KATHOLIEKE UNIVERSITEIT LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

Participation and schooling in a public system of higher education

by

Stijn KELCHTERMANS
Frank VERBOVEN

Econometrics

Center for Economic Studies
Discussions Paper Series (DPS) 06.11
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

April 2006

DISCUSSION PAPER

Participation and Schooling in a Public System of Higher Education

Stijn Kelchtermans and Frank Verboven*

April 2006

Abstract

We analyze the determinants of participation (whether to study) and schooling (where and what to study) in a public system of higher education, based on a unique dataset of all eligible high school pupils in an essentially closed region (Flanders). We find that pupils perceive the available institutions and programs as close substitutes, implying an ambiguous role for travel costs: they hardly affect the participation decisions, but have a strong impact on the schooling decisions. In addition, high school background plays an important role in both the participation and schooling decisions. To illustrate how our empirical results can inform the debate on reforming public systems, we assess the effects of tuition fee increases. Uniform cost-based tuition fee increases achieve most of the welfare gains; the additional gains from fee differentiation are relatively unimportant. These welfare gains are quite large if one makes conservative assumptions on the social cost of public funds, and there is a substantial redistribution from students to outsiders.

**Kelchtermans*: Catholic University of Leuven. Email: Stijn.Kelchtermans@econ.kuleuven.be. *Verboven*: Catholic University of Leuven and C.E.P.R. (London). Email: Frank.Verboven@econ.kuleuven.be. We thank participants at various conferences. Financial support of the Belgian Federal Science Policy Office (Interuniversity Attraction Poles P5/26) is gratefully acknowledged.

1 Introduction

Public systems of higher education are experiencing increased challenges in many European countries. The demand for higher education has more than doubled over the past thirty years. Public spending increased at the same rate, because governments maintained a policy of very low tuition fees and high subsidies per student. At the same time, there is little evidence that the public systems performed significantly better than the more market oriented systems in the Anglo-Saxon countries. If anything, the gap appears to be widening. According to the E.P.I. (Usher et al. (2005)), the countries with low private contributions are also often the countries with low student participation rates, and vice versa.[1] In terms of research performance, most top universities come from the Anglo-Saxon countries. Jacobs and van der Ploeg (2005) compare the relative performance of both systems in more detail. There is an increasing awareness that the public systems should incorporate more market-oriented principles, whether through centralized regulatory or more drastic decentralized reforms. The U.K. is a notable example of this evolution: tuition fees have recently increased rather drastically, accompanied with the introduction of income-contingent student loans in line with reform proposals by Barr (2004) and others.

Against this background we analyze the participation and schooling decisions in a public system of higher education. The participation decision is the decision whether to start with a higher education. The schooling decision is the decision where to study (which college or university) and what to study (which program). We consider the role of the pupils' travel costs, their high school background and demographics, based on a unique data set of all eligible high school pupils in the entire and essentially closed region of Flanders.[2] A key feature of our data set is the information on the pupils' locations, from which it is possible to compute the travel costs to all educational study options. Another key feature is the information on the educational choices at the highly disaggregate level of the study program.

One of our central findings is that pupils perceive the higher education institutions and programs as close substitutes. This implies an ambiguous role for travel costs: college proximity hardly influences the pupils' participation decisions, but it has a strong impact on their schooling decisions. Put differently, pupils are highly cost elastic for the decision where and what to study, but not for the decision whether to study. In addition, we find that the high school background plays an important role in both the participation and schooling decisions.

[1] The E.P.I. points out two exceptions, Finland and the Netherlands, where private contributions are among the lowest and overall participation in higher education among the highest.

[2] Flanders is the Northern part of Belgium, where Dutch is spoken. While access is open, in practice the undergraduate system has been quite closed from the French-speaking part and from other countries.

For example, pupils with a general high school background have an additional willingness to pay (in terms of travel costs) for a university than a college education of more than € 4000, compared with pupils of technical or professional high school backgrounds.

We use our empirical results to draw implications for reforming public systems of higher education. Many issues have occupied the current debate, including the effects of raising tuition fees, rationalizing supply, and moving from input-based to output-based subsidies. We focus on the effects of raising tuition fees, which have been uniform at very low levels in most European countries, including Flanders.[3] Our estimated cost elasticities imply that a uniform fee increase only has a small impact on overall participation, though comparatively more on pupils from the technical and professional high school backgrounds. Differentiated fee increases also have a low overall effect on participation but they imply large substitution effects (colleges versus universities; different study fields).

Our estimated welfare effects of tuition fee reform can be summarized as follows. (1) Uniform cost-based tuition fee increases achieve most of the attainable total welfare gains. The additional gains from fee differentiation by program type or program field (as opposed to uniform fee increases) are relatively unimportant. (2) The welfare gains are quite large if one makes conservative assumptions on the social cost of public funds. (3) If one ignores the social cost of public funds, the welfare gains from fee increases are small, but there is still a substantial redistribution from students to outsiders. The overall conclusion is that uniform fee increases achieve most of the total welfare gains, as well as a fairer distribution between students and outsiders (if properly accompanied with student loans).

Previous empirical research has focused on the effects of tuition fees on overall participation in higher education. These studies typically focus on the U.S. where it is possible to exploit variation in tuition fees at the state level. Estimates on the participation effects of an increase in fees by $ 1,000 are in the range of 3–8%; see Kane (1995), Dynarski (2003) and Cameron and Heckman (2001). We find a considerably lower participation effect of about 1%, which is perhaps not surprising since the current level of fees is low in the public system we consider.

Only a few number of studies have looked at the effects of tuition fees at the level of the institution, though not at the more detailed level of the program field. Most closely related to our work is Long (2004), who conducted a comprehensive study on the role of tuition fees

[3] To identify the effects of tuition fees, we assume that pupils respond in the same way to tuition fees as to mobility or other monetary costs. This is a realistic assumption to the extent that capital constraints do not become more binding after tuition cost increases. We therefore assume that tuition fee increases can be accompanied with income-contingent loans, i.e. loans that can be paid back later depending on the realization of income.

and travel costs at the level the institution.[4] She first estimates a conditional logit model for the schooling decision (where to study) and subsequently a binary logit model for the participation decision. She finds that tuition fees have a higher impact on the schooling than on the participation decision. We extend Long's model in several respects. First, we consider the participation and schooling decisions in an integrated nested logit framework. This approach allows us to infer the impact of tuition fees on both participation and schooling from variation in travel costs among pupils (in the absence of tuition fee variation in a public system).[5] Second, we consider the schooling decision at the even more disaggregate level of both the institution and the program field. Finally, we conduct a total welfare analysis on the effects of uniform and differentiated tuition fee increases, an issue that is perhaps of stronger relevance in the European public systems than in the more market-oriented U.S. system.

The remainder of this paper is organized as follows. In the next section we discuss some key features of the higher education system in Flanders in 2001. This is representative for several other public systems of higher education, and it introduces our subsequent questions. In section 3 we introduce the empirical model of educational choice, and our estimation approach to handle the very large data set. The fourth section discusses the empirical results. We compare the estimates from our disaggregate nested logit model (at the level of the institution and study program) with those of aggregate logit and nested logit models. We also compute the estimated cost elasticities at various levels (total market level, colleges versus universities, and the four main program fields). Finally, we estimate the welfare effects of uniform and differentiated fee increases. Section 5 concludes.

2 The market for higher education in Flanders

We begin with a description of the relevant supply and demand characteristics of this market, which is representative for many other public systems of higher education. Our description applies to the year of our data set, 2001-2002.

[4] Several previous studies incorporated distance to college to explain the participation decision, see e.g. Rouse (1995) or Frenette (2003). However, with the exception of Long (2004) these studies are not at the level of the institution, so that distance is usually proxied by the distance to the most nearby college, regardless of whether that college is chosen.

[5] We are able to compute the distance and travel costs of each pupil to each alternative based on information of the pupils' home address postal code, whereas Long (2004) only observes the pupils' high school address postal code.

2.1 Supply of higher education

Institutions and programs

There are two types of higher education institutions in Flanders: colleges and universities. The colleges focus exclusively on teaching and offer vocational study programs, which are oriented to professional training. Universities are also active in research, and they offer academic study programs. Both colleges and universities either have a catholic or a non-catholic orientation. They sometimes have multiple campuses, especially the colleges.

The vocational programs offered at the colleges are either short programs (one cycle of 3 years), or long programs (two cycles comprising a total of 4 or 5 years). The academic programs at universities are always long programs (two cycles).[6] The programs can be divided into four fields: arts, social sciences, biomedical sciences, and exact sciences. Each field consists of several elemental study options, e.g. nursing (a vocational program in biomedical sciences) or civil engineering (an academic program in exact sciences).

Table 1 provides an overview of the supply of higher education in the year 2001-2002. The top panel shows the number of campuses, broken down by type of institution and program field. The total number of campuses is 53, the majority being college campuses (44 versus 9 university campuses). Colleges more often have a catholic affiliation, whereas universities more often have a non-catholic affiliation. Colleges show a higher degree of specialization, since they typically do not offer all program fields. For example, only 12 out of the 44 college campuses offer arts. Universities tend to be less specialized. All non-catholic universities offer programs in biomedical and exact sciences, and all catholic universities offer programs in arts and social sciences.

The bottom panel of Table 1 shows the corresponding student numbers. The number of students is higher at colleges than at universities (25,182 versus 12,299). But the number of college campuses is comparatively even higher, so that the average scale at college campuses is generally lower than at university campuses. It is also evident that most students choose programs in social sciences, especially at the colleges, followed by programs in the exact sciences.

To give an idea of the geographic coverage of higher education supply across the region

[6] In recent years, the long vocational programs at colleges have shown a trend towards convergence to their academic counterparts at the universities (e.g. the economics or the engineering programs). This development has in part been stimulated by the government. Already in 1991, a Decree stipulated the same rules for two-cycle vocational programs at colleges as for academic programs at universities. Colleges offering two-cycle programs also became entitled to do applied research by means of co-operation agreements with universities. More recently, the Bologna Declaration leading to the Bachelor–Master reforms has strengthened these developments.

of Flanders, Figure 1 shows the locations of the campuses on a map. The upper map refers to the universities, the lower one to the colleges. Focus only on the circle areas for now. Each circle refers to a different campus, and is proportional to the number of first-year students. The figure shows that there is broad geographic coverage of higher education, with the exception of the "corners" in the West and in the East. However, this broad coverage is entirely due to the colleges. University coverage is concentrated around two main university cities (Ghent and Leuven).

The role of the government

As in most other European countries the undergraduate system of higher education in Flanders is entirely public. We only provide a very stylized overview here. Van Heffen and Lub's (2003) country report provides a more detailed description.[7]

Both universities and colleges receive subsidies for teaching; universities in addition receive subsidies for research. The subsidies for teaching consist of a fixed and a variable component. The variable subsidies differ across the various study programs to account for differences in the variable cost per student. According to CHEPS (Deen et al. (2005)), the cost per student tends to be lower for classroom-based programs (arts and social sciences) than for laboratory-based programs (biomedical and exact sciences). To account for this, the Flemish government has traditionally maintained a relatively simple system of four subsidy categories at colleges and three categories at universities. It has recently proposed to revise the rates to distinguish between additional categories, in line with the practices in seven benchmark countries.[8] Table 2 shows the (student-weighted) average variable subsidies per student, for the four program fields at colleges and universities. The first panel shows the averages using the current subsidy rates, the second using the proposed revised rates. The table shows that the subsidies, and hence the estimated variable costs per student, are lower in arts and social sciences than in biomedical sciences and exact sciences. They also tend to be lower at colleges than at universities. The proposed revised rates show a larger variation, especially across the different program fields.

[7]For shorter descriptions covering a large set of countries including the region of Flanders, we refer to Maassen (2000) or Eurydice (2000).

