ_l capture mean preferences for the high track, and the low track, respectively. Parameter α measures average valuations towards other attributes than track denomination. Finally, ϕ_s (ψ_m) are fixed effects at the school management board (municipal) level. The parameters of interest in this estimation stage are $(\gamma_h, \gamma_l, \beta^m)$. Since the large majority of schools offer both tracks, it is fair to consider the availability of such tracks as exogenous. This fact implies that the challenge of our identification strategy relies mostly on computing consistent estimates of β^m. In the next subsection, we discuss our instruments and identification assumptions.

3.4 Identification Strategy

The main objective of this estimation process is to recover consistent estimates of mean utility values (δ_{jt}) and mean preferences for commuting (β^m). Regarding estimation of δ_{jt}, identification is guaranteed since no unobservables at the individual level affect the consistency of parameters which measure deviations from mean preferences. Likewise, our first identification requirement is that individual confounding factors have no role in explaining heterogeneity on preferences for commuting. Under this assumption, mean utilities are identified because of the BLP algorithm implementation. Namely, for a subset of individual interactions, there is a unique vector of mean utilities that equalize the observed market shares with the market shares predicted by the model. As pointed out by Nevo (2000), we only require variation on the choice sets per market, as well as variation of demographic attributes (at least two different individuals) within each market to recover consistent estimates of δ_{jt}.

Identification of β^m is less straightforward. In an ideal experimental setting, households would have been randomly allocated across different residential locations to later record their school-track choices. Residential sorting motivated by neighborhood composition, and strategic location of school facilities implies that the average distance traveled to each alternative is correlated with unobserved elements that influence school choice. To deal with this issue we implement an instrumental variable approach (IV), using two plausibly exogenous sources of variation. The first instrument is the distance from each school to the closest city hall, as a way to capture its degree of market accessibility(Noailly et al., 2012). We argue that this variable is positively correlated with the average commuting distance to each school-track available, but not correlated with unobserved attributes that influence the final schooling decision. Similar instruments have been used in the school choice literature, an example being the distance to the nearest hospital, as a way to control for the endogeneity of traveling costs (Carneiro et al., 2016).

The second set of instruments we implement is motivated by the spatial differentiation of retail markets literature (Davis, 2006; Houde, 2012). We define school-track $j's$ influential area as a concentric circular area with four kilometers radius. Then, we construct concentric rings to count the amount of rival schools located in those rings, just in the same way a dart board is drawn. Hence, we define $rivals(a, b)$ as the number of competing schools located in a concentric ring between a and b kilometers *away* from school-track j. Since these concentric areas are arbitrarily generated, school-track locations in such rings are plausibly exogenous. Thus, our identification assumption is that the larger the number of competing schools concentrated *outside* school-track $j's$ influential area, the lower the average distance commuted to such alternative since pupils living there would prefer to attend education at a nearby school. In that sense, this specific set of instruments are negatively correlated with the average cycling distance to attend education, and not necessarily correlated with other cofounders that influence school choice.

3.5 Recovering Willingness to Travel (WTT) Estimates

It is possible to rearrange equation (7) to express average distance per alternative as a function of school-track characteristics (we ignore fixed effects to ease the exposition):

$$\overline{d_{jt}} = \tilde{\gamma}_h High_{jt} + \tilde{\gamma}_l Low_{jt} + \tilde{\alpha} x_{jt} + \tilde{\xi_{jt}} - \tilde{\delta_{jt}},$$

where all parameters with the ($\sim$) denote that they are divided by the distance coefficient β^m. As Bayer et al. (2007) point out, this equation is nothing more than a corrected hedonic model specification. In such a case, all parameters can be interpreted as the WTT in order to attend (or avoid) any school-track with a particular feature. Thus, from equation (7) is it possible to recover such estimates (and their standard errors using the delta method) by calculating the ratio of each parameter associated to each attribute with respect to the distance coefficient. Naive estimations using OLS-hedonic regressions assume that $\tilde{\delta_{jt}}$ is equal to zero, neglecting any endogeneity problem in the identification strategy.

4 Data Description

In this section we first describe our main sources of information, how these sources were combined to consolidate our final datasets, and the main features of the particular educational system considered for this study. Then, we continue by offering relevant definitions about students, markets, and academic alternatives, as well as summary statistics of our main variables of interest.

4.1 Data Overview

In this study we use a unique dataset on the educational development of elementary and secondary school children who inhabit a statistically representative province of the Netherlands (South-Limburg). These data are collected in a cooperative project between elementary schools, secondary schools, managerial school boards, local authorities, and Maastricht University, in order to assess school performance and track placement, as well as to promote economic and educational development. This project surveys 7 cohorts of students in their final year of elementary school, for the years 2007 through 2013. Hence, the complete dataset comprises a repeated cross-section of 39,794 pupils, roughly 5,600 per cohort, representing up to 95% of all primary school coverage in the region.

Since we are concerned to model parental choices at the secondary school level, we use a particular subsample of students (cohort 2009) who were also surveyed in their third year of secondary education. We are able to identify information about household socio-demographic characteristics, prior pupils' academic performance at elementary school, and final tracking placements from 4,568 students, covering almost 90% of all secondary schools in the province where the cooperation project is developed. We decided to discard pupils' information from four municipalities with only one academic track observed (16 students), as well as all students whose main residence is outside the Netherlands (53 students). With respect to pupils' characteristics and previous academic performance, 464 children were dropped from our sample since information about their standardized test scores and primary school advice is missing. In addition, we cannot identify the gender of 136 students neither from the administrative data, nor from the two surveys conducted. This data cleaning process leaves us with 3,988 usable observations to perform estimations at the individual level.[9]

4.2 Markets

We define a market as a municipality × household's socio-economic background combination. We measure the latter by recording the highest educational level the parents achieved. A household is labeled as "high-educated" if at least one of the parents reported to have completed post-secondary education (regardless of the track they followed), and is tagged as "low-educated" otherwise. With this information we classify students and

[9] However, it is important to point out that, with exception with students living outside the Netherlands and in those municipalities where all students attended the same track, the remaining students were considered in order to calculate the observed market shares of all school-tracks per market.

the school-tracks they selected in 40 different markets (20 municipalities $\times$ 2 education categories), bringing a maximum of 8 alternatives per market, for a total of 286 different alternatives (excluding the outside option).

By using this market definition, we are accounting for geographic and demographic variation that determines the choice set a student faces when registration at the secondary school is required. While the geographic differences might seem obvious (the larger the population within a municipality, the larger its choice set), we claim that parental educational attainments are also essential to determine the availability of options per market that geographic variation cannot capture alone. Table 1 provides evidence about this fact, by reporting the number of alternatives per market. Special attention should be addressed to the variation in the composition of the choice sets. Essentially, markets including students from highly educated families have choice sets with more chances to enroll into the high track. In contrast, we observe that markets where pupils from low-educated families participate offer more alternatives to attend the low track, with few exceptions. In Appendix Table A.1.1 we provide calculations of aggregated market shares per educational track. On average, pupils from high-educated households attend the high track 27% more than students with low-educated parents. Almost half of the latter attend vocational education, contrasting with only 25% of the children from high-educated families choosing that option. Overall, students from low-educated families have lower rates of enrollment into pre-university tracks, a fact that is robust for almost all municipalities in our sample.

4.3 Students

From our survey data we collected information at the pupil and household levels. In particular, we were able to identify the maximum educational level achieved by the household, the student's gender, as well as measurements of pupil previous performance at elementary school. The first performance measure is the CITO test score, which is measured in a 500-550 range with increments of 1 point. The discreteness of this test allows us to compute 6 groups of percentiles to measure differentials in educational outcomes across students. The second performance measure is the qualitative and, to some extent, subjective advice given by primary schools about the most suitable track the pupil should follow afterwards. We create dummy variables to account for these different recommendations, emphasizing the role of three particular outcomes: receiving a positive advice to attend the high track, receiving an advice to follow the low track, and receiving a mixed (inconclusive) advice, suggesting that the student is suited to attend any of the non-vocational tracks available. Finally, we record each student's gender and identify the number of peers from primary school that attended each alternative, as a way to account for potential spillover effects in final schooling choices.

