Student Satisfaction Scores Affect Enrollment in Higher Education Programs

Bart K. de Koning[1], Annemarie Künn-Nelen[2,3,4], and Steffen Künn[2,3,4]

[1] *Cornell University*
[2] *Maastricht University*
[3] *ROA*
[4] *IZA Bonn*

September 19, 2023

Preliminary and incomplete - Please do not cite

Click here for the most recent version of the paper

Abstract

We study the impact of published student satisfaction scores (ranging from 1 to 5) on enrollment of first-year students for the near universe of higher education programs in the Netherlands between 2011 and 2019. We use pageview data from the largest Dutch educational information website to determine each programs' closest substitutes. This allows us to not only analyze the impact of changes in a program's own published student satisfaction score, but also the impact of changes in the student satisfaction scores of its substitutes. We analyze the impact of these satisfaction scores using fixed effects Poisson regressions and exploit rounding discontinuities to identify causal effects. On the whole, our findings show that student satisfaction scores matter for enrollment. An increase in a program's student satisfaction score leads to higher levels of enrollment, whereas an increase in the student satisfaction scores of substitutes leads to lower levels of enrollment. Point estimates of the impact of a program's student satisfaction score being rounded up to the next tenth on first-year enrollment range between 1.70% and 3.52%, depending on the bandwidth around the threshold we consider. Conditional on being above the rounding threshold, a program being rounded up over at least one of its closest substitutes increases first-year enrollment by up to 4.37%.

Keywords: Education choice, university enrollment, student satisfaction
JEL codes: D83, I23

Corresponding author: Bart K. de Koning (dekoning@cornell.edu). We would like to thank Erik Fleur and Jacco Tunzi of the Dutch Education Executive Agency for their help in acquiring data on enrollment. We would like to thank Pauline Thoolen, Bastian Schilderink and Nisan Mol of Studiekeuze123.nl for their help in acquiring data on satisfaction scores as well as pageview data.

1 Introduction

Every year, millions of students worldwide choose different university study programs. An important decision, as education affects many later life outcomes, such as labor market success (Kirkeboen, Leuven, & Mogstad, 2016), the likelihood of committing crimes (Machin, Marie, & Vujić, 2011) and marriage (Lafortune, 2013). However, the value of education does not only come from its returns upon completion. Earlier literature has shown that students enjoy substantial consumption value from their studies (see e.g., Alstadsæter 2011; Oosterbeek and Van Ophem, 2000; Oreopoulos and Salvanes, 2011).

Students decide on what study program to enroll in under imperfect information. Indicators such as rankings, labor market projections and student satisfaction scores are supposed to aid students' decisions. Due to progressing digitalization, students nowadays have easy, user-friendly and instant access to standardized information about universities and programs. The publication of such indicators on university and program quality may lead to better-informed choices, and enhance aggregate human capital and welfare by improving matches (Horstschräer, 2012). At the same time, we observe that markets for higher education have become more competitive, a process which initially started with the Bologna process, aimed at making international degrees more comparable. As such, universities have an incentive to invest in and promote quality indicators in order to attract more and better students (Hazelkorn, Loukkola, & Zhang, 2014).

Given the massive investments by public authorities and universities alike in these indicators, it is crucial to know whether the information provided indeed guides students' educational choices. Based on a number of observational and experimental studies, we know that labor market information about expected earnings and employment opportunities has a moderate, but generally positive effect on students' decisions to enroll in certain programs (see e.g., Bonilla-Mejía, Bottan, and Ham, 2019; Hastings, Neilson, and Zimmerman, 2015; Kerr, Pekkarinen, Sarvimäki, and RoopeUusitalo, 2020). We also know that university rankings play an important role in attracting students (Brewer, Eide, & Ehrenberg, 1999; Griffith

& Rask, 2007; Mueller & Rockerbie, 2005). The recent report by the *European University Association* underlines the increasing importance of university rankings and shows how university strategies are driven by these rankings (Rauhvargers, 2013). While labor market projections and rankings inform students on the likely impact of their decision on later life outcomes, they do not provide information on the more short-term consumption value of education. In this paper, we study how information about student satisfaction scores – determined by factors such as the quality of teaching, knowledge accumulation, the curriculum, community building, fellow student behavior and institutions' responsiveness and concern (Douglas, Douglas, McClelland, & Davies, 2015; Gibson, 2010) – impact enrollment.