[8]These revisions fit in a larger policy reform proposal in 2005, which aims to transform the input-based subsidy system (based on the number of incoming students) to an output-based system (based on the number of outcoming students). To prepare these reforms, the input-based subsidies were frozen in 2000 and should become output-based in 2007. This development is not relevant for our purposes here; we are mainly interested in describing the variable subsidies as our proxy for how the goverment perceives the cost per student.

For a detailed description of the current (frozen) subsidy rates, see van Heffen and Lub (2003), and of the proposed revised rates, see Vandenbroucke (2005).

The variable subsidies are only part of the government's "first flow" budget on higher education.[9] According to Cantillon et al. (2005), the total budget is about € 8600 per student, so that the variable part only accounts for about 38 percent of public spending. The remaining part is independent of the number of students and can be viewed as a measure of the fixed costs to be covered.

In addition to the subsidies, the government intervenes in the tuition fees that the colleges and universities are allowed to set. While the government gives some discretion, the tuition fees show hardly any variation in practice. During the year of our study, 2001, the tuition fees were essentially uniform at € 425 for colleges and € 445 for universities. This shows that private contributions are extremely low, only about 5% of public higher education spending (excluding research).

The policy of high subsidies and low tuition fees may have adverse effects on both the diversity and the quality of the supply. First, colleges and universities may have incentives to offer too much diversity. The government therefore regulates the supply of programs. There is an official list of subsidizable programs, but not all institutions necessarily receive the authorization to offer all programs. The result is a specialization, which we illustrated earlier in Table 1. Second, the institutions may have limited incentives to provide sufficient quality. A system of quality assurance aims to provide sufficient incentives, through self-assessment and external visiting committees. In principle, the government can take away the authorization to offer a study program if quality is insufficient, though this rarely happens in practice.

2.2 Demand for higher education – summary statistics

We now discuss the demand for higher education. This also introduces our data set and subsequent econometric analysis of educational choice.

Every pupil with a high school degree is eligible to start with higher education. This is true regardless of the type of high school degree that has been obtained. There are three main types of high school degrees. A general high school degree provides a broad theoretical training as a basis to continue with higher education. A technical high school degree puts more emphasis on specialized technical-theoretic training but at the same time aims to provide a sufficiently general background to prepare for higher education. A professional high school degree focuses mostly on practical training. Pupils with a professional high school degree can still start with higher education provided that they have taken additional

[9] The first flow budget is the part of the budget that directly goes to teaching. The second flow and third flow budget are devoted to research.

courses during a seventh year of study.[10] In contrast with many other countries, there is no direct rationing of participation in higher education, whether through *numerus clausus* or through minimum course requirements or grades obtained during high school.[11] Each pupil who finishes high school is therefore in the position to make both the participation decision (whether to study) and the schooling decision (where and what to study) by selecting one option out of the full set of all available study alternatives.

To analyze this educational choice process, we combined two basic data sets, covering essentially the entire population of eligible pupils: a "pupils data set" of all 55,905 last year high school pupils in the year 2001; and a "students data set" of the 37,481 participating students.[12] For each pupil we observe sex, nationality, age, the high school institution, the high school degree (program), and the home address. The Appendix provides more detailed information on the two basic data sets, as well as on some additional auxiliary data sets, and how we combined them.

We constructed a number of relevant variables describing the pupils' profile, and we organize them in three groups. The first group consists of general demographics: sex, nationality and the religious affiliation of the high school. The second group contains the scholastic ability variables: years of repetition, the type of high school and the study program followed at high school. Years of repetition is the age minus 18, and measures the number of failures during high school. The type of high school (general, technical or professional) measures the intellectual background, as we discussed above. The study programs at high school offer additional information on ability and intellectual interests. For general high school, we distinguish among the following fields: classical languages, modern languages, economics, sciences and mathematics. These can be combined so they are not mutually exclusive. The brightest pupils often follow either classical languages or mathematics (or both). In technical high schools, there is a very large number of programs. For simplicity, we only distinguish between programs that are "people oriented" (e.g. beautician, interior design) and programs that are "product oriented" (e.g. car mechanics, construction techniques). There is also

[10] There is also a fourth high school category, arts. This also offers a quite practical training. We do not exclude them from our analysis, but since there are relatively few pupils here we include them in our base category.

[11] During the year of our study, some programs (e.g. engineering and medicine) indirectly limited the number of students through an entry examination. This does not function however as a mechanism to directly limit the number of students per year.

[12] These are the pupils who either enroll immediately after highschool in 2001 (36,111 students) or with one period of delay (1,370 students). Hence, the total participation rate of eligible pupils is 67%. Note that the actual participation rate in higher education is lower since only 79% obtain of the people obtain a high school degree.

a large number of programs in professional high schools, but since there are relatively few pupils graduating from professional schools (and even fewer that start with higher education) we do not distinguish explicitly between programs for this category.

The third group of variables refers to transportation costs: distance travelled to the campus (in km), and time travelled either by road or by train (in min). Time travelled by road is the fastest calculated route from the pupil's home postal code to her chosen campus postal code. Time by train is equal to the travel time by road to the nearest well-connected train station, plus the travel time on the train, plus a fixed 10 minutes to incorporate the time to get from the destination station to the campus. Based on these variables, we also constructed an annual commuting cost variable in monetary terms, for a student making 300 trips per year at a cost of € 0.25 per km and an opportunity cost of time of € 8 per hour[13]. This variable will enter our empirical model and it is further motivated at that point.

Table 3 shows the summary statistics: means for all the dummy variables (interpreted as fractions of the population), and means and standard deviations (in brackets) for the continuous variables. The first column shows the unconditional statistics, i.e. for all pupils. The remaining columns show the statistics conditional on the chosen alternative.

Demographics Slightly fewer males than females graduated from a high school in 2001 (48%), and they are comparatively less likely to participate in higher education (45%). More surprisingly, male pupils are less represented in catholic institutions. This may be due to preferences, or to the different supply offered at catholic institutions; our subsequent empiral model distinguishes between these possibilities by accounting for the combined choice of institution and program fields. Similarly, pupils with a foreign nationality are less likely to participate (2% of the pupils is foreign, but only 1% of the students), and those who do participate are less represented at catholic institutions.

The majority of the pupils (75%) attended a catholic high school, and they are more likely to participate (78% of the students). Furthermore, students from a catholic high school are much more likely to choose a catholic institution (87% of the students at catholic institutions have a catholic high school background, compared to 75% in general). This suggests there are still strong links between the catholic high schools and the higher education institutions.

[13] The kilometre cost is a commonly used measure of distance cost for tax purposes (Belgisch Staatsblad (2006)). The opportunity cost of time is representative for the hourly wage of student jobs (Jobdienst KU Leuven (2006)).

Ability The average years of repetitions is 0.6 (with a large standard deviation of 1.06), but it is much lower for participating students (0.36) than for outsiders (1.11).[14] Among the students, the average years of repetition is especially small for students going to universities rather than colleges (0.16 versus 0.48).

Up to 48% of the pupils come from a general high school, slightly more than one third from a technical high school, and only 17% from a professional high school.[15] Pupils from a general high school are more likely to participate (they make up 60% of the students, and only 48% of the pupils). Pupils with a technical high school background are more or less proportionally represented in higher education, but they mainly participate at colleges, and only form a small minority at universities. Pupils from a professional high school (the base) are least represented.

Pupils with a general high school background in mathematics or classical languages have a particularly strong propensity to go to universities rather than to colleges or staying out. Pupils from a "people oriented" technical high school have a strong presence at colleges (over one third).

Mobility Distance travelled is on average 35 kilometers, with a large standard deviation. Car and train travel time are on average 31 minutes and 48 minutes. The average annual commuting costs (including transportation and time costs) is € 3,800. The average travel distance is higher for students going to universities than for students going to colleges (42 versus 31 km), reflecting the lower geographic coverage of universities as shown earlier on the map in figure 1. The rather high standard deviations show there is a lot of variation in distance travelled across students.

Our earlier Figure 1 provides additional information on the role of proximity to institutions in explaining participation. The different shadings shows the different participation rates throughout the region. Part (a) suggests that the university participation rates are related to proximity: high participation rates (above 25%) are especially found around the two main universities, whereas low participation rates (below 15%) occur in the West and East corners of the region. Part (b) also suggests that the college participation rates vary across the region according to college proximity. Nevertheless, the pattern is less pronounced, probably due to the high geographic coverage throughout the region. Note that the areas with little university participation are often areas with a high college participation (see for

[14] As mentioned earlier, for the pupils of a professional highschool, there is one additional preparatory year before becoming eligible to higher education. We include this in our definition of repetitions.

[15] The total fraction of pupils in a professional highschools is much larger than 17%. However, a substantial number does not complete their degree, and we are looking here at last year pupils only.

example the low university and the high college participation in the West). This suggests that proximity may not matter that much for the decision whether to study as for the decision where to study. Whether this is indeed the case, will be addressed in our econometric model.

These summary statistics provide some preliminary suggestive findings on the role of demographics, ability and mobility in participation and schooling decisions. We are now ready to discuss how our econometric framework takes these characteristics into account.

3 Empirical framework

We now specify the empirical model of the pupils' participation and schooling decisions, i.e. their decisions whether to start with higher education studies, and, if so, where and what to study. We model their decision in a discrete choice framework in which individuals choose the alternative that maximizes random utility among the set of available alternatives. We adopt a version of the nested logit model (McFadden (1978)). This model is well suited to deal in a reasonably flexible way with the very large size of our data set: $55,905$ pupils, 563 choice alternatives, and a large set of observed variables describing pupils and choice alternatives. In particular, the model allows for consistent estimation with a large set of choice alternatives, by explicitly aggregating and/or sampling over the alternatives. At the same time, the model incorporates unobserved heterogeneity since individuals may have correlated preferences across alternatives belonging to the same nest. More general models of unobserved heterogeneity, such as mixed or random coefficient logit models (McFadden and Train (2000)) inhibit sampling over alternatives and are therefore not feasible given the size of our choice set. Since we capture a lot of observed individual characteristics anyway, the need for more general models of unobserved heterogeneity is less than in other applications.

3.1 The choice model

Each individual i chooses one out of a large set of choice alternatives. The individuals are the pupils who have finished high school and who are therefore eligible to start with higher education studies. The choice alternatives are defined by the institution (the university or college) and the actual program (e.g. nursing, civil engineering, etc.) There is also one no-study alternative. The total number of choice alternatives is very large (563, including the no-study alternative), and it is typically not feasible to systematically incorporate observable choice-specific variables at this level of detail. Our framework therefore explicitly deals with the possibility of aggregation and sampling over alternatives.

There are J study options, $j = 1 \cdots J$. For each study option j there are K_j variants or "elemental alternatives", $k = 1 \cdots K_j$. Alternative 0 is the no-study option. The total number of alternatives is $\sum_{j=1}^{J} K_j + 1 = 563$. The empirical analysis aggregates the K_j variants to the level of the study option j. For example, the study option may be defined at the level of the institution, or at the level of the institution, program type (short vocational, long vocational or academic) and program field (arts, social sciences, biomedical sciences or exact sciences). In both cases, the variants per study option are defined accordingly, i.e. as the actual programs offered under the defined study option. Note that this framework also covers the general case of no aggregation: simply define the study options as the institution and the actual program. Each study option then goes with a single variant, so that $K_j = 1$ for all j.