Table 2 reports pupils attributes, making the distinction between the full sample (column (1)), and the groups of students coming from low and high-educated households, respectively (columns (2) and (3)). In column (4) we report the differences in means between both types of pupils. Children from educated parents are on average 7.5 percentage points more likely to obtain results in the upper 10th percentile of the CITO standardized test scores distribution. Moreover, students from low-educated families are 20 percentage points less likely to receive a positive advice for enrolling in the high track. Regarding mixed advices, understood as the inconclusive recommendation from the elementary school suggesting that the pupil can attend both tracks, children from high-educated families received on average 7.6 percentage points more recommendations than pupils from low-educated backgrounds. Not surprisingly, students with low educated parents are 3.5 percentage points more likely to receive an advice to attend the vocational track. Virtually no difference is found in the gender composition among both types of households. On average, students commute 4.4 kilometers to attend secondary education, but students from highly educated families travel slightly longer distances (300 meters).

4.4 Alternatives

In addition to our individual dataset, we merged student information with data on school-track characteristics using administrative records from the Dutch Inspectorate of Education (in Dutch: DUO). We recorded information on school denomination and whether the school offered an alternative pedagogical method, the type of track offered, the managerial board in charge of each school, and two measures of observed quality. The first measure is the standardized high-school exit test scores that students took at the end of their studies. The second measure is the proportion of pupils that fail to graduate in time per track selected (6 years for the high track, 5 years for the low track). Regarding traveling costs, we use information on the location of pupils and schools at the 6-digit postcode level precision and construct cycling distance measurements to each alternative-household combination.[10] With this information, we compute the average cycling distance per alternative available in each pupil's choice set.

Table 3 shows the distribution of students per school-level attributes and track attended (excluding the outside option). Catholic schools represent 70 percent of total schools' supply. Schools with an alternative teaching methodology account for 15 percent of all schools. While there are virtually no differences in the fail rate among schools offering both tracks (8% on average), students in the lower track score on average 0.15 standard deviations higher in the central exit exam. Regardless of the track, these exams are

[10] To measure cycling distances, we recover information on postcodes of schools and pupils at the six digit level from our survey dataset. Next, we calculated the distance between these postcodes using the STATA module *traveltime3,* which gathers distance measures and travel time by mode of transport using Google Maps geocoded information (Ozimek and Miles, 2011).

graded in the same zero-to-ten scale, suggesting that pupils enrolled in the low track are scoring better. However, since the exam content is not necessarily the same for each track, another much more plausible explanation is that students enrolled in the high track might face a more demanding test. Regarding commuting costs, on average students travel seven kilometers to attend non-vocational tracks, while travelling slightly larger distances for to attending the lower track (500 meters).

Table 4 reports the distribution of school-track attributes per household educational background. Columns (1) and (2) report means and standard deviations of all characteristics considered, discriminating by families with low and high education level, respectively. Columns (3) and (4) show regressions on observed market shares as a function of school-track characteristics for both groups of pupils. Students from highly educated families are two times more likely to be enrolled in the high track, relative to students from low educated backgrounds. In addition, no difference is found in enrollment rates for the low track between the two groups of students considered. Relative to pupils with low-educated parents, students from a higher economic background fail more often to graduate in time, but also obtain better results in the central standardized test scores after finishing their studies. Furthermore, they travel 6.6 kilometers on average (roughly one kilometer more than pupils from low-educated families) to attend education at secondary schools. Regarding results from columns (3) and (4), a large majority of school-track attributes is negatively correlated with the observed market shares per type of parents. However, the negative correlation is stronger among pupils from low-educated families. Remarkably, 70% of the variation in market shares can be explained by the entire set of independent variables considered. That is, all attributes considered are powerful enough to explain variation in the market shares per type of household.

Appendix Table A.1.2 presents summary statistics of school-track attributes by students' characteristics. Considering CITO test scores, results suggest that there is a positive sorting on track enrollment, where pupils with better (worse) scores are more likely to register into the high (low) track. The same seems to hold for pupils receiving a positive advice to enroll in the high track. Interestingly, a very low proportion of students receiving a mixed advice enrolls in non-catholic schools. Failing rates among female students are 3 percentage points lower than the average for the entire set of pupils in our sample. Considering high-school exit exam scores, larger scores are positively associated with better results in both objective and subjective performance measurements at the primary school level. Finally, considering commuted distances, at least 36% of students cycle more than four kilometers to reach their preferred schools. Overall, these findings support the idea that individual heterogeneity plays a key role in determining preferences for school-track placements.

5 Results

5.1 Recovering Mean Utility Values

Table 5 provides estimates of the full parental choice model in order to compare them with traditional, conditional logit demand specifications. Column (1) reports a conditional logit with no individual heterogeneity. Column (2) displays results of the same model incorporating observed individual heterogeneity. Finally, Column (3) shows results from our suggested specification, while implementing the BLP contraction algorithm to recover fixed effects per alternative that can be interpreted as mean utility levels. Results from our preferred model (3) suggest that the set of attributes considered are positively correlated with the average distance pupils have to cycle in order to attend secondary education. The correlation is stronger in size the better the pupil's performance at primary school is. Thus, students with CITO scores above the mean have significantly higher correlations than pupils with low scores. Regarding primary school tracking recommendations, the less optimistic the advice the higher the correlation with traveling costs. Other attributes, such as the number of peers and gender, are statistically significant, but their correlations are relatively low, suggesting that heterogeneity in traveling preferences are mostly driven by differences in prior academic performance across students.

When we consider all specifications, some inconsistent results emerge. In particular, estimating a conditional logit model without interaction terms implies recovering a *negative* mean willingness to travel for attributes that commonly are valued positively by parents (10 kilometers on average for both tracks). Moreover, this anomaly is not solved by including interaction terms to account for student heterogeneity (Column (2)). Overall, these results point out the empirical problem faced in this setting: While gaining simplicity and tractability, demand estimation by traditional methods will lead to inconsistent estimates of parental preferences for school-track attributes. In contrast, the method we adopt here aggregates all observed and unobserved variation that is constant across individuals, in order to obtain reliable estimates of substitution patterns that depend on students' characteristics.

To better illustrate the advantages of our model, Table 6 presents estimations of student enrollment probabilities for the high track. We estimate these probabilities for all specifications considered in Table 5. Next, we estimate OLS regressions to account for differences in the mean probability of enrollment that can be explained by changes in students' features. Hence, columns (1)-(3) correspond to the enrollment probabilities derived from a conditional logit without individual heterogeneity. Columns (4)-(6) relax that restriction by incorporing the enrollment probabilities from a model that interacts school-track attributes with pupils' characteristics. Finally, columns (7)-(9) display the results when the dependent variable is the enrollment probability estimated using the

BLP algorithm. In these regressions, the coefficient of interest is the one reported in the second row, which accounts for the enrollment gap between pupils from high and low educated households. A positive value indicates that the enrollment gap favours students with high educated parents.

The findings reported in columns (1)-(6) show that, on average, children have a 20% chance to be enrolled into the high track. However, after controlling for parental educational background, pupils from high-educated families have a higher likelihood to attend that option. The difference in the size of this effect depends on the specification considered. For traditional conditional logit models the difference is rather small (between 2% and 3% after controlling for pupils' covariates). In contrast, the conditional logit model under the BLP contraction algorithm, students with low educated parents are between 20% and 24% less likely to attend the high track relative to students from better family backgrounds. This last result is robust either when other students' characteristics are included (columns (7) and (9)).

In general, the key aspect to highlight from these findings is that traditional individual-choice models cannot outperform the BLP approach. The contraction algorithm implemented allows us to recover a unique vector of mean utility values that i) fits the best the distribution of students' attributes per school-track selected and ii) equalizes the market shares observed in the data with the aggregated market shares computed using the individual enrollment probabilities the model predicts. Figure 1 portrays graphic evidence for this claim, by showing aggregated market shares from all models against the aggregated market shares from the data. Thus, the closer the predicted shares to the 45-degree line, the better the prediction from the model. As argued, traditional specifications predict market shares relatively poorly, especially for larger values. Contrarily, implementing the BLP contraction algorithm provides almost full convergence, as long as a reasonable convergence criterion is set.[11]

5.2 Instrumental Variable Selection And Discussion of Potential Violations

As we are interested in consistently recovering the structural parameter β^m to calculate willingness to travel valuations, the instrumental variable that we should implement needs to satisfy two requirements. First, the instrumental variable has to be relevant, in a sense that it is a strong predictor of the average commuting distance to each school-track in the data. Second, the instrumental variable should be exogenous. In particular, we expect the instrument not to be correlated with the academic track supply, as we are concerned to recover preferences over non-vocational denominations.