Universities invest significant resources to improve student satisfaction. While student satisfaction may be important to universities in and of itself, they may also expect it to have spillover effects on more tangible metrics, such as student performance, student outcomes, student retention, and enrollment of new students. The evidence on these spillover effects is scant, however. While many studies in educational sciences have shown a strong correlation between student satisfaction and performance, there is no evidence for a causal interpretation. In one article, Bean and Bradley (1986) argue "that satisfaction had a greater influence on performance than performance had on satisfaction". However, their methodology would likely not hold up to today's standards.[1] In a recent study, Britton et al. (2021) show that student satisfaction is not related to student outcomes as measured by graduates' labor market earnings. This (disappointing) evidence suggests that the only remaining economic argument for universities to invest in student satisfaction is based on the assumption that student satisfaction is key to retain students and attract new ones. However, population-representative, causal evidence on this assumption is missing. To the best of our knowledge, there is only one study, by Horstschräer (2012), who shows a positive impact of student satisfaction score on applications. However, Horstschräer uses time variation to identify the effects of student satisfaction, and acknowledges that endogeneity concerns may still exist.

[1] They use a 2-stage least squares estimation strategy. They assume that institutional fit and utility influence satisfaction, but not GPA, and that high-school performance will influence GPA, but not satisfaction.

Moreover, this evidence is restricted to one specific program (medicine) with a strong selection of high-ability students. This subpopulation can be assumed to be well informed and hence particularly responsive to variations in publicly available indicators, including student satisfaction.

To our knowledge, we are the first to provide a detailed causal analysis on the role of student satisfaction in determining students' educational decisions. We rely on administrative records documenting the first-year enrollment figures of almost all university level study programs in the Netherlands between 2011 and 2019. We link these enrollment figures to satisfaction scores (ranging from 1 to 5) published on the largest Dutch educational information website (Studiekeuze123.nl) in June of the year before, from 2010 to 2018. These satisfaction scores are rounded before being published, providing us with a source of exogenous variation to assess their impact. Satisfaction scores are prominently displayed on the website and therefore likely used to compare study programs to one another. To fully capture their impact on enrollment, one needs to have information on satisfaction score of programs' close substitutes as well. To this end, we construct a substitutability matrix using the Studiekeuze123.nl's pageview data. Specifically, we determine how often two program pages are viewed by the same user. We argue that the more often two pages are viewed in tandem, the closer substitutes they are. These three data sources allow us to provide a detailed look into the impact of programs' student satisfaction scores, as well as the scores of their substitutes, on subsequent enrollment.

We start our analysis by running fixed effects Poisson regressions of first-year enrollment on published satisfaction scores. We find that there is a clear positive relationship between the two, that is surprisingly linear. In a next step, we also include the national average satisfaction score of the same degree program at different universities. It has a clear, negative relationship with enrollment. We further extend our analysis by adding a number of measures of the average satisfaction score of a program's closest substitutes as determined by our substitutability matrix. These measures also show a negative relationship with first-year enrollment. By adding these measures, the negative effect of the degree-level national average moves towards zero. Our interpretation of this finding is that our measure of substitutes' satisfaction scores is more informative than a 'naive' metric such as the degree-level national average. It underlines the importance of properly identifying substitutes.

While the fixed effects Poisson regressions take care of many potential sources of endogeneity, we cannot interpret these results as causal. In a next step, we therefore exploit the fact that satisfaction scores are rounded to the nearest tenth before being published on Studiekeuze123.nl. These rounding discontinuities allow us to compare programs when they were just below the rounding threshold to when they were just above the rounding threshold. We estimate a positive effect of being rounded up to the next tenth on subsequent first-year enrollment of 1.70% to 3.52%, but the effects are marginally significant and sensitive to the chosen bandwidth. Next, we look at programs whose satisfaction score is very close to the degree-level national average. We compare programs when they were rounded to the same satisfaction score as the national average to when they were rounded up over the national average. We find no impact of being just above the national average, but argue that this may be because programs very close to the national average have, by definition, very average satisfaction scores. Lastly, we study how important it is to receive a higher student satisfaction score than a close substitute. We show that being rounded up over a least one close substitute, conditional on being above the rounding threshold, increases first-year enrollment by up to 4.37%.