An individual i's utility for study option j and variant k, U_{ijk}, is the sum of a deterministic component V_{ijk} and a random component ε_{ijk}, i.e. $U_{ijk} = V_{ijk} + \varepsilon_{ijk}$. Assume that there is no choice-specific information at the level of the variant k. Hence, the deterministic component of utility is the same for each variant k of study option j, $V_{ijk} = V_{ij}$ for all k. Individual i's utility for study option j and variant k is then given by:

$$U_{ijk} = V_{ij} + \varepsilon_{ijk}.$$

Individual i's utility for the no-study option is:

$$U_{i0} = V_{i0} + \varepsilon_{i0}.$$

The random component of utility follows the distributional assumptions of a three-level nested logit model (McFadden (1978)). At the highest level, there are two nests: the study nest S, which includes the study options and their variants; and the no-study nest, which is a degenerate nest with only alternative 0. At the lower level, the study nest S consists of the J different study option nests, $j = 1 \cdots J$. At the lowest level, each study option nest j consists of the K_j variants. The distribution of the ε_{ijk} and ε_{i0} has a sequential structure with two parameters, ρ and σ. The parameter ρ roughly measures the extent to which the ε_{ijk} show correlation between the J study option nests within the study nest S (i.e. correlation relative to the no-study option 0). Similarly, the parameter σ measures the extent to which the ε_{ijk} show additional correlation across the K_j variants within a given study option nest j.[16] A parameter close to zero means that the correlation is weak, while a parameter close to one means that the correlation is strong. To illustrate, if $\rho = \sigma = 0$ the model reduces to

[16] There is no parameter for the no-study nest, since it is a degenerate nest. Furthermore, we assume that the parameter σ is common for all J study option nests. It would be straightforward to allow this parameter to vary across the study option nests, e.g. according to the institution or program type or program field.

a simple logit model with no correlation between the ε_{ijk}. As another example, if $\rho = 0$ and $\sigma = 1$, then there is no correlation of the ε_{ijk} between the study option nests, while there is perfect correlation between the variants within each study option nest.

Individuals choose the alternative that maximizes random utility U_{ijk}. The nested logit model is consistent with random utility maximization if $0 \leq \rho \leq \sigma \leq 1$ (McFadden (1978)), i.e. correlation parameters are between zero and one, with a weaker correlation between than within the study option nests. Because the distribution of the ε_{ijk} and ε_{i0} has a sequential structure, the nested logit model yields simple expressions for the conditional choice probabilities: the probability of choosing a variant k within a study option nest $P_{ijk|j}$, the probability of choosing a study option j within the study nest $P_{ij|S}$, and the probability of choosing the study nest P_{iS}. We are not interested in the probability that individual i chooses variant k of study option j, i.e. $P_{ijk} = P_{ijk|j}P_{ij|S}P_{iS}$, since there is no choice-specific variation at the level of the variant k, $V_{ijk} = V_{ij}$. Our interest is thus only in the aggregate probability that individual i chooses any variant of study option j, i.e. $P_{ij} = P_{ij|S}P_{iS}$. Applying the formulas for the three-level nested logit model, the probability P_{ij} for $j = 1 \cdots J$ is:

$$P_{ij} = \frac{\left(\sum_{k=1}^{K_j} \exp(V_{ijk}/(1-\sigma))\right)^{\frac{1-\sigma}{1-\rho}}}{\sum_{j=1}^{J}\left(\sum_{k=1}^{K_j} \exp(V_{ijk}/(1-\sigma))\right)^{\frac{1-\sigma}{1-\rho}}} \frac{\left(\sum_{j=1}^{J}\left(\sum_{k=1}^{K_j} \exp(V_{ijk}/(1-\sigma))\right)^{\frac{1-\sigma}{1-\rho}}\right)^{(1-\rho)}}{\left(\sum_{j=1}^{J}\left(\sum_{k=1}^{K_j} \exp(V_{ijk}/(1-\sigma))\right)^{\frac{1-\sigma}{1-\rho}}\right)^{(1-\rho)} + \exp\left(V_{i0}\right)}.$$

Since the variants k within study option nest j have the same utility, $V_{ijk} = V_{ij}$ for all k, this can be simplified to:

$$P_{ij} = \frac{\exp(V_{ij}^*/(1-\rho))}{\sum_{j=1}^{J}\exp(V_{ij}^*/(1-\rho))} \frac{\left(\sum_{j=1}^{J}\exp(V_{ij}^*/(1-\rho))\right)^{(1-\rho)}}{\left(\sum_{j=1}^{J}\exp(V_{ij}^*/(1-\rho))\right)^{(1-\rho)} + \exp\left(V_{i0}\right)}, \tag{1}$$

where $V_{ij}^* = V_{ij} + (1-\sigma)\ln(K_j)$ can be interpreted as the aggregate utility of a study option j. This shows that we can consider a simplified two-level nested logit model at the level of the study option j, after simply including $\ln(K_j)$ as a correction term to V_{ij}; see also Ben-Akiva and Lerman (1985). The correction term captures the extent of unobserved heterogeneity within a study option. It drops out if $(1-\sigma)$ is equal to zero: the utilities of the variants of the same study option are perfectly correlated (homogeneous), so that additional variants do not lead to a higher aggregate utility of that study option. Note that the correction term also drops out in the general case of no aggregation, since in this case $K_j = 1$ for all j.

The probability that i chooses the no-study option P_{i0} is simply:

$$P_{i0} = 1 - P_{iS} = \frac{\exp(V_{i0})}{\left(\sum_{j=1}^{J} \exp(V_{ij}/(1-\rho))\right)^{(1-\rho)} + \exp(V_{i0})}. \quad (2)$$

It is instructive to compare this model to Long (2004), who has been the only author to consider the study options at the level of the individual institution. First, she does not include a correction term $(1-\sigma)\ln(K_j)$ to the utility terms of the study options. This is a special case of our model if $\sigma = 1$, i.e. homogeneity of the variants within the institution. Second, Long estimates her model sequentially. In a first step, she models the probability where to study, conditional on choosing to study, i.e. $P_{ij/S}$ in our notation. In a second step, she models the probability whether or not to study, i.e. P_{iS} and P_{i0}, using the characteristics of the predicted most preferred alternative as explanatory variables in the deterministic component of utility. This may also be viewed as a special case of our model for $\rho = 1$. Indeed, as ρ goes to one, we have $(1-\rho)\ln\left(\sum_{j=1}^{J} \exp(V_{ij}^{*}/(1-\rho))\right) = \max(V_{i1}^{*} \cdots V_{iJ}^{*})$, so that P_{iS}, as given by the second part of (1), reduces to Long's binary choice model.[17]

3.2 Indirect utility

We specify the deterministic component of utility V_{ij} of individual i for alternative j as a conditional indirect utility function. It depends on the expected benefits, including monetary returns in the form of increased future salaries, and on the expected costs, i.e. the non-monetary costs of studying and the monetary costs in the form of tuition fees and travel costs. In our application, tuition fees are low and do not show any variation across alternatives. However, individuals pay an implicit price in the form of travel or commuting costs: transportation costs and the opportunity cost of time. We consider the following specification of V_{ij}:

$$V_{ij} = \beta_j + w_i'\gamma_j + \alpha_i(y_i - t_j - g(x_{ij})), \quad (3)$$

where w_i is a vector of individual characteristics (sex, age, high school background, etc.), y_i is individual i' s annual income, t_j is the tuition fee for study option j (currently uniform at about € 500 for all $j \neq 0$, but possibly differentiated across study options after policy reform), and $g(x_{ij})$ is the implicit price paid by individual i for alternative j, which is a function of the annual travel or commuting costs x_{ij}.

In principle, one may include a full set of alternative-specific intercepts β_j and alternative-specific slope vectors γ_j (relative to one base alternative). For example, women may have a

[17] This comparison is not entirely accurate. Long's specification includes variables such as distance in both steps of the estimation. For her approach to be exactly a special case of our random utility framework, the coefficients of these common variables should have been restricted to be the same.

different valuation for every alternative than men. In practice, such flexibility would imply a very large number of parameters to be estimated, since the set of alternatives is large, and there are many individual characteristics in the vector w_i. To reduce the number of parameters, we will summarize the alternative-specific effects β_j and γ_j by a more limited set of characteristics, i.e. the institution's religious orientation, the program type and the program field. A more precise discussion of the included variables is given in the next section.

The second part of (3) refers to the utility derived from the consumption on other goods (i.e. other than the educational choice), after spending an implicit price $g(x_{ij})$, which is an increasing function of the annual commuting costs x_{ij}. The parameter α_i can be interpreted as the marginal utility of income of individual i. It can be used to quickly reinterpret the other parameters such as the γ_j in (3) in monetary terms (by simply dividing γ_j by α_i), and to conduct a more complete welfare analysis.

The annual commuting costs x_{ij} of individual i for alternative j consist of two components: transportation costs and the opportunity cost of time (McFadden and Train (1978)). The transportation costs (in Euro) are proportional to the distance per trip d_{ij} (in km). The opportunity cost of time (also in Euro) is proportional to the travel time per trip t_{ij} (in min). More precisely, specify the annual commuting costs as $x_{ij} = 75d_{ij} + 40t_{ij}$.[18] Each individual has two options: commute or go on residence. If she commutes, her implicit price for alternative j is simply $g(x_{ij}) = x_{ij}$. If she goes on residence, she saves a fraction ϕ of the trips, but pays an extra annual cost on rent r_j. Her implicit price is correspondingly $g(x_{ij}) = (1 - \phi)x_{ij} + r_j$. A cost-minimizing individual commutes if $\phi x_{ij} \leq r_j$, and goes on residence otherwise. Intuitively, commuting is preferred if the annual commuting costs are sufficiently small relative to the annual cost of rent. The deterministic component of utility (3) for a cost-minimizing individual can then be written as:

$$V_{ij} \; = \; \beta_j + w_i'\gamma_j + \alpha_i \left(y_i - t_j - x_{ij}\right) + \alpha_i(\phi x_{ij} - r_j)\, I(\phi x_{ij} - r_j), \tag{4}$$

where $I(\cdot)$ is an indicator function equal to 1 if the expression inside the brackets is positive, and equal to 0 otherwise. Utility thus decreases in the commuting costs x_{ij} in a piecewise linear way: at a steeper rate α_i for low values of x_{ij} (when the pupil commutes), and at a flatter rate $\alpha_i\phi$ for high values of x_{ij} (when the pupil goes on residence).

The utility specification (4) holds for the J study options, $j = 1 \cdots J$, as well as for the no-study option 0. For the study options, x_{ij} has the clear interpretation of the commuting

[18] This assumes that a commuter engages in 10 trips per week during 30 weeks of the year, at a transportation cost of 0.25 Euro/km and an opportunity cost of time of 8 Euro/hour. The annual transportation cost per kilometer (in Euro) is then $10 \cdot 30 \cdot 0.25 = 75$, and the annual opportunity cost of time per minute (in Euro) is $10 \cdot 30 \cdot (8/60) = 40$.

costs of individual i to study option j. This may clearly vary across individuals and study options. For the no-study option, x_{i0} can be interpreted as the commuting costs to work. We assume this to be constant across individuals, $x_{i0} = x_0$, and to be sufficiently small so that the last term does not enter in V_{i0}.[19]

It is again instructive to compare this specification to Long (2004). She allows utility to vary quadratically with distance, and finds that utility decreases with distance at a decreasing rate. Our commuter/resident specification yields the same degree of flexibility. In fact, a specification of V_{ij} that is quadratic in x_{ij} indeed gave a similar fit. The advantage of our approach in this context is that the coefficient of x_{ij}, i.e. α_i, can also be interpreted as the marginal utility of income. In Long's specification, the price (i.e. tuition fee) coefficient can be directly interpreted as the marginal utility of income, but this coefficient is not identified here since tuition fees show no variation across alternatives. In sum, our specification allows us to capture the role of distance in the same flexible way as Long's, while at the same time providing us a marginal utility of income coefficient in the absence of any variation of a direct measure of price.

3.3 Estimation

The choice probabilities (1) and (2) may be used to construct the likelihood function and estimate the model. There are, however, practical difficulties due to the size of our data set.

- There is a very large number of individuals $(55,905)$.
- Each individual can choose from a very large number of alternatives (563, including the no-study alternative).
- The associated choice-specific variables need to be interacted with many individual characteristics (the demographic and high school background variables).