[11] According to the literature on demand estimation, a reasonable convergence criterion is to set $\varepsilon = 10^{-16}$. This was the convergence criterion we selected to implement the BLP contraction algorithm.

Relevance: In Table 7 we report first stage regressions results using all non-vocational tracks available (286 in total). The dependent variable in each of these estimations is the cycling distance (in kilometers) students, on average, travel to attend a particular school-track. Each regression corresponds to a different instrument and includes all remaining school-track features, plus a set of school board and municipality fixed effects. We use as plausible instruments several combinations of the "concentric rings" rival schools variables to evaluate their relative strength, as well as the distance from each school-track to the closest city hall. Finally, goodness of fit diagnostics and F-statistics used to test the null hypothesis are reported in columns (4) and (5).

Two conclusions can be extracted from the above exercise. First, the distance of each school to the closest city hall appears to be a weak instrument, as the effect derived from the first stage is too imprecise to be considered as statistically different from zero. Second, regarding the number of rival schools in concentric proximity areas, we have at least six excluded instruments that we can consider relevant since their corresponding F-statistics are above traditional critical values (Staiger and Stock, 1997). From these, we decided to choose the three most relevant, $rivals(3,6)$, $rivals(4,6)$, and $rivals(4,8)$ to conduct further endogeneity, tests and use them in the central estimation procedure to recover the structural parameters of the mean utility equation (7).

Exogeneity: While a direct test of the exogeneity assumption is not available, it is possible to argue that each school's location in the artificially-build concentric rings is uncorrelated with mean preferences for track denomination. To further develop this idea, we implement the procedure followed by Altonji et al. (2005) to show that, conditional to non-track related school attributes, the instruments selected from the relevance analysis are uncorrelated with the academic track supply. This test works in a two step estimation process: First, we regress each instrument on school attributes unrelated to the track denomination, plus a set of fixed effects at the school board and municipal levels. Then, the resulting residuals are regressed on the set of dummy variables that indicate whether a particular alternative corresponds to a higher or a lower track. Not being able to reject the joint null hypothesis that the coefficients from the latter regressions are equal to zero provides evidence that the exclusion restriction for each instrument holds. Table 8 display our findings on this exogeneity test. As expected, none of the three instruments previously selected appear to be correlated with the track denomination. This fact suggest that, while school-track location might be endogenous, location in the artificially built concentric areas is plausible random.

5.3 Parental Willingness to Travel for School-Track Attributes

Table 9 reports the reduced form (columns (1)-(3)), first stage (columns (4)-(6)), instrumental variables (IV) (columns (7)-(9)), and naive OLS (column (10)) results from estimating equation 7. In the structural, reduced, and OLS forms, the dependent variable is the mean utility level of each alternative that was recovered using the BLP algorithm. In the first stage specifications, the dependent variable is the average cycling distance (measured in kilometers) that pupils have to commute to attend education at a particular school-track. Given our central research questions we include other relevant school-track attributes as controls, but we abstain to interpret any other parameters than $(\gamma_h, \gamma_l, \beta^m)$.

Regarding valuations for track denominations, coefficients for the low and high track barely differ. The coefficient associated to the average cycling distance variable seems to be overestimated under OLS. However, this bias is on the order of 10%. Regarding the instrumental variables requirements, while the reduced and first stage forms suggest that all instruments previously selected are statistically significant, in the structural form it is clear that $rivals(4, 8)$ is the strongest instrument as it has associated the highest F-statistic. Hereafter, the results from this study will be obtained using this instrument alone.

In Table 10, we provide an economic interpretation of our results by showing willingness to travel estimates for all non-vocational tracks. Column (1) and (2) report valuations that are recovered from an OLS hedonic travel model, either by using the average cycling distance, or the current cycling distance as a dependent variable, respectively. Column (3) presents the implied willingness to travel from the IV approach using mean utility levels as a dependent variable, and the number of rival schools located in a concentric ring between four and eight kilometers, $rivals(4, 8)$, as an instrument. All specifications contain fixed effects at the school board and municipality levels. These coefficients measure the distance in kilometers that a pupil is willing to cycle in order to attend (or avoid) a school-track with a particular attribute. Our results suggest that parents hold positive preferences for non-vocational tracks. On average, parents are willing to let their child commute 5.2 (4.8) kilometers to attend the higher (lower) track, which supports the common idea that academically demanding tracks are (individually) valued much more in the market. Moreover, ignoring unobserved school-track features implies an overestimation of the reserve price (measured in kilometers) students will pay to receive education by roughly fifty percent.

5.4 Heterogeneous Preferences for Track Placement

Table 11 extends the previous findings by comparing willingness to travel for low and highly educated parents. Panels A and B display valuations for the higher and lower tracks, respectively. The parameter of interest is the interaction between the parental educational attainment and the dummy variable indicating the track denomination (third and sixth rows). Positive (negative) coefficients will suggest that highly educated parents value more (less) each alternative, relative to less educated parents.

Notably, while all parents value both tracks positively, the latter are willing to let their child cycle roughly 0.8 kilometers more in order to attend the higher track, relative to parents from lower educational backgrounds. This difference in mean preferences cannot be recovered by a simple OLS-hedonic travel regressions. This fact suggests that OLS estimates present a downward bias in the reserve price of commuting for highly educated parents, as well as in the difference in valuations among parents with diverse degrees of education. The fact that the interaction coefficient in Panel B is the negative of the one obtained for the higher track allows us to conclude that there is educational sorting. Children from highly educated families will cycle to avoid attending the lower track and can travel twice the distance if that guarantees to be enrolled in the higher track.

In order to support the previous finding we also calculated the elasticity distance-demand for all parents in our sample. This elasticity reflects the percentage points the enrollment in a particular track decreases when the average distance to commute increases by 10 meters (1%). Figure 2 shows how this elasticity distributes for high and low-educated parents. Although no difference is spotted for elasticities regarding the low track, low-educated parents appear to be more sensitive to changes on average commuting costs regarding the high track. on average, we obtain elasticities of -4.46 and -3.43 for low and high-educated families, respectively. These large values reflect how, in sharp contrast with other studies in the literature of school choice, sensitivity to commuting costs is especially high in school systems where the mode of transportation is uniform, relatively easy to access and, from the supply side, schools cannot create legal barriers on enrollment other than minimum academic requirements to be fulfilled by students at the end of elementary school.

To assess the extent to which heterogeneity in preferences is sensitive to students' attributes, we report differences in mean willingness to travel between pupils from low-educated parents vs. high-educated (Table 12). Negative values indicate that, relative to parents with post-secondary education, parents from lower educational backgrounds are less inclined to let their child travel to attend a school-track far away from home. Our findings suggest that, for the high track, parents whose pupils excel at the CITO test have comparable preferences to those with higher educational attainments. By including primary school recommendations, it is clear that the heterogeneity is driven mostly by the advice pupils receive, rather than their achievements at the CITO test. Furthermore, contrary to other studies on school choice, neither primary school peer effects nor gender have a role explaining differences in preferences among parents. We conjecture that this partially reflects the fact that pupils from low-educated parents might receive more conservative recommendations than pupils from better household backgrounds.

5.5 Welfare Considerations

Using estimates from our model we can simulate a shutdown policy of all academic options concerning pre-university tracks to assess how much parents value the counterfactual system implied by the policy, relative to the original setting they based their choices from. To understand the relevance of this exercise, consider the situation where in a particular market the higher track is no longer available. For those students who originally chose a lower track, this policy intervention will have no effect. However, students that originally might have selected the high track will decide whether to travel longer distances and find a comparable alternative, or selecting a lower track available close home. As children with high educated parents are willing to travel longer distances, we might expect that the policy intervention has a welfare-decreasing effect that largely impact pupils from low educated backgrounds that could have attended the higher track otherwise.

We use a standard measure of compensated variation (CV), which consists in the amount of *additional* kilometers a student should commute that equalizes utility across both states of nature: the one implied by the shutdown policy, and the one implied by the status quo. To translate these costs in monetary terms we apply the following rule: we calculate the amount of euros that a student should be compensated using as a benchmark an average speed of 16 km/hour, and an hourly compensation of 2.7 Euros per hour implied by the minimum wage regulations in the Netherlands. Then, the number obtained is multiplied by two, assuming the student is compensated for the whole commuting costs of attending school (i.e. the round trip).