We contribute to the literature by providing the first causal evidence of the impact of student satisfaction scores on enrollment. We base this on observational data that include the near universe of existing university programs in an industrialized country. Our findings contribute to our understanding of the impact of information dissemination and provide new insights into the educational choice of students that goes beyond later life returns. The Netherlands is an ideal country to study this question, as there is strong competition between programs, because of geographical proximity as well as homogeneous university quality.

The rest of this paper is structured as follows. Section 2 explains the institutional context: the Dutch higher education system and the available informational resources. Section 3 describes the data we use and provides a first look into the relationship between satisfaction scores and enrollment. In Section 4, we present a more sophisticated data analysis, lay out how we identify the causal impact of satisfaction scores on first-year enrollment and present the results. Section 5 concludes.

2 Institutional Context

2.1 The Dutch higher education system

To understand how student satisfaction scores may impact enrollment, it is important to discuss the higher education market in the Netherlands. It is a highly competitive market on the supply side: tuition fees are fixed at the national level, admission to most programs is open conditional on prerequisites (i.e., no selection), university quality is relatively homogeneous, and the country is geographically small. Students thus have many options available to them, without major differences in selectivity, quality or costs. Relevant information on the expected utility they would enjoy at different available programs is therefore likely to drive their decisions.

In total, there are 37 universities of applied sciences and 12 research universities, which are all public entities. All universities are subject to monitoring by the Ministry of Education, with periodic accreditation rounds. This is to ensure homogeneous educational quality and standards across all public universities within the Netherlands (Inspectorate of Education, 2022). Students from the European Union enrolled at accredited public universities are eligible to subsidized tuition fees of about €2,000.- per year.[2] Both universities of applied

[2]Tuition fees have increased gradually over the years. For the 2010/2011 academic year, the subsidized tuition fee was €1,627 per year (Ministry of Education, Culture and Science, 2010). For the 2018/2019 academic year, it was €2,060 per year (Ministry of Education, Culture and Science, 2018). Students outside of the EU/EEA pay sticker price.

sciences as well as research universities generally offer a broad range of degree programs, with the exception of three research universities of technology.

Students' level of secondary education is decisive for the type of university students can apply for. Universities of applied sciences are open to graduates of higher general continued education, preparatory scientific education and middle-level applied education. Research universities are only open for graduates of preparatory scientific education and under certain conditions to graduates of universities of applied sciences. Most students entering universities of applied sciences come from higher general continued education. For research universities, most students come from preparatory scientific education (Inspectorate of Education, 2022). Both at higher general continued education and preparatory scientific education tracks, students have to choose a specific educational profile in their third year. An educational profile consists of a number of fixed subjects, that prepare pupils for specific areas of study. There are four different profiles: culture and society, economy and society, nature and health, nature and technology (Nuffic, 2019). The different profiles provide access to different study programs in higher education. In most cases, having the right degree level as well as educational profile provides access to a study program. Since 2014, students are required to complete an advisory intake procedure before enrollment, but its results are not binding (Chu & Westerheijden, 2018). Only in exceptional cases, universities can apply a 'numerus fixus', limiting the number of students that can enroll. They can do so because of excessive demand, high costs of education or poor labor market prospects. Some degrees, such as medicine, have a numerus fixus at all universities. Others, such as psychology and international business only have a numerus fixus at particular universities. Once a numerus fixus has been instated, the enrolment process changes as well. Students that apply for a program with a numerus fixus have to write a letter of motivation and are often required to take a test. Aside from that, there is a maximum number of numerus fixus study programs that students can apply for (generally two or three).

The uniform governmental policies have prevented the rise of an institutional hierarchy

(Veerman et al., 2010). All Dutch research universities are within the global top 250, but none of them is in the global top 50 according to the Times Higher Education ranking (2022). The fact that these universities all rank fairly high, arguably without any university being 'world class', means that students are unlikely to consider particular universities to be more desirable than others based on prestige. The competition among universities is further strengthened because of the relatively small geographical size of the Netherlands, that makes spatial accessibility high (Sa, Florax, & Rietveld, 2004). Even though the universities are not equally spread throughout the Netherlands, the size of the country enables students to reach any university either by commuting or small distance relocation.