There are several ways to reduce the size of the data set: aggregation over alternatives, sampling over alternatives, and sampling over individuals. We adopt a combination of these approaches.

First, we aggregate over alternatives. Using the framework in section 3.1, we aggregate the variants k to the level of the study option j by adding the correction term $(1-\sigma)\ln(K_j)$ to V_{ij}. We consider two definitions of the study option j. Our "aggregate" model defines the study options as the 53 different institutions. This model serves as a useful benchmark since it considers the same aggregation level as in Long (2004). It discards, however, almost

[19] The exact value of x_0 is irrelevant and can be normalized since it is not separately identified from β_0.

all of the information on the study programs (except for including the correction term for the number of programs offered at the institution $\ln(K_j)$). Since this is a main source of richness of our data, which we want to exploit in our empirical analysis, we will focus on a "disaggregate" model. This model defines the study options as the 154 different institutions, program types, and program fields (though it still aggregates over the actual programs within a field, e.g. nursing which is one of the vocational social sciences programs). Utility V_{ij} can then be specified to depend on both the institutions' characteristics and on the program characteristics down to the level of the program field.

Second, we sample over alternatives; this is only necessary to estimate our disaggregate model. In simple logit models one can randomly select a reduced choice set for each individual, and define the choice probabilities as if the individuals only faced this reduced choice set. Maximum likelihood estimation based on these as-if choice probabilities yields consistent estimates (McFadden (1978)). We extend this approach to the nested logit model by exploiting its sequential structure. We first consider the probability of choosing a study option j, conditional on choosing to study, i.e. $P_{ij|S}$ entering as the first term in (1). This is a simple logit probability, so that it is possible to sample over alternatives and obtain consistent estimates. More precisely, for each student we construct a reduced choice set of 20 study options, i.e. the chosen study option and a random sample of 19 other study options. This gives consistent estimates for the parameters entering the study option utilities V_{ij}. We subsequently consider the probability whether to study, i.e. $P_{i|S}$ entering as the second term in (1). Provided that the utilities of all the study options are now included, as computed from the parameter estimates of the first stage, this yields consistent estimates of the parameters entering the no-study option V_{i0} and of the distributional parameter ρ. Since this is a two-step estimation procedure, the standard formulas for the standard errors of the parameters computed in the second step (i.e. those for the no-study option and ρ) are not correct. We follow the general procedure of Murphy and Topel (1985) to obtain the corrected standard errors.

Finally, we sample over individuals. In general, there is a trade-off between sampling over alternatives and sampling over individuals (doubling the size of the sampled choice set implies halving the number of individuals to keep the size of the data set fixed). Our experience showed that it is more efficient to sample over the alternatives than over the individuals, in particular to identify the utility effects of some relatively unpopular alternatives with few observations. We therefore sampled much less heavily over individuals than over alternatives. For both the aggregate and the disaggregate models we sampled about $20,000$ out of the $55,905$ individuals, as compared to a sampled choice set of 20 out of the 155 alternatives.

3.4 Alternative specifications

We compare three different models.

- In the aggregate logit model the choice alternatives are at the level of the institution. The choice set thus consists of 54 alternatives, i.e. 53 study options and one no-study option.
- In the aggregate nested logit model the choice alternatives are again at the level of the institution, but there is now a nesting parameter ρ that may be different from 0. This allows for unobserved heterogeneity in that individuals may have correlated preferences across the 53 study options, even after conditioning on the observable characteristics.
- In the disaggregate nested logit model the choice set consists of 155 alternatives (including one no-study option), referring to the institutions, the program types (long and short vocational, and academic) and the program fields (exact sciences, biomedical sciences, social sciences and arts). There is again a nesting parameter ρ to allow for correlated preferences across the 154 study options.

The first two specifications correspond to the previous literature, which also looked at the participation and schooling decisions at the aggregate institution level. As mentioned earlier, there are still two important differences with the most comprehensive study to date, i.e. Long (2004): the inclusion of the correction term $\ln(K_j)$ to account for the aggregation over the different programs within each institution; and the integrated consideration of the participation and schooling decisions. The third specification is at the more detailed level of the institution and the program field, which has not been considered in previous work. Since this third specification is the most general, we will focus our discussion around it, and use the results from the two aggregate models mainly as a point of comparison with previous work.

It remains to specify the variables entering the indirect utility V_{ij} as given by (4), i.e. the individual characteristics entering w_i and α_i, and the choice characteristics entering β_j and γ_j. The vector of individual characteristics w_i includes the following 12 variables: sex, nationality, years of repetition during high school (age $-$ 18), high school's religious orientation (catholic or not), and 8 variables referring to the type of high school education (i.e. various forms of general and technical high school, relative to a professional high school education). Similarly, we specify the marginal utility of income or the commuting cost parameter α_i to depend on the full vector of individual characteristics w_i. Hence, $\alpha_i = w_i'\alpha$, where α is the corresponding vector of parameters.

The number of alternative-specific intercepts β_j and slope vectors γ_j is very large. In particular, in the disaggregate model with 155 choice alternatives and 12 individual characteristics in w_i, there are up to $154 + 154 * 12 = 2002$ parameters to be estimated. To make estimation and interpretation feasible, we therefore put some structure on β_j and γ_j. The intercepts β_j are specified to depend on a full set of institution, program type and program field dummy variables. The slope γ_j^l corresponding to each individual's characteristic l depends on the following choice characteristics: a dummy variable for the no-study option (γ_0^l), and a set of dummy variables characterizing the study options: the religious affiliation of the institution (catholic or not), the program type (short-term vocational, long-term vocational or academic) and the program field (exact sciences, biomedical sciences, social sciences, or arts). We take the following study option as the base: a non-catholic institution offering a short-term vocational program in the field of arts. Hence, all estimated slopes γ_j^l should be estimated relative to that base.[20]

Descriptive statistics on the individual characteristics entering w_i and α_i, unconditional and conditional on the chosen alternative, were presented and discussed earlier in Table 3.

4 Empirical results

We begin with a discussion of the parameter estimates, to uncover the determinants of the participation and schooling decisions. Next, we summarize our key results through the cost elasticities implied by our estimates. Finally, we draw some implications on the welfare effects of uniform and differentiated tuition fee increases.

4.1 Parameter estimates

Tables 4 and 5 present the empirical results. Table 4 compares the estimates of several parameters across the three different models (aggregate logit, aggregate nested logit, and disaggregate nested logit): the nesting parameters ρ and σ, the slope parameters entering the utility of the no-study option (γ_0) and the commuting cost parameters (α). Table 5 shows the slope parameters entering the utility of the study options (γ_j, $j \neq 0$). This table only

[20]In the two aggregate models with 54 alternatives, the number of alternative-specific intercepts β_j and the alternative-specific slopes γ_j is lower, i.e. $53 + 53 * 12 = 702$, so that estimation of all parameters would be easier. However, for ease of comparison and interpretation we adopt a more parsimonious specification that is similar to the disaggregate model. The only difference is that the β_j and γ_j obviously no longer include dummy variables referring to long-term vocational program type and to the program fields.

shows the estimates for the disaggregate nested logit model, since there are many of these parameters.[21]

We first discuss the observed and unobserved determinants of participation (γ_0 and ρ in Table 4). Next, we highlight the role of mobility costs in both participation and schooling (α in Table 4). Third, we discuss the determinants of schooling (γ_j, $j \neq 0$ in Table 5). Finally, we briefly discuss the aggregation parameter (σ in Table 4).

Observed and unobserved determinants of participation (γ_0 and ρ)

The top part of Table 4 shows how individuals differ in their valuation of the no-study option. We focus our discussion on the estimates of the most general disaggregate nested logit model (third column). Males and especially foreigners have a significantly higher utility from staying out of higher education. The same holds true for older pupils, i.e. those who experienced repetitions during high school. Pupils from a catholic high school have a significantly lower utility from staying out than others. This is consistent with the reputation of the catholic high schools in providing a strong preparation for a higher education.

The most important individual characteristics affecting the participation decision relate to the pupils' type of high school. Pupils with a technical and especially those with a general high school education have a substantially lower utility from staying out relative to the pupils from a professional high school. This may be either due to the acquired or due to the intrinsic skills of these pupils. So one should be cautious and not conclude that promoting general high school education will improve participation in higher education. What matters for our purposes, is only that the type of high school background does play an important role, which will be reflected in our estimated cost elasticities. Finally, while the type of high school plays a crucial role, the specific discipline followed at the high school does not matter much in the participation decision. None of the so-called more difficult general high school disciplines, such as mathematics or classical languages, matter in the participation decision.

Most of the estimated γ_0 are of a similar order magnitude in the more restrictive aggregate models, but there are some important differences. For example, the aggregate logit model estimates pupils with a catholic high school background to have a higher utility from staying out, in contrast with common wisdom. This illustrates the importance of accounting for unobserved heterogeneity affecting the participation decision, as captured by the nesting parameter ρ. While the logit model restricts ρ to be equal to zero, the nested logit model estimates it as 0.9 and 0.95 in the aggregate and disaggregate versions. Hence, pupils have

[21] A comparison of the parameters in γ_j across models does not yield any main additional insights. Many of the parameters do not enter in the aggregate models, and the ones that do (relating to the institutions' religious affiliation and program type) were usually estimated to be similar.

strongly correlated preferences across all study options and view them as close substitutes relative to the no-study option.[22]

In sum, while several observed individual characteristics affect the utility of the no-study option, there remains a lot of unobserved heterogeneity affecting the participation decision.

The role of mobility costs in participation and schooling (α)

The next question concerns the role of the annual commuting costs, affecting the valuations of the study options relative to the no-study option. Table 4 shows that the annual commuting costs have a negative and highly significant effect on utility (an estimate of -6.46 and a t-statistic of -17.84 in the third specification). Furthermore, the parameter $\phi = 0.49$ shows that the effect of commuting costs is not linear but decreasing. Students who live sufficiently far and go on residence save 49% on the commuting costs (to be traded off against their fixed renting costs). The effect of the commuting costs differs across individuals in some respects, for example pupils from a catholic high school are somewhat less cost sensitive. Pupils from a technical high school with a social orientation appear to be more cost sensitive than others. Overall, however, mobility costs do not show much significant variation across individuals.

Commuting costs thus have a highly significant effect, but do they play a quantitatively important role in the pupils' participation decision? Or are they more relevant for the schooling decision, i.e. the decision where and what to study? Since the commuting costs enter the utility of the no-study option and the study options with a common parameter α_i, it would appear that they may have a similar effect on both the participation and the schooling decisions. However, this is only the case in the logit model. The nested logit models showed that pupils value the various study options as close substitutes for each other, relative to the no-study option. Our estimate of $\rho = 0.95$ implies that pupils are actually up to $1/(1-0.95) = 20$ times more responsive to commuting costs in their schooling decision than in their participation decision (see (1)). An increase in the mobility costs of one of the study options would thus generate substantial shifts in demand to other study options. But an increase in the mobility costs of the no-study option would have much smaller effects. In this sense, the pupils' mobility is a relative matter. Most pupils choose a study option close to their homes because they have a lot of study options in their neighborhood and they view these as close substitutes to more distant alternatives. But those pupils who do not have nearby access to any study option, would be willing to travel high amounts. These findings will be confirmed in our subsequent analysis, where we report the cost elasticities implied

[22] The parameter is significantly different from one at the 5% level, so that perfect correlation can be rejected.

by our estimates.

Determinants of schooling (γ_j)

Table 5 shows how individuals differ in their valuations of the various study options. The base study option is a non-catholic institution offering a short-term vocational program, in the field of arts. To obtain an idea of the quantitative importance of the parameter estimates, one may compute the additional willingness to pay relative to the base study option in monetary terms (in € 10,000's), by simply dividing the coefficients by the marginal utility of income α_i. Table 5 reveals several interesting findings.