Hence, following Nevo (2000) and Train (2009), the compensated variation (in Km.) per student i is given by:

$$CV_i = \frac{ln[\sum_{j=0}^{\check{J}_t} exp\left(\check{V_{ijt}}\right)] - ln[\sum_{j=0}^{J_t} exp\left(V_{ijt}\right)]}{-\beta^m},$$

where $\left(\check{V_{it}}\right)$ denotes the utility perceived for alternative j in market t under the shutdown policy, and (V_{ijt}) is the utility level enjoyed under the original choice set. Table 13 presents this set of results. Both tracks are socially valuable since their absence implies a positive CV for all students. On average, students attending the high track have to commute an additional kilometer to attend education if such option is no longer available. If society is willing to compensate students using as a benchmark the welfare loss of the worst-off student, our estimates imply that each student should be compensated by roughly 1 Euro per day attended at school. We find similar results for the low track, implying a social value of 0.76 euros per student/day at school. As observed, our policy experiment reflects that not only parents with better educational attainments value more pre-university options, but also that society place more value on those programs.

6 Concluding Remarks

In this chapter, we have estimated a demand model to consistently recover parental preferences for academic options under an early tracking education system. The methodology implemented here allows us to identify potential educational sorting incentives mediated not only by previous academic performance, but by parental socio-economic background. Our results suggest that, conditional on students' performance, children from lower educated households tend to face small probabilities of enrollment into the tracks that have better economic prospects. Moreover, pupils with highly educated parents are willing to face larger costs (in terms to commuting) to attend academic offers that are more selective and exclusive, or to avoid programs deemed less attractive. We also present evidence on how recovering preferences is not a trivial task, by showing how OLS hedonic travel models, or traditional discrete choice demand models, neglect any educational sorting explained by non-academic determinants. Finally, our welfare analysis evidence that the educational sorting of pupils with highly educated parents is directed towards those options that society values the most. All these findings combined, show that inequalities on educational opportunities can arise, even in those cases where schooling systems are fairly designed, with almost no restrictions to choice in place.

The findings from this study motivate interesting policy implications. First, it establishes the ground to design interventions aimed to reduce the existing enrollment gap between children of comparable ability but different socio-economic background. While student-based programs may offer short term benefits in terms of school enrollment and completion (e.g. Rodriguez-Planas, 2012), our research advocates for a more comprehensive approach that includes parents and schools alike. Given the moderate success of interventions aimed to provide information and create awareness about school quality (Hastings and Weinstein (2008)), an alternative is to apply such programs on the specific subset of households from lower educational backgrounds whose children's academic performance is at the top of the ability distribution. Second, schools should discourage the track misallocation, by giving more weight to objective performance measurements that are less likely to be precisely manipulated by parents with high socio-economic status. Finally, as our welfare analysis suggests, the education system should encourage (discourage) upward (downward) track mobility.

Our results are subject to the identification assumptions we have made. In particular, the source of endogeneity we have assumed is not individually driven. Only unobserved cofactors at the school-track level are relevant to explain the variation on preferences for commuting. While we decided to avoid any ad-hoc assumptions about how students' unobserved variables distribute, we cannot ignore the fact that our estimates are not robust to such source of unobserved heterogeneity. To our best, we consider this decision as optimal since there is not a clear transmission channel where this unobserved variation, if it exists, affects the substitution patterns reflected in the data. In addition, while recovering mean utility values, we can guarantee that these substitution patterns are identified for the particular set of interactions we included in the individual choice model. Designing and implementing an empirical strategy that accounts for both, observed and unobserved variation, and how these affect final schooling decisions, is an area of future research.

It is also important to remind that this is a demand-driven analysis. While unlikely, it is plausible that some schools engage in strategic behavior to discourage students from specific socio-economic backgrounds to pursue further studies into the higher track. In our model we assumed that this discriminatory behavior, if it exists, it is internalized by the household, modifying their willingness to travel for a particular school-track. We consider as a mandatory extension of this study the inclusion of the supply side of the educational relationship, to understand the incentives at the school level, that will in turn shape the design of policies aimed to promote social mobility through education. Overall, this study brings some light to understand the extent to which parental valuations of schooling attributes shape their children's future outcomes. Incentives to engage in suboptimal schooling decisions arise in a free school-choice system and they should determine the creation of policy instruments that improve intergenerational mobility and promote equity on the access to economic opportunities.

References

Alderman, H., Orazem, P. F., and Paterno, E. M. (2001). School quality, school cost, and the public/private school choices of low-income households in pakistan. *Journal of Human Resources*, 36(2):304–326.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *Journal of Human resources*, 40(4):791–821.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Bayer, P. and McMillan, R. (2005). Choice and competition in local education markets. *NBER Working Paper*, 11802.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Black, S. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

Borghans, L., Golsteyn, B., and Zolitz, U. (2015). Parental preferences for primary school characteristics. *The BE Journal of Economic Analysis & Policy*, 15(1):85–117.

Burgess, S. and Briggs, A. (2010). School assignment, school choice and social mobility. *Economics of Education Review*, 29(4):639–649.

Cabus, S. and Cornelisz, I. (2014). On competition, student sorting, and performance gains in local education markets: The dutch secondary sector. *TIER-Maastricht University working paper.*

Carneiro, P., Das, J., and Reis, H. (2016). The value of private schools: Evidence from pakistan. *IZA Discussion Paper*, 9960.

Chumacero, R., Gomez, D., and Paredes, R. (2011). I would walk 500 miles (if it paid): Vouchers and school choice in chile. *Economics of Education Review*, 30(5):1103–1114.

Cullen, J., Jacob, B., and Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, 74(5):1191–1230.

Davis, P. (2006). Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics*, 37(4):964–982.

Figlio, D. and Hart, C. (2014). Competitive effects of means-tested school vouchers. *American Economic Journal: Applied Economics*, 6(1):133–156.

Hastings, J., Kane, T. J., and Staiger, D. (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER Working Paper*, 2145.

Hastings, J. and Weinstein, J. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics*, 123(4):1373–1414.

Houde, J. F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *The American Economic Review*, 102(5):2147–2182.

Hoxby, C. M. (2003). School choice and school competition: Evidence from the united states. *Swedish Economic Policy Review*, 10:9–65.

Hsieh, C. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of Public Economics*, 90(8):1477–1503.

Koning, P. and Van der Wiel, K. (2013). Ranking the schools: How school-quality information affects school choice in the netherlands. *Journal of the European Economic Association*, 11(2):466–493.

Korthals, R. (2015). The pre-tracking effects on parental background. *ROA Research Memorandum*, 5.

Manski, C. F. (1992). Educational choice (vouchers) and social mobility. *Economics of Education Review*, 11(4):351–369.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328.

McMillan, R. (2005). Competition, incentives, and public school productivity. *Journal of Public Economics*, 89(5):1133–1154.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy*, 9(4):513–548.

Noailly, J., Vujić, S., and Aouragh, A. (2012). The effects of competition on the quality of primary schools in the Netherlands. *Environment and Planning A*, 44(9):2153–2170.

Ozimek, A. and Miles, D. (2011). Stata utilities for geocoding and generating travel time and travel distance information. *Stata Journal*, 11(1):106.

Rincke, J. (2006). Competition in the public school sector: Evidence on strategic interaction among us school districts. *Journal of Urban Economics*, 59(3):352–369.

Rodriguez-Planas, N. (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the united states. *American Economic Journal: Applied Economics*, 4(4):121–39.

Ruijs, N. and Oosterbeek, H. (2014). School choice in amsterdam: Which schools are chosen when school choice is free? *TIER Working Paper*.