2.2 Informational resources

To aid prospective students in their decision making, The Dutch Ministry of Education and Science, together with students and higher education institutions, founded the 'Studiekeuze123' foundation. The aim of this foundation is to provide students with objective information about post-secondary education programs at research universities and universities of applied sciences. To this end, the foundation does two things: it maintains the national student survey and operates the Studiekeuze123.nl website.

The national student survey is an annual survey sent out to all individuals enrolled in higher education in the Netherlands.[3] All students enrolled at government funded universities are invited to participate in the survey. More than 200,000 students answered the survey each year between 2010 and 2018. The survey asks students how satisfied they are about the study programs they are enrolled in on a number of dimensions (content, professors, facilities, general atmosphere, etc.). The most important and widely publicized metric is students' general satisfaction with their study program. The relevant question reads *"How satisfied are you with your course programme in general?"*. Students can answer on a scale from 1 (very dissatisfied) to 5 (very satisfied).

[3] There was some discussion about the participation of universities of applied sciences in 2020, but this is beyond the scope of our data.

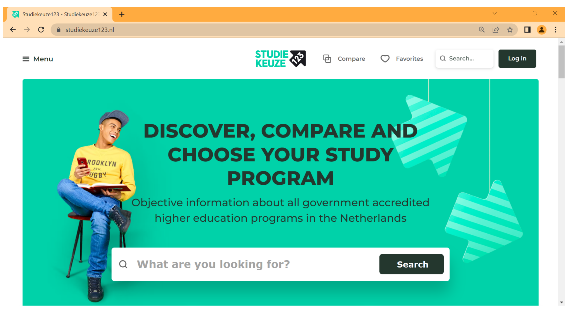
Studiekeuze123 visualizes and publishes the information on their website. The Studiekeuze123 website is the largest website for higher education information in the Netherlands, boasting 632,884 unique visitors between May of 2019 and June of 2020. Figure A1 in the Appendix displays the number of daily visitors during this time period.

Figure 1 shows a number of important pages on the Studiekeuze123.nl website. Panel (a) shows the Studiekeuze123.nl homepage. The website is designed in such a way that looking up programs is front and center. It uses a smart search bar, that directs students to the available programs based on the keywords they put in. Once students find a program, they are redirected to the program page, displayed in Panel (b). Students can find a wide range of information on this page: information about the program, requirements for enrollment, important dates for their application, statistics on the number and demographic characteristics of enrolled students, student satisfaction, graduation rates, labor market statistics and contact information. Note that the general satisfaction score of a program is listed and immediately visible as students open the page.

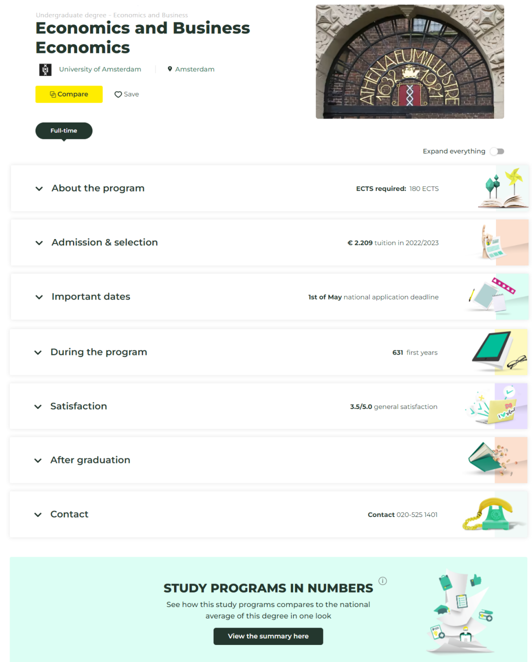
On the program page, students have a number of further options. Two are of particular interest to our study. Firstly, they can open the 'Study Programs in Numbers' sheet for the program. This is a quick summary of all the program's relevant statistics. Panel (c) provides a translated example of (part of) the sheet. The first thing students see when they open the sheet is the general student satisfaction score of the program they are looking at, as well as the national average for the degree. The page further provides information on the number of first year students, contact hours, as well as (not displayed) the share of students who continue on to the second year, share of students who receive degrees within four years, further education and labor market information[4]. Another option students have is to compare the program to other programs they are interested in. Panel (d) provides an example of such a comparison between two programs. Again, the student satisfaction scores of both programs are the first thing students see when they navigate to this page. Figure

[4]If the information pertains to a program at a University of Applied Sciences or a Masters' program.