The first column shows how individuals value catholic institutions of higher education. Few variables play a role, but the one exception is the religious orientation of the pupil's former high school. The coefficient of 1.05 is highly significant and it is also quantitatively important. It amounts to an additional willingness to pay for a catholic institution by pupils from a catholic high school of € 10,000 $\times$ 1.05 / (6.46-0.45) = € 1,750. This indicates that there are still strong linkages between the religious networks.

The next two columns show the valuations for the program type, i.e. academic or long term-term vocational programs (relative to short-term vocational). Interestingly, males and foreigners have a higher valuation for academic or long-term vocational programs. Hence, while they have a lower utility from participation (as we saw before), they do have a stronger preference for the long-term programs conditional on participating. Pupils who experienced years of repetition during high school have a lower utility from participating in the long-term programs, whether vocational or academic. The type of high school plays a significant role: pupils with the intellectually more demanding general high school background are not only more likely to participate, but they also choose the more demanding vocational long and especially academic program types. Their additional willingness to pay for academic programs at universities than for short-term vocational programs at colleges amounts to €10,000 $\times$ 2.71 / (6.46-0.36) = € 4,440. Furthermore, while we earlier found that the specific discipline taken at a general high school does not matter for participation, it does matter for the type of higher education program. Pupils with a general high school background in science, mathematics and classical languages have a substantially higher valuation for the academic or long-term vocational programs than for the short-term programs. Most notably, pupils who took classical languages would be willing to pay an additional € 10,000 $\times$ 1.88 / (6.46-0.53) = € 3,170 to follow an academic rather than a short-term vocational program, relative to comparable other general high school pupils.

The final three columns show the valuations for the specific program fields. The type of high school and the specific discipline followed at high school play the quantitatively most important roles. Generally speaking, pupils prefer the program fields that closely match the discipline they followed at high school. For example, pupils who followed science or mathematics at a general high school prefer sciences, whereas pupils who followed classical languages prefer arts. These findings stress the central importance of the high school background in the subsequent higher education decision, in contrast to some claims that the general high school leaves the options open for all study options at the higher education level.

It is interesting to point out that the gender effect also comes out strong in explaining differences in valuations across the study fields. Females have a strong preference for arts and especially biomedical sciences, and the weakest preference for exact sciences, as compared to males. While this seems to simply confirm common wisdom, it is important to stress that these gender effects are found even after having controlled for gender differences in the high school education background. For example, females have a lower willingness to pay for exact sciences than arts of € 2,400 relative to males, even if they both have the same science high school background.

Aggregation (σ)

Finally, consider the aggregation parameter σ, which is interacted with the log of the number of variants K_j available at aggregate study option j. It measures the degree to which preferences are correlated across the K_j variants over which we aggregated. In the two aggregate models, $\sigma < \rho$, which is inconsistent with the restrictions of the nested logit model.[23] In the third, "disaggregate" model, $\sigma = 0.955 > \rho$, but it is still quite close to ρ. This means that preferences only show weak additional correlation over the variants available at each study option. This finding suggests the need for even more detailed disaggregate analysis, to study educational choices at even more disaggregate levels. We leave this as a topic for further research.

4.2 Cost elasticities

Many of our empirical findings can be summarized by the own- and cross-cost elasticities of demand. We focus on the semi-elasticities, defined as the percentage change in the number of students in response to an absolute increase in the monetary costs x_{ij} of a given study option or subset of study options. We consider here an absolute cost increase by € 1,000 (or equivalently, given our definition of x_{ij}, an increase in the daily commuting distance by 9km

[23] T-statistics also showed that the difference is significant.

for a pupil traveling at a speed of 60km/h).[24]

The estimated cost elasticities provide information as to how participation and schooling would change in response to uniform or differentiated tuition fee increases. They may also give a first impression on the possible effects of a rationalization of supply, which is essentially a very large cost increase for a subset of study options. Finally, the cost elasticities are informative in interpreting the results of a more complete welfare analysis of reform.

We compute the cost elasticities at three different levels: the level of the market, the program type (colleges versus universities), and the program fields.

Uniform cost increase

Table 6 considers the effects of a uniform cost increase by € 1,000 on colleges (first column), universities (second column), and overall participation (third column). The first panel compares the elasticities as implied by the aggregate logit, the aggregate nested logit and the disaggregate nested logit. The results differ dramatically. While the logit model would predict overall participation to drop by a substantial 13.79%, the aggregate nested logit model predicts a drop of only 1.62%, and the disaggregate nested logit model a drop of an even lower 0.91%. These large differences across models follow from our finding that pupils have quite strongly correlated preferences across study options, i.e. they perceive the various study options as close substitutes ($\rho = 0.95$). Hence, while mobility costs matter in their schooling decisions, they play only a limited role in their participation decisions. In all models the relative drop in students is larger at colleges than at universities, e.g. -1.09% versus -0.51% in the disaggregate nested logit model; the absolute drop at colleges is even higher since colleges have a larger market share.

The second panel shows how the elasticities differ among pupils. Males and foreigners are more likely to drop out than others. For example, male students would drop their overall participation by 1.07% compared to a drop of 0.78% for female students.[25] Students with no repetitions are less likely to drop out than students with one or two years of repetition (-0.65% versus -1.41% and -2.14%). The most important differences are found between students from different high school background: students with a general high

[24] More formally, denote the predicted probability that individual i chooses study option j by $\widehat{P}_{ij}(x_i)$, where $x_i = (x_{i1} \cdots x_{iJ})$ is the $J \times 1$ vector of individual i's monetary costs for the various study options. Let δ be a $J \times 1$ vector of ones and zeros, where the ones denote the study options for which there is a cost increase by € 1000. The semi-elasticity of demand for all study options $j \in A$ with respect to a cost increase 1000δ is then defined as $\sum_i \sum_{j \in A} \left(\widehat{P}_{ij}(x_i) - \widehat{P}_{ij}(x_i + 1000\delta) \right) / \sum_i \sum_{j \in A} \widehat{P}_{ij}(x_i)$. For example, suppose that the set A consists of all study options and all elements in δ are equal to 1. The semi-elasticity then refers to the percentage change in the total number of students in response to a uniform cost increase by € 1000.

[25] The ratio of these changes is similar to the odds ratio, which is equal to $\exp(\gamma_0^l)$ for the l-th individual characteristic.

school background would reduce participation by only 0.5%, whereas students from technical and professional high schools would drop participation by, respectively, 1.35% and 2.58%. In sum, the empirical results show that the overall demand for higher education is highly inelastic, although there are some clear differences between individuals.

Cost increases by program type: colleges versus universities

Additional insights are obtained by considering the cost elasticities at lower levels of aggregation. Table 7 presents the semi-elasticities at the level of the program type: vocational programs at colleges versus academic programs at universities. Cost increases by colleges only or by universities only evidently have even smaller effects on total participation (last column). However, underlying these small total effects, there are large shifts in demand. A € 1,000 cost increase to all colleges reduces college demand by almost 13%, and a € 1,000 cost increase to all universities reduces university demand by an even larger 24%. These findings are consistent with our earlier discussion on the role of mobility costs. Since pupils have strongly correlated preferences across study options, they are quite willing to substitute between colleges and universities in response to a differentiated cost increase, even though they are unlikely to refrain from participation altogether.

We can use these results to assess the government's historic efforts to promote participation by investing in a large college network with a broad geographic coverage. Our estimated cost elasticities suggest that these efforts only have a negligible effect on total participation. The investment efforts thus mainly lead to a substitution from universities to colleges. In this sense, we may conclude that the government's policy has essentially not lead to more democratization, but rather to a diversion away from universities. This relates to Rouse's (1995) findings on the impact of U.S. community colleges on educational attainment. She also found that colleges did not increase the likelihood of attendance (though they may have led to an increase in the number of years of schooling). However, even if the promotion of colleges did not have an effect on democratization, this does not mean it is undesirable. The diversion to colleges may be efficient or inefficient depending on the benefits to pupils and the different costs of supplying education at colleges and universities. We will turn to that question in our welfare analysis below.

Cost increases by program field

Cost elasticities at the lower level of the program field provide interesting additional information. For example, they are relevant in assessing regulatory policies to promote certain study programs. They are also of interest in assessing the effects of introducing differentiated tuition fees. To our knowledge, there are no previous estimates of elasticities in higher education at the level of the program field.

Table 8 shows the estimated semi-elasticities for eight disciplines (arts, social sciences, biomedical sciences, and exact sciences; of either the vocational or the academic type). The own-cost elasticities on the diagonals show that pupils are quite cost-sensitive for all disciplines. The elasticities tend to be lower for the disciplines with the higher market share. For example, vocational biomedical sciences have by far the highest market share and also the lowest own-cost elasticity. However, market share is not the only relevant factor. For example, academic social sciences and vocational exact sciences have a similar market share, but pupils are much more cost sensitive towards the former.

The cross-cost elasticities reveal several additional interesting patterns. It is best to read the cross-elasticities by row. Notice first that the cross-cost elasticities would be the same for all fields on the same row if all individuals were identical (no observed or unobserved heterogeneity). This follows from the IIA property of the logit model, saying that identical pupils shift proportionally to other alternatives. We earlier found however that pupil heterogeneity does matter in explaining educational choices. This is indeed reflected in the pattern of cross-cost elasticities, which vary widely within each row of the table. For example, the last row shows that a rise in the costs of academic exact sciences would generate much more substitution to other academic sciences than to vocational sciences; and among the vocational sciences the gains would mainly go to the vocational exact sciences.

This example illustrates a more general pattern for the cross-cost elasticities: pupils tend to mainly substitute within the academic or within the vocational program types, and to the extent that they substitute across types they would especially choose the "twin" program field of the other type. There is thus generally a dominance of the program type dimension over the program field dimension. In a few cases this dominance is somewhat weak: vocational exact sciences loose a comparatively high amount to academic sciences (fourth row), and academic arts loose relatively much to vocational arts (fifth row). In one case the dominance of the program type over the program field is actually reversed: vocational arts loose more to academic arts than to any other vocational program (first row). This is consistent with the high quality reputation of the vocational arts programs relative to academic arts (e.g. the language interpreter programs).

As a final remark, note that exact sciences and biomedical sciences are the closest neighbors, and so are social sciences and arts: in almost all cases, substitution mainly occurs to these neighbors. The only exception occurs on the second and third columns, where it appears that vocational biomedical and vocational social sciences are closest neighbors. This can be explained by the nature of many of these vocational programs (e.g. nursing).

Concluding remarks

The estimated elasticities show that uniform cost increases have little overall effects on the demand for higher education, though they can somewhat change the composition of demand. In contrast, cost increases by program type and program field lead to large shifts in the composition. One may use these findings to draw some tentative conclusions on policy reform. On the one hand, the low market-level elasticities suggest that a uniform tuition fee may generate large distributional effects, and comparatively lower total welfare improvements. On the other hand, the large elasticities at the level of the program type and program field, suggest that differentiated tuition fee increases may involve additional distributional and welfare effects. To obtain more insights, we turn to an illustrative welfare analysis next.

4.3 Implications for policy reform

To illustrate how our empirical results can be useful in assessing policy reform, we focus on the effects of raising tuition fees within a centralized public system, i.e. keeping other things such as the quality and the diversity of supply constant. Such an analysis is highly relevant since, as we have seen in section 2, fees are currently far from sufficient to cover even the variable cost per student, let alone the fixed costs. An analysis of more drastic reforms, such as decentralizing decision making to the institutions, is beyond the scope of this paper, since it would require a better empirical understanding on how the institutions of higher education compete.

Framework

We make the following assumptions.

- Pupils can borrow at a competitive interest rate to finance their educational expenses. This ensures that increases in tuition fees do not cause pupils to drop out because of capital constraints. Barr (2004) and others have discussed how this can be accomplished through appropriately designed income-contingent student loans.