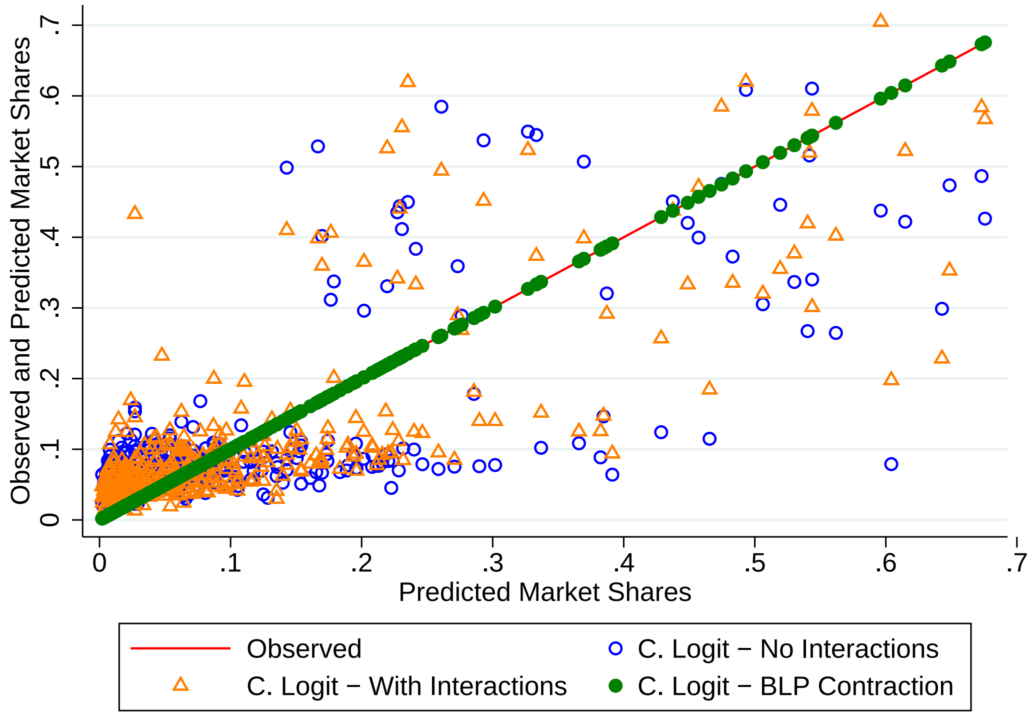
Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557.

Train, K. (2009). *Discrete choice models with simulation*. Cambridge university Press.

Van Goeverden, C. and De Boer, E. (2013). School travel behaviour in the netherlands and flanders. *Transport Policy*, 26:73–84.

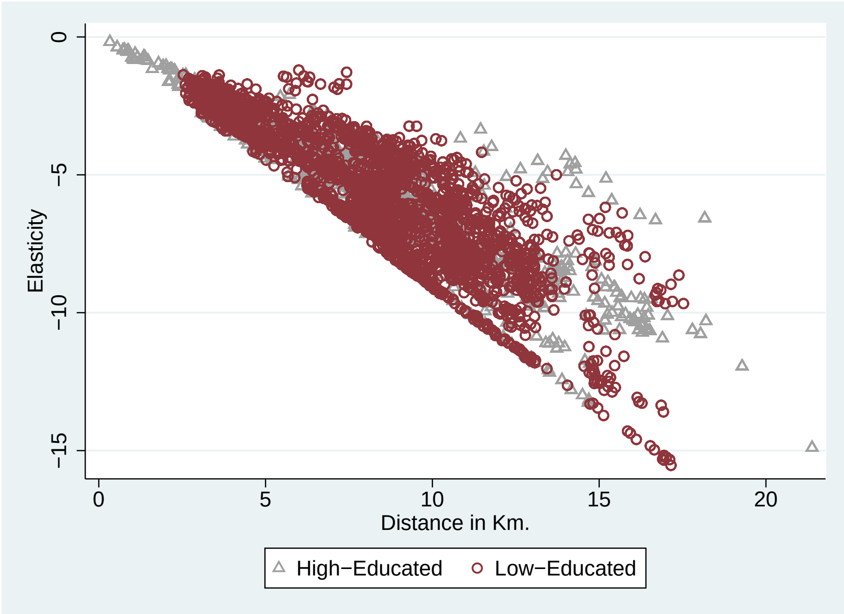
Tables and Figures

Figure 1: Observed Vs. Predicted Market Shares - All Specifications

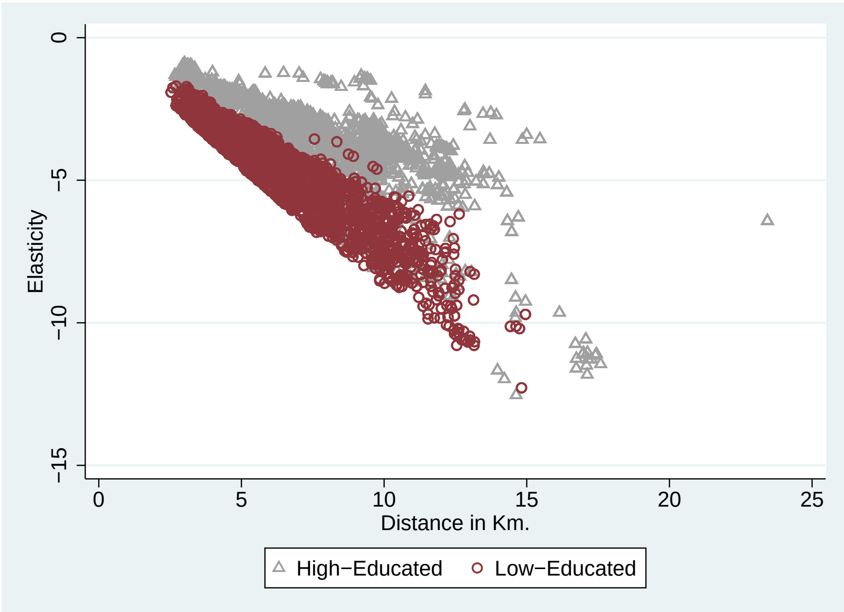


Notes: This figure reports fitted values of aggregated market shares (all tracks) vs. observed market shares. The 45-degree line depicts (sorted) observed values.

Figure 2: Elasticity of Enrollment respect to Cycling Distance



(a) Low Track



(b) High Track

Notes: These figures represent the demand elasticity with respect the average distance pupils need to commute to attend the low track (a), and the high track (b). This elasticity reflects the percentage points decrease in the enrollment rate of a particular track when the average distance to cycle increases by 10 meters (1%).

Table 1: Summary Statistics - Options per Academic Market

Municipality	low-educated parents Low Track	low-educated parents High Track	low-educated parents Total	high-educated parents Low Track	high-educated parents High Track	high-educated parents Total
1	2	2	4	2	3	5
2	3	3	6	2	3	5
3	1	3	4	1	3	4
4	4	3	7	5	4	9
5	2	2	4	2	3	5
6	5	4	9	5	7	12
7	6	4	10	5	6	11
8	3	5	8	3	3	6
9	3	3	6	1	3	4
10	4	4	8	5	5	10
11	4	6	10	4	4	8
12	6	5	11	3	5	8
13	4	4	8	3	4	7
14	4	4	8	3	4	7
15	3	4	7	3	4	7
16	2	3	5	3	4	7
17	3	4	7	2	3	5
18	2	2	4	2	3	5
19	6	5	11	5	5	10
20	2	3	5	5	4	9
Observations	69	73	142	64	80	144
Mean	4.04	3.96	4	3.78	4.30	4.07
Std. Dev.	1.43	1.06	1.25	1.28	1.23	1.27
Min	1	2	1	1	3	1
Max	6	6	6	5	7	7

Notes: Each column reports the number of school-tracks available per market, and the total number of alternatives per municipality. Each market is a municipality-Parental educational attainment combination, for a total of 40 markets.