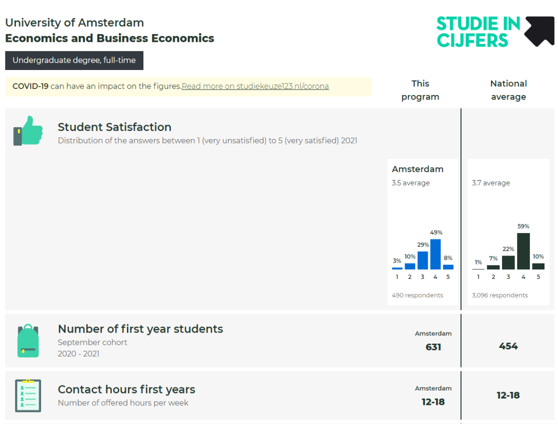
Figure 1: Studiekeuze123.nl website



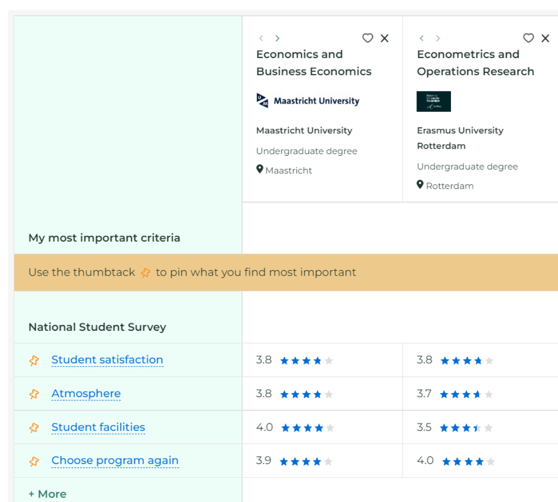
(a) Studiekeuze123.nl homepage



(b) Studiekeuze123.nl program page



(c) ‘Study Programs in Numbers’ page



(d) Program comparison page

Note: Translated versions of original Studiekeuze123.nl webpages.

1, combined with the large number of visitors on Studiekeuze123.nl, suggests that students are likely to be aware of the satisfaction scores of the programs they are interested in, and could thus have a real impact on enrollment decisions.

3 Data

3.1 Enrollment

Our outcome of interest is the enrollment numbers of Dutch study programs. For this, we rely on data we received from the Dutch Education Executive Agency (DEEA). On the first of October of each year[5], the DEEA records the enrollment status of all students in higher education. For each student, they register in which degree program(s) they are enrolled, as well as at which university. We refer to the degree program-university combination as a study program. Aside from the enrollment status of each student, the DEEA records a number of background characteristics, such as gender, migration background and secondary school results. We have access to data from 2011 to 2019. As our analyses are at the study program level, we reshape the data to a panel at that level. For each year, we record the number of (first-year) students enrolled in the program and how many have certain characteristics. We have data on 1,894 programs over a period of 9 years in our sample.

The top part of Table 1 provides some descriptive statistics on enrollment. The mean 'cohort size' is about 72, but the standard deviation is just above 91. The size of the standard deviation makes clear that there are large differences between programs in the number of first-year students that enroll. Figure A2 in the Appendix shows the distribution of the number of first-year enrolled students. The distribution is similar to a Power law distribution; there are a large number of small programs and a small number of large programs. The largest programs enroll more than 800 new first-year students on a yearly basis. The data look the same for the total number of students enrolled, as well as the number of first-year male students and students with a migration background; the standard deviation is large compared to the mean. The average high school grade of enrolled students is 6.57, with a standard deviation of just 0.31. It shows that programs are fairly homogeneous in terms of student ability; likely a product of the open admission policy in Dutch higher education.

[5]This is about a month after the official start of the academic year.