- Pupils can deduct their educational expenses from their (future) taxable income. This ensures that a progressive income tax system does not distort the incentives to invest in higher education; see Bovenberg and Jacobs (2005) and Jacobs and Vanderploeg (2005).[26]

[26] The result that educational expenses should be tax deductable, or equivalently that the educational subsidy rate should be equal to the marginal income tax, closely relates to Diamond and Mirlees' (1971) result that the optimal tax rate on intermediate goods is zero.

- The private (Mincer) returns to higher education are equal to the social returns, i.e. there are no spillovers to others from investing in higher education. Based on a discussion of the available micro and macro empirical evidence, Jacobs and van der Ploeg (2005) conclude that this is a realistic assumption. It allows us to directly use the estimates of our random utility discrete choice model to compute the changes in consumer surplus after a change in tuition fees.

- The government can regulate the colleges and universities, so that they do not change the quality or diversity of supply in response to an increase in tuition fees. Variable subsidies are granted on a cost basis. This allows us to treat the government and the higher education institutions as an integrated entity, and use the government's net revenues (tuition fee revenues minus the subsidy costs) as a measure of producer surplus.

- Progressive income taxes ensure the socially desirable income distribution. Atkinson and Stiglitz (1976) provide conditions under which this is the case. This allows us to abstract from equity considerations.

Based on these assumptions we can assess the welfare effects of tuition fee increases in terms of consumer and producer surplus. Write the indirect utility $V_{ij}^* = V_{ij} + (1-\sigma)\ln(K_j)$, where V_{ij} is given by (4), as a function of the $J \times 1$ tuition fee vector $t = (t_1 \cdots t_J)$, i.e. $V_{ij}^* = V_{ij}^*(t)$ for $j \neq 0$. The nested logit model gives the following expression for pupil i's expected surplus as a function of t (see e.g. McFadden (1981)):

$$CS_i(t) = \frac{1}{\alpha_i}\left(\sum_{j=1}^{J} \exp(V_{ij}^*(t)/(1-\rho))\right)^{(1-\rho)} + \exp\left(V_{i0}\right).$$

The average consumer surplus per pupil is $CS = \sum_{i=1}^{I} CS_i/I$, where I is the number of pupils. The variable part of producer surplus per pupil (or government's variable net revenue) for a specific study option j, as a function of t, is equal to

$$PS_j(t) = (t_j - c_j)\sum_i P_{ij}(t),$$

where $P_{ij}(t)$ is the probability that individual i chooses program j as a function of t, and c_j is the constant variable cost per student of program j. Total producer surplus per pupil is $PS(t) = \sum_{j=1}^{J} PS_j(t)/I$.

Direct estimates of the variable cost per student c_j are not available. As an indirect measure we use the government's variable subsidy per student, since these are granted on a cost basis as discussed in section 2. Table 2 summarized how the variable subsidies, and hence the estimated variable costs, tend to be higher for the biomedical and exact sciences, especially at the universities. We will present the results from the estimated variable costs of the second panel of Table 2, but the results are very similar when the alternative estimates in the first panel were used.

The most simple welfare analysis would simply look at the sum of consumer and producer surplus. However, this would ignore that the social costs of public funds λ may be greater than zero, i.e. that the government may need to levy distortionary taxes elsewhere to finance a higher education deficit. Estimates of the social costs of public funds vary widely, from € 0.17 to over € 1.65 per euro of public funds raised as discussed in Bird (2005). Diewert, Lawrence and Thompson (1998) suggest using a number of at least € 0.23. We take a conservative approach and compute total welfare as $CS(t) + (1 + \lambda)PS(t)$, where λ is either equal to 0, or equal to Diewert et al.' s 0.23.

Findings

Table 9 shows the effects of both uniform and differentiated tuition fee increases. The first two columns show the size of the considered fee increases. The next four columns show the effects on per pupil consumer surplus, net revenues and total welfare assuming the cost of public funds λ is either zero or € 0.23. The final two columns show the associated effects on participation.

The first row shows the effects of a "small" uniform tuition fee increase of € 1,000. This is small in the sense that it is still insufficient to cover the variable costs per student (see Table 2). Consumer surplus drops by € 657 per pupil.[27] Producer surplus increases by an amount of € 670 per pupil. Hence, there are large distributional effects from students to the government (tax payers). The total welfare increase depends on how one values the increase in producer surplus: it increases by a small € 13 per pupil if the social costs of public funds are zero ($\lambda = 1$), but by a much higher amount of € 167 if $\lambda = 0.23$.

The second row compares the € 1,000 uniform increase with a differentiated fee increase that yields an equivalent increase in producer surplus: we take a fee increase of € 750 at colleges, which requires a fee increase of € 1,579 at universities to keep producer surplus constant. This differentiated fee increase implies that welfare increases by an additional € 12 per pupil relative to the uniform fee increase. This increase is due to a shift in demand from the universities to the colleges, which operate at a lower variable cost (Table 2). This shows that the earlier discussed diversion effects to colleges are not necessarily bad from a total welfare point of view. Nevertheless, it is also striking that differentiating fees between

[27] This is roughly proportional to the fraction of eligible pupils that choose to study, as expected from a discrete choice model.

colleges and universities only improves total welfare by a small amount: we attribute this to the fact that the shift to the colleges does not only imply a lower variable cost of supply, but also a lower benefit to the pupils.

To gain further insights we subsequently consider the effects of more drastic cost-based fee increases, i.e. fee increases that are sufficient to exactly cover the variable part of producer surplus. The third row of Table 9 considers the effect of a uniform cost-based fee increase, which amounts to a required fee increase by € 2,810. Since the variable costs per student vary across program types and fields, such a fee increase implies some cross-subsidization from colleges to universities, and from arts and social sciences to biomedical sciences and exact sciences. By construction, (variable) producer surplus becomes zero, an increase by € 1,852 per pupil relative to the status quo. There is therefore a large shift in distribution from students to producers (the government). Total welfare increases by the small amount of € 22 per pupil if $\lambda = 0$, and by the much larger amount of € 448 if $\lambda = 0.23$.

The fourth row shows how welfare further improves after a differentiated cost-based fee increase, i.e. such that the variable costs of each individual program are covered (implying no longer a cross-subsidization). If $\lambda = 0$, these tuition fees are also the first-best levels. Total welfare now increases by an additional € 100 relative to the uniform cost-based fee increase. This shows that fee differentiation has some modest effect on total welfare, but the effect should not be exaggerated.

The final row of Table 9 considers the welfare effects of adding a uniform markup over the cost-based levels. To illustrate, we consider a uniform markup of € 5,000, which is roughly sufficient to cover the fixed costs of higher education (in addition to the variable costs). If $\lambda = 0$, such a uniform markup lowers welfare relative to the first-best cost-based fees (by about € 80 per pupil), but it still raises welfare relative to the status quo (by the small amount of € 36 per pupil). In contrast, if $\lambda = 0.23$ total welfare per pupil further increases by € 1,150 relative to the status quo.[28] Note that we also considered the welfare effects from introducing non-uniform markups over marginal costs (i.e. Ramsey pricing), to exploit differences in the elasticities; we found that the additional gains are negligible, which is due to the fact that the estimated program-level cost elasticities are quite large and similar to each other (see Table 8).

This discussion focused on the effect of tuition fee increases on total welfare. To gain additional intuition, the last two columns of Table 9 show how participation changes in response to the tuition fee increases. Consistent with our earlier discussed cost elasticities,

[28] We also computed the optimal uniform markup under $\lambda = 0.23$. This amount to a markup over costs of € 15,200. This very high number is due to the very low cost elasticity with respect to the overall participation decision.

overall participation generally does not drop by very much for the small fee increases (drop by less than 1%) but also not for the more drastic cost-based fee increases (drop of slightly more than 2%). Only if fees increase to have a € 5,000 markup, participation drops considerably (−7%). As a final point, the differentiated fee increases are accompanied by substantial student shifts from the universities to the colleges. This again emphasizes that a diversion of students from university to colleges is not necessarily harmful from a total welfare perspective, due to the associated variable cost savings.

Concluding remarks

We can summarize this discussion as follows. First, uniform cost-based tuition fee increases achieve most of the welfare gains; the additional gains from fee differentiation are relatively unimportant. The welfare increases are relatively large even under conservative assumptions on the social cost of public funds. If one ignores the social cost of public funds, the welfare gains are relatively small, but there is still a substantial redistribution from students to outsiders.

5 Conclusions

We have analyzed the determinants of participation and schooling in a public system of higher education, using a unique data set on pupils' study choices. One of our central findings is that pupils perceive the available institutions and programs as close substitutes, implying an ambiguous role for travel costs: they hardly affect the participation decisions, but have a strong impact on the schooling decisions. In addition, high school background plays an important role in both participation and schooling. Our empirical analysis generalizes previous work, which has focused on the participation rather than the schooling decision (where and what to study). Based on information of travel costs, we can indirectly infer the effects from raising costs including tuition fees at a high level of detail.

Our empirical results can contribute to informing the debate on reforming public systems of higher education. As an illustration, we have assessed the effects of tuition fee increases. Uniform cost-based tuition fee increases achieve most of the welfare gains. The additional gains from fee differentiation are relatively unimportant. The welfare gains are quite large if one makes conservative assumptions on the social cost of public funds, and there is a substantial redistribution from students to outsiders. Our empirical framework may be used to assess the effects of additional policy reforms of public systems. For example, in several countries governments aim to rationalize the supply by reducing the number of institutions through associations and/or reducing the number of duplicated programs. In future research, it would be of strong interest to evaluate the efficiency and distributional effects of these and other, more drastic reforms, such as decentralizing decision-making to the universities and colleges.

6 References

Atkinson, A.B. and Stiglitz, J.E. 1976. "The design of tax structure: Direct versus indirect taxation" Journal of Public Economics 6(1-2): 55-75.

Barr, N. 2004. "Higher education funding" Oxford Review of Economic Policy 20(2): 264-283.

Belgisch Staatsblad. 2006. "Koninklijk besluit tot wijziging van het KB/WIB 92, op het stuk van de voordelen van alle aard" 2006-02-17.

Ben-Akiva, M. E. and S. R. Lerman. 1985. "Discrete Choice Analysis: Theory and Application to Travel Demand". Cambridge, Ma.: MIT Press.

Bird, R.M. 2005. "Evaluating public expenditures: Does it matter how they are financed?" Working Paper 0506. International Tax Program, Institute for International Business, Joseph L. Rotman School of Management, University of Toronto

Bovenberg, A.L. and B. Jacobs. 2005. "Redistribution and education subsidies are Siamese twins". Journal of Public Economics 89: 2005-2035.

Cameron, S.V. and J.J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic and White Males". Journal of Political Economy 109(3): 455-499.

Cantillon, B.; G. Verbist and I. Segal. 2005. "Student in de 21ste Eeuw. Studiefinanciering voor het hoger onderwijs in Vlaanderen." Report for Onderwijskundig Beleids- en Praktijkgericht Wetenschappelijk Onderzoek (OBPWO), Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs.

Deen, J.; B. Jongbloed and H. Vossensteyn. 2005. "Bekostigingstarieven in het hoger onderwijs. Een vergelijking tussen 7 landen." Studie i.o.v. Ministerie van de Vlaamse Gemeenschap. Center for Higher Education and Policy Studies (CHEPS), Universiteit Twente.

Diamond, P.E. and J.A. Mirrlees. 1971. "Optimal taxation and public production: I. Production efficiency." American Economic Review 61: 8-27.

Diewert,W.E. , D.A. Lawrence and F. Thompson. 1998. "The Marginal Cost of Taxation and Regulation". In *Handbook of Public Finance*, ed. F. Thompson and M.T. Green, 135-71. New York: Marcel Dekker.