Table 2: Summary Statistics - Pupils Characteristics

	(1) All Students	(2) Low Educated	(3) High Educated	(4) Difference (2)-(3)
CITO Score $> P_{90}$	0.077	0.049	0.124	-0.075***
	(0.267)	(0.217)	(0.330)	0.009
CITO Score $(P_{75}, P_{90}]$	0.157	0.117	0.226	-0.109***
	(0.364)	(0.321)	(0.418)	0.012
CITO Score $(P_{50}, P_{75}]$	0.232	0.204	0.278	-0.075***
	(0.422)	(0.403)	(0.448)	0.014
CITO Score $(P_{25}, P_{50}]$	0.266	0.283	0.236	0.047***
	(0.442)	(0.451)	(0.425)	0.014
CITO Score $(P_{10}, P_{25}]$	0.164	0.209	0.091	0.117***
	(0.371)	(0.406)	(0.288)	0.012
CITO Score $\leq P_{10}$	0.103	0.139	0.044	0.095***
	(0.304)	(0.346)	(0.206)	0.010
1[Advice=High Track]	0.220	0.143	0.348	-0.205***
	(0.414)	(0.350)	(0.477)	0.013
1[Advice=Mixed]	0.142	0.113	0.189	-0.076***
	(0.349)	(0.317)	(0.392)	0.011
1[Advice=Low Track]	0.153	0.148	0.163	-0.016
	(0.360)	(0.355)	(0.370)	0.012
1[Advice=Vocational Track]	0.114	0.128	0.092	0.035***
	(0.318)	(0.334)	(0.290)	0.010
1[Woman]	0.514	0.519	0.505	0.013
	(0.500)	(0.500)	(0.500)	0.016
#Peers (Primary School)	5.968	6.965	4.304	2.661***
	(6.017)	(6.280)	(5.136)	0.192
Cycling distance to	4.372	4.261	4.555	-0.294***
selected option (Km.)	(2.974)	(2.874)	(3.126)	0.097
Observations	3988	2494	1494	3988

Notes: This table reports in columns (1)-(4) means and standard deviations of student's characteristics, classified by parental educational background. Column (1) reports descriptive statistics for the full sample. Column (2) reports results for those students whose parents' highest educational achievement is high-school or below. Column (3) reports results for students whose parents' highest educational level is above high-school. Column (4) presents differences in means, whith standard errors reported in brackets. Standard deviations of all variables are reported in parentheses. * p-value $<$ 0.1 ** p-value $<$ 0.05 *** p-value $<$ 0.01.

Table 3: Summary Statistics - School's Attributes by Academic Track

Variables	Low Track (1)	High Track (2)	Difference (1)-(2)
1[Catholic School]	0.732	0.752	-0.020
	(0.443)	(0.432)	[0.018]
1[Public School]	0.119	0.118	0.000
	(0.324)	(0.323)	[0.013]
1[Alternative School]	0.149	0.130	0.019
	(0.356)	(0.337)	[0.014]
Fail Rate (10%)	0.857	0.819	0.039
	(0.401)	(0.766)	[0.026]
Central Exam Score	0.189	0.035	0.154***
	(0.745)	(1.181)	[0.042]
Average Distance (Km.)	7.440	6.946	0.494***
	(2.527)	(1.585)	[0.086]
Observations	1052	1276	2328

Notes: This table reports means and standard deviations of school's characteristics, classified by the type of track the student is enrolled. Column (1) reports results for those pupils enrolled in the Low track (HAVO). Column (2) reports results for those pupils enrolled in the high track (VWO). The outside option is to attend the vocational track (VMBO), therefore students who selected that option are excluded from this table. Column (3) presents differences in means, with standard errors reported in brackets. Standard deviations reported in parenthesis. * p-value $<$ 0.1 ** p-value $<$ 0.05 *** p-value $<$ 0.01.

Table 4: Summary Statistics - School-Track Attributes by Parental Educational Background

	Summary Statistics		Regressions on Market Shares	
	Low Educated (1)	High Educated (2)	Low Educated (3)	High Educated (4)
1[High Track]	0.213	0.499	-0.4967***	0.0472***
	(0.409)	(0.500)	(0.0171)	(0.0134)
1[Low Track]	0.257	0.275	-0.5165***	-0.0598***
	(0.437)	(0.447)	(0.0217)	(0.0129)
1[Public School]	0.0473	0.106	-0.0548***	-0.0902***
	(0.212)	(0.308)	(0.0147)	(0.0073)
1[Alternative School]	0.0666	0.105	-0.1956***	-0.0831***
	(0.249)	(0.307)	(0.0207)	(0.0139)
Fail Rate	0.413	0.613	-0.0233	-0.0427***
	(0.609)	(0.653)	(0.0144)	(0.0087)
Central Exam Score	0.0151	0.138	-0.0349***	-0.0121***
	(0.659)	(0.929)	(0.0093)	(0.0041)
Average	5.570	6.621	0.0936***	0.0321***
Distance (Km.)	(2.467)	(2.278)	(0.0013)	(0.0013)
R-Squared	-	-	0.7047	0.6909
Observations	2,494	1,494	2,494	1,494

Notes: This table reports summary statistics of school-track characteristics, classified by parental educational background. Columns (1) and (2) report means and standard deviations for pupils with parents whose educational level is high-school or below (low educated), and parents whose educational attainments are above high-school (high educated), respectively. Columns (3) and (4) report regressions where the dependent variable is the observed market shares. Column (3) reports regressions results for low educated parents. Column (4) reports regression results for high educated parents. Standard errors reported in parentheses. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 5: Conditional Logit Estimations

	No Interactions (1)	With Interactions (2)	BLP Algorithm (3)
Average Distance ×			
CITO Score $[P_{10}, P_{25})$		0.315***	0.095**
		(0.065)	(0.038)
CITO Score $[P_{25}, P_{50})$		0.686***	0.484***
		(0.059)	(0.027)
CITO Score $[P_{50}, P_{75})$		0.853***	0.686***
		(0.059)	(0.025)
CITO Score $[P_{75}, P_{90}]$		0.992***	0.864***
		(0.058)	(0.019)
CITO Score $> P_{90}$		0.999***	0.839***
		(0.058)	(0.022)
1[advice: mixed]		0.169***	0.198***
		(0.020)	(0.024)
1[advice: low track]		0.253***	0.339***
		(0.021)	(0.027)
1[advice: vocational track]		0.190***	0.223***
		(0.036)	(0.039)
#Peers primary school		0.024***	0.011***
		(0.001)	(0.001)
1[Female]		0.053***	0.057***
		(0.014)	(0.017)
1[High Track]	-1.284***	0.092	
	(0.073)	(0.094)	
1[Low Track]	-1.343***	0.004	
	(0.086)	(0.104)	
1[Public School]	-0.260***	-0.233***	
	(0.070)	(0.072)	
1[Alternative School]	0.049	-0.081	
	(0.066)	(0.066)	
Fail Rate (10%)	-0.044	-0.095*	
	(0.057)	(0.057)	
Central Exam Score	0.041	0.039	
	(0.036)	(0.035)	
Average Distance	-0.127***	-1.108***	
	(0.008)	(0.060)	
N	38,373	38,373	38,373
Log-Likelihood	-7549	-6688	-5809

Notes: This table presents coefficients from different conditional logit estimations. The dependent variable in all specifications is a dummy variable that takes value one if a particular school-track was selected and zero for the remaining alternatives. Column (1) reports a conditional logit specification with no individual heterogeneity. Column (2) displays results of the same model incorporing observed individual heterogeneity. Column (3) shows results from our prefered specification, in which we implement the BLP contraction algorithm to recover fixed effects per alternative that can be interpreted as mean utility levels. The outside option consists in attending a vocational track. The categories excluded are the bottom 10th percentile on the CITO exam scores, the primary school advice indicating the student should follow the high track, and whether the academic track is offered by a roman-catholic school. Robust standard errors reported in parentheses. p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 6: High Track's Enrollment Probabilities - Differences by Students' Attributes