Table 1: Descriptive statistics on enrollment and satisfaction scores

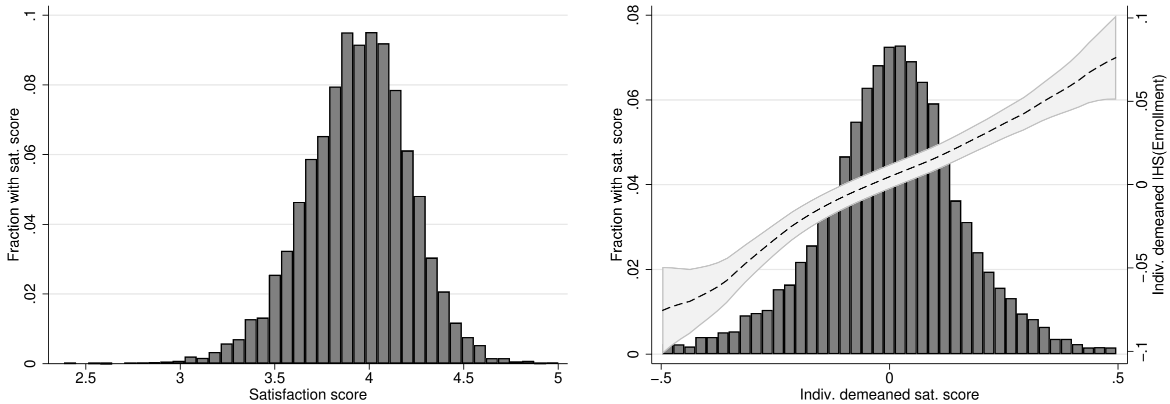
	Mean	Std. dev.	Within std. dev.
Enrollment			
No. of first-year students	71.86	91.14	22.34
Total no. of students	328.93	360.34	70.33
No. of male first-year students	33.58	51.89	12.01
No. of first-year students w/ migration background	25.52	44.60	14.43
Avg. high school grade	6.57	0.31	0.174
Sat. scores			
Student satisfaction score	3.94	0.28	0.167
National avg. sat. score of degree	3.92	0.198	0.103
Avg. sat. score. no. 1 substitute	3.94	0.254	0.148
Avg. sat. score. top 3 substitutes	3.94	0.190	0.103
Avg. sat. score. top 5 substitutes	3.94	0.168	0.088
Avg. sat. score. top 10 substitutes	3.93	0.147	0.075
Avg. sat. score. top 20 substitutes	3.94	0.132	0.066
Weighted avg. sat. score of substitutes	3.81	0.130	0.097
Observations	12821		

Note: data at program-year level. *No. of first-year students*, *total no. of students*, *no. of male first-year students*, *no. of first-year students w/ migration background* and *avg. high school grade* are based on data from the Dutch Education Executive Agency. *Student satisfaction score*, *national avg. sat. score of degree*, *weighted avg. sat. score of substitutes*, *weighted avg. sat. score top 20 subst.* are based on the 'Studiekeuzedatabase' from Studiekeuze123.nl. In case of the latter two variables, the substitutes are determined as described in Section 3.3, using Studiekeuze123.nl's pageview data. Within standard deviation is the standard deviation of the individually (program level) demeaned variables.

3.2 Student satisfaction scores

We obtain information on the historical satisfaction scores for the different study programs directly from Studiekeuze123's database records from June of every year. For each study program, we have access to the unrounded student satisfaction score. These scores are rounded to the nearest tenth before they are uploaded to the Studiekeuze123 website. To analyze the impact of the satisfaction scores on subsequent enrollment, we match the satisfaction scores published in June of year $t-1$ to enrollment in year t. Students likely decide on their study program before June and are even forced to register before the 1^{st} of May since the 2014/2015 academic year. Figure A1 in the Appendix shows that in 2019 and 2020, the high traffic months on Studiekeuze123.nl are October, January and April.

Figure 2: Distribution of satisfaction scores



(a) Distribution of satisfaction scores

(b) Sat. scores & first-year enrollment

Note: The x-axis displays the satisfaction scores. Panel (a) displays distribution of raw satisfaction scores. In Panel (b) the satisfaction scores are demeaned at the program level. The right-hand side y-axis of panel (b) shows the inverse hyperbolic sine of enrollment, again demeaned at the individual level. The dashed line displays the polynomial fit, with the gray area displaying the 95% confidence interval.