Dynarski, S.M. 2003. "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion" American Economic Review 93(1): 279-288.

European Commission. 2003. DG Education and Culture. Communication from the Commission "Education & Training 2010". The Success of the Lisbon Strategy Hinges on Urgent Reforms.

European Commission. 2004. DG Education and Culture. Implementation of "Education & Training 2010" Work Programme. Working Group E "Making The Best Use Of Resources". Progress Report.

Eurydice 2000. "Two decades of reform in higher education in Europe: 1980 onwards" Eurydice, the Information Network on Education in Europe, http://www.eurydice.org

Frenette, M. 2003. "Access to College and University : Does distance matter?" Report No.11F0019 201. Business and Labour Market Analysis, Statistics Canada.

Jacobs, B. and F. van der Ploeg. 2005. "Guide to Reform of Higher Education: A European Perspective" CEPR Discussion Paper 5327.

Jobdienst KU Leuven. 2006. "Overeenkomst voor tewerkstelling van studenten" http://www.kuleuven.be/jobstudent/contract.htm [24/04/2006].

Kane, T.J. 1995. "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College?" NBER Working Paper 5164.

Long, B. T. 2004. "How have college decisions changed over time? An application of the conditional logistic choice model," Journal of Econometrics 121 (1-2): 271-296.

Maassen, P. 2000. "Models of Financing Higher Eduation in Europe" A Background Paper Commissioned by the Nordic Council of Ministers for the Conference on “Financing Higher Education”.Held in Reykjavík 3-4 April 2000. Center for Higher Education and Policy Studies (CHEPS), Universiteit Twente.

McFadden, D. 1978. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models*. Edited by L. Lundqvist F. Snickars and J. Weibul A.Karlqvist. New York: North-Holland.

McFadden, D. 1981 "Econometric Models of Probabilistic Choice". In *Structural Analysis of Discrete Data and Econometric Applications* (eds. Manski, C.F. and McFadden, D.) Cambridge, Ma.: MIT Press.

McFadden, D. and K. Train. 2000. "Mixed MNL Models for Discrete Response" Journal of Applied Econometrics 15(5): 447-470.

Murphy, K.M. and R.H. Topel. 1985. "Estimation and Inference in Two-Step Econometric Models" Journal of Business & Economic Statistics 3(4): 370-379.

OECD. 2004. "Education at a Glance: OECD Indicators 2004".

Rouse, C. E. 1995. Democratization or diversion? The effect of community colleges on educational attainment. Journal of Business Economics and Statistics, 13(2), 217-224.

Train, K. 2003. Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press.

Train, K. and D. McFadden. 1978. "The goods/leisure trade-off and disaggregate work trip mode choice models," Transportation Research 12: 349-353.

Usher, A. & Cervenan, A. 2005. Global Higher Education Rankings 2005. Toronto, ON: Educational Policy Institute.

Vandenbroucke, F. 2005. "De financiering van het hoger onderwijs: voorstel aan de werkgroep financiering. Vervolgnota en simulaties." Vlaams Ministerie van Onderwijs (http://www.ond.vlaanderen.be/hogeronderwijs/Vervolgnota9-12-05.htm).

van Heffen, O. and Lub, A. Higher Education in Flanders. Country Report. 2003. CHEPS - Higher Education Monitor.

7 Appendix. The data

We combine two main data sets: “pupils” and “students”. Both are made available by the Flemish Ministry of Education. In addition to these data sets, we constructed a number of auxiliary data sets, describing additional choice and/or demographic characteristics.

1. Pupils

   The pupils data set contains information on all 55,905 pupils who attended the last year of secondary school in the year 2001. For each pupil there is information on five variables, defining the pupil's *profile*: secondary school institution, study program during last year of secondary school, postal code, age (birth year), sex and nationality.

2. Students

   The students data set contains information on all students who first registered for a higher education program in either 2001 or in 2002. Information is available on each student's profile, according to the same five variables as in the pupils data set. In addition, the data set contains each student's choice of higher education institution, campus and study program. Finally, there is information on the year of graduation from secondary school. We use this last variable to extract the subset of students who graduated from a secondary school in the year 2001. This amounts to a total number of 35,562 first registering students, out of which the large majority (34,395 students) immediately registered after secondary school graduation in 2001, and a small group registered with one year of delay in 2002 (1,167 students). Hence, it is reasonable to assume that the fraction of 2001 pupils that first registers to a higher education institution after 2002 is negligible.

Our task is to distinguish between pupils who become students and pupils who remain "outsiders", i.e. who do not register at a higher education institution in 2001 or 2002. In principle, this can be done by combining the pupils and the students data sets, based on their profiles as defined by the five common variables: a pupil for which there is a successful match with a student can be identified as a student, while a pupil for which there is no successful match with a student can be identified as an outsider. In practice, we also found a small number of students for which there is no successful match with a pupil. Based on correspondence with the ministry of education, we attribute this to some inconsistencies in the definition of the pupils' profiles rather than in the definition of the students' profiles. We therefore adopt the following approach when matching the pupils and students data sets. We identify as students all individuals in the students data set (even if no successful match with a pupil was found). We then identify as outsiders all pupils for which there is no successful match with a student. This will generate a data set with slightly too many individuals that are identified as outsiders: this excess number is equal to the number of students for which no successful match with a pupil was found. We then randomly drop this excess number of individuals from the outsiders. Our combined data set then contains information on 55,905 pupils: 35,562 students and 20,343 outsiders.

In addition to the two main data sets, we constructed a number of auxiliary data sets. First, we have information on various characteristics of the secondary school institution and the study program during the last year of high school. The information on the secondary school institution includes the postal code of the secondary school and the network affiliation (free subsidized, official subsidized, or community). The information on the study program during the last year of secondary school includes the main category (i.e. general, technical, arts or professional high school program), as well as more specific information (focus on languages, math, science, etc.). Second, we have information on the analogue characteristics of the higher education institution and study program. Third, we have information on the distribution of several demographic variables: average income by postal code, and average commuting distance and commuting time of the active labor force by postal code. Fourth, we make use of Microsoft's route planning software to compute the distance and car travel time between each individual's postal code address and each higher education institution's postal code address. The distance and travel time between an individual's postal code and the no-study alternative is set equal to the average distance travelled to work, by postal code. Finally, we use information provided by the Belgian railroad company to compute train travel times, i.e. train travel time between each individual's closest train station and each higher education institution's closest train station, plus car travel time to and from the respective stations.

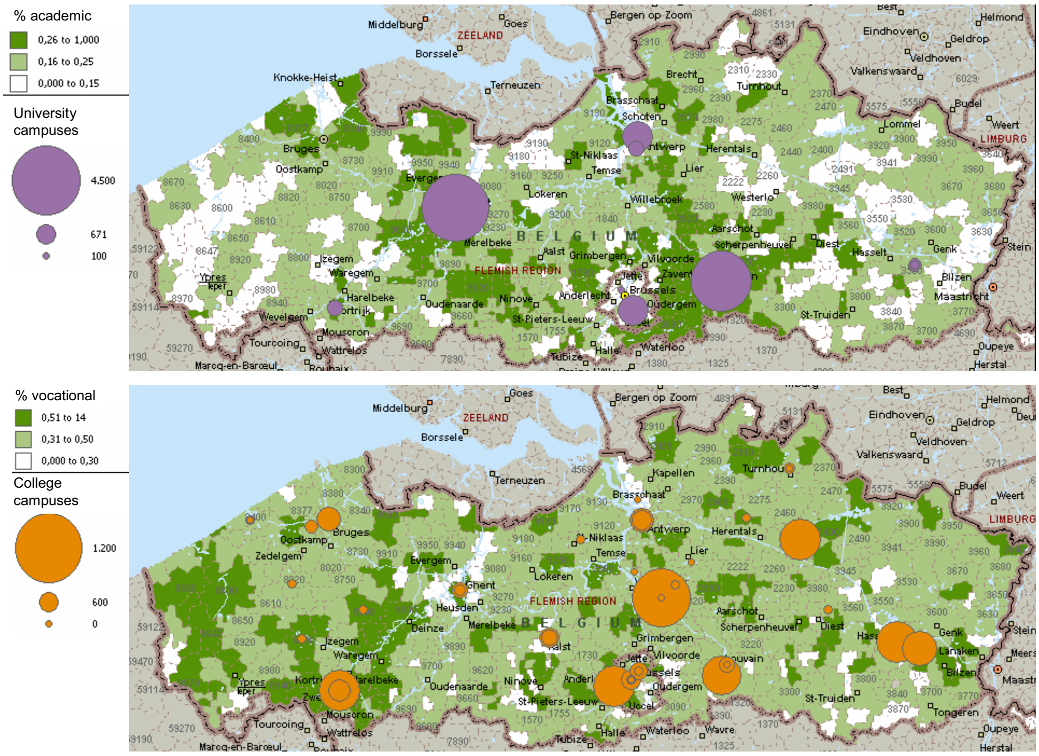

Figure 1: Participation in higher education, by postal code

Table 1: Supply of Higher Education in Flanders (2001)

	Colleges (vocational education)		Universities (academic education)	
	Non-catholic	Catholic	Non-catholic	Catholic
Number of campuses				
Total	16	28	5	4
offering degrees in				
Arts	5	7	2	4
Social Sciences	12	21	3	4
Biomedical Sciences	10	17	5	2
Exact Sciences	14	15	5	2
Number of students				
Total	9,899	15,283	6,658	5,641
enrolled for a degree in				
Arts	1,016	791	943	933
Social Sciences	5,538	9,952	2,996	3,116
Biomedical Sciences	839	1,709	1,312	798
Exact Sciences	2,506	2,831	1,407	794

Note: Own calculations based on our dataset from the Flemish Ministry of Education.

Table 2: Variable subsidies per student in Euros

	Colleges	Universities
2001 subsidy scheme		
Total	2,973	3,891
Arts	2,456	2,593
Social Sciences	2,772	2,601
Biomedical Sciences	3,671	5,186
Exact Sciences	2,994	5,186
2005 subsidy scheme		
Total	3,113	3,937
Arts	2,456	2,652
Social Sciences	2,865	2,807
Biomedical Sciences	3,683	5,000
Exact Sciences	3,448	5,290

Note: The top panel is based on the variable subsidy scheme for higher education in vigour until 2001 (Universiteitendecreet, 1991; Hogescholendecreet, 1994). The bottom panel is based on the proposed new subsidy scheme (Vandenbroucke, 2005). We report student-weighted averages of subsidies per study field for both colleges and universities.

Table 3: Summary statistics of 2001 eligible pupils

	All Pupils	Students	Outsiders	College	University	Non-catholic	Catholic
Demographic							
male	0.48	0.45	0.56	0.45	0.45	0.47	0.43
foreign	0.02	0.01	0.04	0.01	0.01	0.02	0.01
catholic high school	0.75	0.78	0.70	0.79	0.76	0.67	0.87
Ability							
years of repetition	0.61	0.36	1.11	0.46	0.16	0.40	0.34
	(1.06)	(0.95)	(1.09)	(0.99)	(0.83)	(1.05)	(0.87)
general high school	0.48	0.60	0.21	0.44	0.94	0.63	0.58
classical languages	0.11	0.14	0.04	0.05	0.33	0.15	0.13
modern languages	0.19	0.24	0.09	0.22	0.27	0.23	0.24
economics	0.15	0.19	0.07	0.19	0.17	0.17	0.20
sciences	0.16	0.20	0.07	0.11	0.40	0.24	0.18
mathematics	0.23	0.30	0.09	0.15	0.60	0.34	0.27
technical high school	0.35	0.33	0.39	0.47	0.04	0.29	0.35
´product'-focused	0.14	0.12	0.19	0.17	0.02	0.11	0.12
Mobility							
Distance (kms) by road to campus	n/a	34.71	n/a	30.96	42.38	35.73	33.90
	n/a	(28.17)	n/a	(25.65)	(31.37)	(28.19)	(28.13)
Time (mins) by road to campus	n/a	30.74	n/a	28.33	35.67	32.13	29.64
	n/a	(17.33)	n/a	(16.2)	(18.47)	(17.59)	(17.03)
Travel cost to campus (x10,000€)	n/a	0.38	n/a	0.35	0.46	0.40	0.37
	n/a	(0.28)	n/a	(0.25)	(0.31)	(0.28)	(0.28)
Time (mins) by train to campus	n/a	47.60	n/a	44.33	54.30	47.62	47.59
	n/a	(28.04)	n/a	(26.19)	(30.41)	(27.83)	(28.20)
Number of observations	55,905	37,481	18,424	25,182	12,299	16,557	20,924

Note: Standard errors for the continuous variables are in parentheses. Demographic and ability data are based on our data set from the Flemish Ministry of Education; mobility statistics are based on own calculations using postal code information.