	No Individual Heterogeneity			Individual Heterogeneity			BLP		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.308***	0.296***	0.224***	0.283***	0.213***	0.163***	0.201***	0.099***	0.126***
	(0.001)	(0.004)	(0.004)	(0.003)	(0.010)	(0.009)	(0.002)	(0.006)	(0.007)
1[HIgh Educated]	0.033***	0.031***	0.027***	0.099***	0.048***	0.019***	0.283***	0.238***	0.198***
	(0.002)	(0.002)	(0.003)	(0.005)	(0.004)	(0.007)	(0.004)	(0.003)	(0.006)
CITO Score > P_{90}		0.021***	0.004**		0.183***	0.206***		0.201***	0.167***
		(0.004)	(0.002)		(0.011)	(0.009)		(0.008)	(0.006)
CITO Score $(P_{75}, P_{90}]$		0.013***	0.004**		0.171***	0.204***		0.177***	0.162***
		(0.004)	(0.002)		(0.010)	(0.008)		(0.006)	(0.006)
CITO Score $(P_{50}, P_{75}]$		0.011***	0.004**		0.147***	0.184***		0.145***	0.143***
		(0.004)	(0.002)		(0.009)	(0.007)		(0.006)	(0.005)
CITO Score $(P_{25}, P_{50}]$		0.010***	0.004**		0.124***	0.146***		0.114***	0.113***
		(0.004)	(0.002)		(0.009)	(0.006)		(0.006)	(0.004)
CITO Score $(P_{10}, P_{25}]$		0.005	0.002		0.047***	0.043***		0.021***	0.016***
		(0.004)	(0.002)		(0.009)	(0.006)		(0.006)	(0.004)
1[Advice=Mixed]		0.002	-0.002**		0.017***	0.031***		0.036***	0.030***
		(0.003)	(0.001)		(0.006)	(0.005)		(0.005)	(0.003)
1[Advice=Low Track]		0.006**	-0.001		0.043***	0.049***		0.048***	0.043***
		(0.003)	(0.001)		(0.005)	(0.005)		(0.004)	(0.004)
1[Advice=Vocational Track]		-0.005*	-0.001		0.034***	0.050***		0.039***	0.048***
		(0.003)	(0.001)		(0.006)	(0.005)		(0.005)	(0.004)
#Peers (Primary School)		0.000*	-0.000*		-0.007***	-0.010***		-0.002***	-0.004***
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
1[Female]		-0.001	0.001		0.006*	0.007***		0.009***	0.005**
		(0.002)	(0.001)		(0.004)	(0.003)		(0.003)	(0.002)
Observations	3,988	3,988	3,988	3,988	3,988	3,988	3,988	3,988	3,988
R-squared	0.082	0.092	0.900	0.101	0.435	0.749	0.586	0.768	0.913
Municipality FE	No	No	Yes	No	No	Yes	No	No	Yes
School-Track FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This table reports estimations of OLS regressions where the dependent variable is the (estimated) individual probability to be enrolled into the high track. Columns (1)-(3) present results when the individual probability is estimated using conditional logit models with no individual heterogeneity. Columns (4)-(6) show the estimates when the individual probability is recovered using a conditional logit with individual heterogeneity. Columns (7)-(9) report the results when the individual probability is estimated using the BLP algorithm. The categories excluded are the bottom 10th percentile on the CITO exam scores, the primary school advice indicating the student should follow the high track, and whether the track is offered by a roman-catholic school. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 7: Instrument Selection - First Stage Regressions

Instrument	Estimate (1)	S.E. (2)	Observations (3)	R-squared (4)	F-stat (1,285) (5)
Distance to city hall	-0.305	(0.186)	286	0.945	2.695
$rivals(0,2)$	0.642	(0.422)	286	0.945	2.318
$rivals(2,4)$	0.196	(0.281)	286	0.945	0.491
$rivals(4,6)$	-1.231***	(0.135)	286	0.958	83.52
$rivals(0,3)$	0.843***	(0.278)	286	0.948	9.227
$rivals(3,6)$	-0.778***	(0.084)	286	0.954	86.54
$rivals(6,9)$	-0.651***	(0.079)	286	0.955	67.45
$rivals(0,4)$	0.230	(0.219)	286	0.945	1.101
$rivals(4,8)$	-0.533***	(0.059)	286	0.956	82.49
$rivals(8,12)$	-0.303***	(0.090)	286	0.947	11.38
$rivals(0,5)$	0.085	(0.149)	286	0.945	0.331
$rivals(5,10)$	-0.416***	(0.046)	286	0.955	81.97
$rivals(10,15)$	-0.060	(0.078)	286	0.945	0.583
$rivals(0,6)$	-0.440***	(0.115)	286	0.948	14.69
$rivals(6,12)$	-0.303***	(0.052)	286	0.950	33.47
$rivals(12,18)$	0.013	(0.061)	286	0.945	0.049

Notes: This table reports first stage regressions where the dependent variable is the average commuting distance per school-track measured in kilometers. Each row corresponds to a single regression where the specific instrument was implemented. All specifications include all school-track attributes included in the conditional logit estimations, plus a set of school board and municipality fixed effects. We report in the last column the F-statistic used to test whether the instrument can be considered as a weak one. Robust standard errors are reported in parentheses. * p-value $<$ 0.1 ** p-value $<$ 0.05 *** p-value $<$ 0.01.

Table 8: Instrument Exogeneity

	$rivals(3,6)$ (2)	$rivals(4,6)$ (1)	$rivals(4,8)$ (3)
1[High Track]	-0.061	-0.066	-0.125
	(0.084)	(0.059)	(0.129)
1[Low Track]	0.070	0.076	0.144
	(0.084)	(0.064)	(0.140)
Observations	286	286	286
R-squared	0.004	0.009	0.007
F-stat (2,284)	0.615	1.309	0.993
p-value	0.541	0.272	0.372

Notes: In each specification the dependent variable is the fitted residuals from regressing each instrument against a set of school board and municipality fixed effects, plus all school-track attributes with the exception of the academic track type. Standard errors are clustered at the school-track level and reported in parentheses. * p-value $<$ 0.1 ** p-value $<$ 0.05 *** p-value $<$ 0.01.

Table 9: Mean Utility Regressions

	Reduced Form			First Stage			IV			OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Average Distance (β^m)							-0.916***	-0.898***	-0.906***	-0.808***
							(0.120)	(0.132)	(0.103)	(0.039)
1[High Track] (γ_h)	4.556*	-2.974***	-3.266***	5.082**	9.883***	10.203***	4.807***	4.689***	4.742***	8.906***
	(2.697)	(0.293)	(0.308)	(2.459)	(0.414)	(0.454)	(0.955)	(1.035)	(0.828)	(1.337)
1[low Track] (γ_l)	3.665	-3.833***	-4.162***	5.730**	10.497***	10.857***	4.510***	4.382***	4.439***	8.547***
	(2.709)	(0.272)	(0.299)	(2.456)	(0.379)	(0.439)	(1.026)	(1.112)	(0.885)	(1.356)
1[Public School]	0.421	-0.041	-0.521	-0.696	-0.199	0.334	-0.216	-0.220	-0.218	-0.241
	(0.410)	(0.409)	(0.413)	(0.470)	(0.487)	(0.504)	(0.311)	(0.304)	(0.313)	(0.296)
1[Alternative School]	-3.058***	-2.607***	-3.643***	2.586***	2.093**	3.236***	-0.690*	-0.727**	-0.710**	-0.916***
	(0.764)	(0.743)	(0.776)	(0.885)	(0.877)	(0.901)	(0.367)	(0.365)	(0.329)	(0.266)
Fail Rate (10%)	-0.318	-0.404	-0.415	0.256	0.345	0.360	-0.084	-0.093	-0.089	-0.142
	(0.306)	(0.293)	(0.288)	(0.281)	(0.265)	(0.259)	(0.212)	(0.213)	(0.210)	(0.200)
Central Exam Score	0.117	0.061	0.081	-0.102	-0.043	-0.064	0.023	0.022	0.022	0.013
	(0.204)	(0.201)	(0.204)	(0.151)	(0.147)	(0.154)	(0.149)	(0.149)	(0.149)	(0.150)
Instruments										
$rivals(4,6)$	0.808***			-0.882***			Yes	No	No	No
	(0.165)			(0.174)						
$rivals(3,6)$		0.522***			-0.581***		No	Yes	No	No
		(0.107)			(0.109)					
$rivals(4,8)$			0.406***			-0.448***	No	No	Yes	No
			(0.069)			(0.063)				
Observations	2,328	2,328	2,328	2,328	2,328	2,328	2,328	2,328	2,328	2,328
R-squared	0.824	0.823	0.828	0.964	0.964	0.965	0.928	0.928	0.928	0.930
F-stat(1st Stage)	-	-	-	-	-	-	25.78	28.47	50.34	-

Notes: This table reports results of mean utility regressions. Columns (1)-(3) present reduced form estimates. Columns (4)-(6) show first stage results. Columns (7)-(9) display instrumental variables coefficients. In column (10) we report OLS results. In the reduced, structural, and OLS specifications, the dependent variable is the mean utility levels recovered by using the BLP algorithm. In the first stage forms, the dependent variable is the average cycling distance (in kilometers) to each school-track in the sample. All specifications include a set of school board and municipality fixed effects. Robust standard errors are reported in parentheses. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 10: Willingness to Travel for School-Track Features

	Hedonic Model I (OLS)	Hedonic Model II (OLS)	BLP (IV)
Dependent Variable	Average Cycling Distance	Current Cycling Distance	Mean Utility Values
	(1)	(2)	(3)
1[High Track]	9.514***	11.674***	5.233***
	(0.557)	(0.819)	(0.446)
1[Low Track]	10.088***	11.569***	4.898***
	(0.548)	(0.822)	(0.520)
1[Public School]	0.226	0.704*	-0.241
	(0.521)	(0.360)	(0.343)
1[Alternative School]	2.095**	0.483	-0.784*
	(0.896)	(0.437)	(0.424)
Fail Rate (10%)	0.542*	0.347	-0.098
	(0.284)	(0.243)	(0.236)
Central Exam Score	0.095	0.126	0.025
	(0.172)	(0.145)	(0.164)
Observations	2,328	2,328	2,328

Notes: This table report wilingness to travel estimates from OLS hedonic model regressions (Columns (1) and (2)) and IV regressions (Column (3)). The interpretation of the coefficients is the kilometers parents are willing to let their children cycle to attend a school-track with a particular attribute. Standard errors are clustered at the school-track level, computed using the delta method, and reported in parentheses. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 11: WTT for Pre-University Tracks by Parental Educational Background

	Hedonic Model I (OLS)	Hedonic Model II (OLS)	BLP (IV)
Dependent Variable	Average Cycling Distance	Current Cycling Distance	Mean Utility Values
	(1)	(2)	(3)
Panel A: High Track Valuations			
1[High Track]	6.775***	4.300***	4.230***
	(0.540)	(0.697)	(0.403)
1[High Educated]	-0.093	0.118	0.957***
	(0.421)	(0.243)	(0.161)
1[High Track]×1[High Educated]	0.192	-0.156	0.763***
	(0.487)	(0.320)	(0.188)
Panel B: Low Track Valuations			
1[Low Track]	7.444***	4.125***	4.519***
	(0.539)	(0.695)	(0.430)
1[High Educated]	0.099	-0.038	1.719***
	(0.246)	(0.205)	(0.177)
1[Low Track]×1[High Educated]	-0.192	0.156	-0.763***
	(0.487)	(0.320)	(0.188)
Observations	2,328	2,328	2,328

Notes: This table report wilingness to travel estimates from OLS and IV regressions, by parental educational attainments. Panel A presents valuations for the high track. Panel B shows valuations for the lower track. Columns (1) and (2) present OLS estimates from hedonic regressions where the dependent variables are the average and current commuting distances, respectively. Column (3) displays IV estimates where the instrumental variable is the number of rival schools located in a concentric ring between 4km. and 8km far from each school. The interpretation of the coefficients is the kilometers parents are willing to let their child cycle to attend a school-track with a particular attribute. Standard errors are clustered at the school-track level, computed using the delta method, and reported in parentheses. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 12: Differences in WTT for Pre-University Tracks by Students' Features

low vs. high-educated	Low Track (1)	High Track (2)
CITO Score > P_{90}	0.3007	-0.1571
	(0.3715)	(0.1887)
CITO Score $[P_{75}, P_{90}]$	-0.2018	-0.2511
	(0.1376)	(0.1874)
CITO Score $[P_{50}, P_{75})$	-0.1170	-0.3596**
	(0.0862)	(0.1737)
1[advice: high track]	-0.3626**	-1.4656***
	(0.1632)	(0.2381)
1[advice: mixed]	-0.2570	-1.1825***
	(0.1656)	(0.2328)
1[advice: low track]	-0.4677***	-1.4374***
	(0.1523)	(0.2237)
1[Female]	-0.3886***	-0.1147
	(0.0968)	(0.0733)
#Peers primary school	0.0067	0.0336
	(0.0211)	(0.0226)
Observations	2328	2328

Notes: This table report differences in the wilingness to travel from IV regressions, for pupils with low and high-educated parents. Column (1) presents results for the high track. Columns (2) presents result for the low track. The excluded instruments in all specifications are the distance (in km.) from each school to the closest city-hall, and the number of rival schools located in a ring between 4km. and 8km far from each school. The interpretation of the coefficients is the kilometers low-educated parents are willing to let their child cycle to attend a school-track with a particular attribute, relative to high-educated parents. Standard errors are clustered at the school-track level, computed using the delta method, and reported in parentheses. * p-value < 0.1 ** p-value < 0.05 *** p-value < 0.01.

Table 13: Individual Compensated Valuation for the Absence of one Academic Track

	Mean	Max	Min
No High Track			
Km	0.9447	2.796	0.2113
Euros	0.3212	0.9505	0.0718
No Low Track			
Km	0.7288	2.259	0.0367
Euros	0.2478	0.7684	0.0125

Notes: This table report the compensated variation per student if a particular track is not longer available in the pupil's original choice set. This compensation is measured per day attended at school. In the computation was assummed a minimum hourly wage of 2.70 Euros, and a average speed of 16km/hour.

Appendix: Tables and Figures

Table A.1.1: Aggregated Market Shares

Municipality	Low Educated Parents		High Educated Parents	
	Low Track	High Track	Low Track	High Track
1	0.223	0.233	0.190	0.517
2	0.237	0.090	0.327	0.346
3	0.081	0.189	0.154	0.577
4	0.248	0.215	0.339	0.459
5	0.253	0.253	0.167	0.667
6	0.240	0.143	0.273	0.453
7	0.267	0.184	0.368	0.355
8	0.291	0.189	0.228	0.402
9	0.313	0.250	0.071	0.786
10	0.284	0.203	0.307	0.511
11	0.210	0.250	0.293	0.488
12	0.210	0.200	0.216	0.588
13	0.190	0.167	0.290	0.290
14	0.257	0.286	0.292	0.479
15	0.212	0.192	0.176	0.588
16	0.230	0.224	0.224	0.538
17	0.283	0.260	0.275	0.391
18	0.162	0.162	0.478	0.261
19	0.235	0.235	0.379	0.362
20	0.254	0.254	0.302	0.528

Notes: In this table we report market shares of non-vocational tracks only. The market share of the outside option should be calculated as 1 + sum of market shares for the high and low tracks. For instance, the market share of the outside option at the market composed by high educated families at municipality one is 1-(0.233+0.223).

Table A.1.2: Summary Statistics -School-Track Attributes per Student Features

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CITO Score $> P_{90}$	CITO Score $(P_{75}, P_{90}]$	CITO Score $(P_{50}, P_{75}]$	Advice: High Track	Advice: Mixed	Advice: Low Track	Female	Distance $> 4km.$
1[High Track]	0.955	0.793	0.413	0.875	0.559	0.224	0.334	0.361
	(0.208)	(0.405)	(0.493)	(0.331)	(0.497)	(0.417)	(0.472)	(0.481)
1[Low Track]	0.0421	0.191	0.449	0.119	0.370	0.637	0.267	0.254
	(0.201)	(0.393)	(0.498)	(0.323)	(0.483)	(0.481)	(0.442)	(0.436)
1[Public School]	0.129	0.115	0.105	0.128	0.0991	0.101	0.0654	0.0711
	(0.336)	(0.319)	(0.307)	(0.334)	(0.299)	(0.302)	(0.247)	(0.257)
1[Alternative School]	0.155	0.140	0.107	0.162	0.0903	0.142	0.0883	0.0447
	(0.363)	(0.347)	(0.309)	(0.369)	(0.287)	(0.349)	(0.284)	(0.207)
Fail Rate (10%)	0.769	0.798	0.744	0.723	0.673	0.815	0.508	0.527
	(0.715)	(0.673)	(0.669)	(0.568)	(0.527)	(0.712)	(0.624)	(0.666)
Central Exam Score	0.318	0.0605	0.0493	0.241	0.120	0.0678	0.0419	0.0531
	(1.190)	(1.045)	(0.935)	(0.984)	(0.862)	(0.999)	(0.753)	(0.841)
N	309	628	924	877	565	612	2049	1857

Notes: This table reports means and standard deviations of school-track attributes by pupil's characteristics. Columns (1)-(3) present results for pupils whose CITO test score is above the 90th Percentile, between the 90th and 75th Percentile, and between the 75th and 50th Percentile, respectively. Columns (4)-(6) show summary statistics for pupils receiving a primary school advice advocating to follow the high track, both tracks, and the Low track, respectively. Column (7) report summary statistics for female students. Column (8) shows results for those pupils who commute more than four kilometers to their prefered alternative. Standard deviations reported in parentheses.