Table 4: Participation and schooling decisions - Comparison of alternative models

	Aggregate logit model		Aggregate nested logit model		Disaggregate nested logit model	
Parameter	Estimate	t	Estimate	t	Estimate	t
Outside option (γ_0)						
intercept	4.10*	(26.52)	1.51*	(12.39)	1.18*	(14.60)
male	0.22*	(3.85)	0.27*	(5.61)	0.31*	(7.55)
foreign	1.17*	(6.23)	1.00*	(6.15)	0.77*	(6.59)
catholic high school	0.20*	(3.28)	-0.33*	(-6.20)	-0.35*	(-7.90)
years of repetition	0.47*	(13.04)	0.45*	(15.57)	0.47*	(20.61)
general high school[1]	-1.81*	(-13.70)	-2.19*	(-20.44)	-2.14*	(-21.00)
classical languages	0.82*	(6.21)	-0.34*	(-3.01)	-0.14	(-1.55)
modern languages	0.03	(0.24)	0.10	(1.02)	0.05	(0.63)
economics	-0.36*	(-2.94)	-0.17	(-1.62)	0.08	(0.89)
sciences	0.29	(2.46)	-0.07	(-0.67)	0.02	(0.17)
mathematics	0.34*	(3.08)	-0.08	(-0.80)	0.00	(-0.05)
technical high school[1]	-1.73*	(-17.41)	-1.38*	(-18.55)	-1.19*	(-16.88)
´product'-focused	0.57*	(6.37)	0.21*	(2.94)	0.24*	(3.85)
Travel cost (α_i)						
intercept	5.44*	(18.14)	6.34*	(19.44)	6.46*	(17.84)
φ	-0.41*	(-2.20)	-0.42*	(-2.09)	-0.49*	(-2.12)
male	-0.10	(-0.86)	-0.06	(-0.53)	0.24	(1.83)
foreign	-0.63	(-1.24)	-0.22	(-0.38)	0.63	(1.30)
catholic high school	-0.30*	(-2.13)	-0.62*	(-3.96)	-0.45*	(-2.68)
years of repetition	-0.01	(-0.18)	0.04	(0.49)	0.07	(0.74)
general high school[1]	0.28	(1.02)	0.06	(0.21)	-0.36	(-1.09)
classical languages	-0.58*	(-2.93)	-0.75*	(-3.59)	-0.53*	(-2.36)
modern languages	-0.01	(-0.07)	0.07	(0.33)	0.42	(1.89)
economics	0.29	(1.38)	0.25	(1.11)	0.50*	(2.06)
sciences	-0.57*	(-2.89)	-0.62*	(-2.96)	0.01	(0.06)
mathematics	-0.20	(-1.06)	-0.17	(-0.84)	-0.17	(-0.77)
technical high school[1]	1.42*	(6.13)	1.67*	(6.58)	1.41*	(4.80)
´product'-focused	-1.05*	(-5.50)	-1.30*	(-6.32)	-1.41*	(-6.33)
Nesting parameters						
ρ	0 (not estimated)		0.898*	(33.90)	0.946*	(42.80)
σ	-0.048*	(2.01)	0.888*	(30.42)	0.950*	(46.60)
Slope parameters (γ_j, $j \neq 0$)						
Catholic Institution[2]	*included*		*included*		*included, see table 5*	
Academic program[3]	*included*		*included*		*included, see table 5*	
Vocational long program[3]	*not included*		*not included*		*included, see table 5*	
Study field characteristics[4]	*not included*		*not included*		*included, see table 5*	
Fixed effects (β_j)	*included*		*included*		*included, see table 5*	
Observations	778,464		778,464		2,981,735	
individuals	*14,416*		*14,416*		*19,237*	
alternatives	*54*		*54*		*155*	
Mean log likelihood	-0.0383068		-0.037747		-0.0032785	

Notes: t-statistics in parentheses. * statistical significance at 5% level

[1] base category = professional/arts high school

[2] base category = non-catholic study option

[3] base category = vocational study option (aggregate models), vocational short study option (disaggregate model)

[4] base category = arts study option

Table 5: Schooling decisions - Results from disaggregate nested logit model

	Religious orientation[1]	Type of Higher Education[2]		Study Field[3]		
Parameter ($\gamma_j, j \neq 0$)	Catholic	Vocational long	Academic	Social	Biomedical	Exact
intercept	-0.74*	-0.52*	-2.99*	-1.70*	-3.12*	-2.33*
	(-4.48)	(-2.12)	(-9.88)	(-6.13)	(-8.55)	(-8.59)
male	-0.03	0.52*	0.23*	0.31*	-0.62*	1.43*
	(-0.49)	(6.38)	(3.12)	(3.29)	(-5.41)	(13.85)
foreign	-0.16	0.50	1.07*	0.03	0.08	-0.40
	(-0.67)	(1.41)	(3.44)	(0.08)	(0.17)	(-0.95)
catholic high school	1.05*	0.19	0.15	0.08	0.20	0.33*
	(15.27)	(1.79)	(1.49)	(0.64)	(1.36)	(2.53)
years of repetition	-0.10*	-0.28*	-0.22*	-0.02	-0.14	-0.14
	(-2.50)	(-4.30)	(-3.68)	(-0.25)	(-1.56)	(-1.79)
general high school[4]	0.10	0.23	2.71*	2.40*	2.65*	0.71*
	(0.78)	(0.96)	(10.16)	(9.64)	(7.86)	(2.67)
classical languages	0.08	0.92*	1.88*	-0.52*	-0.48*	-0.81*
	(0.88)	(5.82)	(14.83)	(-3.83)	(-2.92)	(-5.12)
modern languages	-0.08	0.41*	-0.03	-0.53*	-0.73**	-0.65*
	(-0.94)	(3.03)	(-0.32)	(-4.08)	(-4.43)	(-4.07)
economics	0.11	0.84*	0.11	1.35*	0.67*	0.43*
	(1.19)	(5.85)	(0.96)	(8.90)	(3.48)	(2.36)
sciences	-0.06	0.89*	1.01*	0.57*	1.73*	1.39*
	(-0.67)	(6.34)	(8.80)	(3.98)	(10.37)	(8.87)
mathematics	-0.19*	1.60*	1.37*	0.91*	1.22*	1.98*
	(-2.13)	(12.09)	(12.81)	(7.02)	(7.53)	(12.61)
technical high school[4]	-0.14	-0.42	0.32	2.86*	3.45*	1.12*
	(-1.23)	(-1.87)	(1.17)	(10.84)	(10.13)	(4.26)
'product'-focused	-0.11	0.81*	-0.17	-0.09	1.31*	2.94*
	(-1.19)	(5.26)	(-0.91)	(-0.34)	(4.66)	(10.77)

Notes: t-statistics in parentheses. * statistical significance at 5% level

[1] base category = non-catholic study option

[2] base category = vocational short study option

[3] base category = arts study option

[4] base category = professional/arts high school

Table 6: Cost elasticities at market level

	Effect on[1]:		
	Colleges	Universities	Overall
All pupils			
Logit	-16.43	-0.91	-13.79
	(0.29)	(0.24)	(0.25)
Aggregate nested	-1.95	-0.87	-1.62
	(0.51)	(0.23)	(0.43)
Disaggregate nested	-1.09	-0.51	-0.91
	(0.45)	(0.21)	(0.38)
Pupils by profile (disaggregated nested logit model only)			
male	-1.28	-0.60	-1.07
female	-0.94	-0.44	-0.78
Belgian	-1.08	-0.50	-0.90
foreign	-2.11	-1.24	-1.80
no repetition	-0.78	-0.44	-0.65
1 year repetition	-1.51	-0.91	-1.41
2 years repetition	-2.20	-1.54	-2.14
catholic high school	-1.00	-0.47	-0.83
non-catholic high school	-1.49	-0.74	-1.29
general high school	-0.54	-0.46	-0.50
technical high school	-1.36	-1.33	-1.35
professional/arts high school	-2.58	-2.54	-2.58
Current market share	45.05	22.00	67.04

Notes: Standard errors in parentheses.
[1]Reported as semi-elasticities: % change in market share given a uniform cost increase of € 1,000.

Table 7: Cost elasticities at program type level

	Effect on[1]:		
	Colleges	Universities	Overall
Increase for colleges	-12.88	26.30	-0.72
	(0.44)	(0.54)	(0.30)
Increase for universities	10.61	-24.03	-0.13
	(0.21)	(0.40)	(0.06)
Current market share	45.53	20.47	66.00

Notes: Standard errors in parentheses.
[1] Reported as semi-elasticities: % change in market share given a cost increase of € 1,000.

Table 8: Cost elasticities at study field level

	Effect on[1]: Vocational study fields				Academic study fields			
	Arts	Social	Biomedical	Exact	Arts	Social	Biomedical	Exact
Increase for:								
Vocational * Arts	-43.03	2.87	1.99	2.08	4.09	2.48	1.74	1.43
Vocational * Social Science	28.34	-26.60	37.89	18.69	15.84	17.86	10.55	8.57
Vocational * Biomedical Science	2.36	4.15	-45.53	3.31	1.87	2.07	2.87	1.89
Vocational * Exact Science	5.40	5.43	8.03	-33.46	2.00	3.09	3.29	7.41
Academic * Arts	3.79	1.55	1.43	0.72	-39.17	4.26	3.29	2.19
Academic * Social Science	5.90	4.38	4.09	2.92	10.91	-38.28	10.76	9.78
Academic * Biomedical Science	1.72	1.13	2.42	1.29	3.52	4.51	-38.49	5.25
Academic * Exact Science	2.08	1.41	2.33	4.10	3.50	6.12	7.80	-36.44
Current market share	3.73	28.33	4.46	9.02	3.35	8.35	3.61	5.17

[1] Reported as semi-elasticities: % change in market share given a cost increase of € 1,000.

Table 9: Welfare effects of tuition fee increases

	Fee increase		Welfare effects[1]				Participation effects[2]		
	colleges	universities	consumer surplus	producer surplus	welfare (λ=0)	welfare (λ=0.23)	colleges	universities	total
Uniform small increase	+ €1,000	+ €1,000	-657.0	670.1	13.1	167.2	-1.09	-0.51	-0.91
Differentiated small increase	+ €750	+ €1,579	-645.1	670.1	25.0	179.2	8.04	-20.46	-0.80
Uniform cost-based increase	+ €2,810	+ €2,810	-1,830.8	1,852.3	21.6	447.6	-3.08	-1.46	-2.58
Differentiated increase to variable cost	$t_j = c_j$	$t_j = c_j$	-1,736.8	1,851.8	115.1	541.0	3.32	-14.98	-2.35
Uniform big increase	c_j + €5,000	c_j + €5,000	-4,881.7	4,918.2	36.5	1,167.7	-2.49	-17.31	-7.09

[1] The changes are in €/pupil and relative to the status quo.

[2] Reported as semi-elasticities: % change in market share given a cost increase of € 1,000.