Panel (a) of Figure 2 shows the distribution of student satisfaction scores. The first row of the bottom part of Table 1 provides further details. The Figure shows that satisfaction scores are approximately normally distributed, with a mean just below 4.0 (3.94). The standard deviation is equal to 0.28. Panel (b) provides a first look into the relationship between student satisfaction scores and first-year enrollment. The x-axis shows the program level demeaned satisfaction score. The left y-axis shows the distribution of these scores and the right y-axis shows the program-level demeaned inverse hyperbolic sine[6] of enrollment. The reason we demean at the program level is that the variance of enrollment is very high, as discussed in Section 3.1. The within-program variance is substantially lower, as shown in Table 1. Moreover, by only exploiting within-program variation in our analyses, we take care of any time-invariant endogeneity concerns. For instance, one may expect that programs with larger budgets have higher student satisfaction, but are able to invest more in marketing as well. Time-varying endogeneity concerns, such as budgetary changes, do remain. Keeping

[6]The inverse hyperbolic sine is a transformation similar to the natural logarithm. However, contrary to the natural logarithm, it allows for values of zero.

these caveats in mind, Figure 2 shows a clear positive relationship between enrollment and a program's satisfaction score.

3.3 Substitutes

While satisfaction scores may be an important determinant of a prospective student's enrollment decision, their value does not have a clear interpretation. Higher levels of satisfaction scores are unlikely to drive larger number of students to enroll in higher education. It is more likely that students use these numbers to compare different programs they are interested in to each other. Therefore, it is key to identify programs' substitutes for the empirical analysis.

Ideally, we would be able to calculate the *cross-price elasticity* between programs. However, this would require (i) exogenous price changes and (ii) enough data to properly estimate the elasticity. Our setting provides neither. As stated in Section 2, the cost of enrollment is the same for almost all study programs. In addition, we only have nine years of enrollment data available to us, for fewer than 2,000 programs.

An alternative to the cross-price elasticity would be to define substitutes by their characteristics. For instance, it is likely that a program's closest substitutes are the same degree programs at different universities close by, or same-sector programs at the same university. However, this would cause us to potentially miss a lot of substitutes that do not fit these criteria. For instance, students interested in the math-heavy degrees in the economics and business sector, may also be interested in engineering degrees. These degrees are not in the same sector, however, and generally not available at the same universities. We thus require a more sophisticated approach.

To determine each program's closest substitutes, we make use of pageview data from Studiekeuze123.nl between May of 2019 and June of 2020. These data allow us to observe the behavior of prospective students looking for information about study programs. Apart from providing us with insight into how prospective students use these types of websites, it also allows us to construct a substitutability matrix. We construct the substitutability matrix by looking at which two program pages are viewed by the same user most often. We argue that if students who look at a certain program are highly likely to look at another program, these are likely to be close substitutes. To be more precise, we calculate the following:

$$S_{j,k} = \frac{Views_k | Viewed_j}{Views_j}. \tag{1}$$

Here, $S_{j,k}$ is our substitutability index. $Views_k | Viewed_j$ is the number of views to program k's page by individuals who have also viewed program j's page. $Views_j$ indicates the number of views that program j's page received in total. The higher this number, the larger the share of individuals who visit program j's page as well as program k's page.

A potential issue with this metric is that it is attenuated by individuals who look at a very large number of programs. We may want to assign more weight to individuals looking at just two programs, than to those looking at twenty programs. We therefore construct a weighted substitutability index:

$$S^w_{j,k} = \frac{\sum_{i=1}^{N} \frac{1}{Page\ Visits_i} Viewed_{i,k} | Viewed_{i,j}}{Views_j}. \tag{2}$$

Here, $\sum_{i=1}^{N} \frac{1}{Page\ Visits_i} Viewed_{i,k}$ adds a weight to individual i's visit that is the inverse of the number of pages individual i visited. For each program, we can then rank the substitutes in descending order of $S^w_{j,k}$.

Table B1 in the Appendix provides an example for the *Econometrics and Operations Research* degree at Erasmus University Rotterdam; one of the twelve research universities of the Netherlands. Its number one ranked substitute is *Econometrics* at the Free University of Amsterdam. Unsurprisingly, this is the same degree program at a university close by. The substitutes ranked number two and four are Econometrics at the University of Amsterdam and *Economics* at Erasmus University Rotterdam, respectively. What's interesting about these programs is that the program page of Economics at Erasmus University Rotterdam was viewed more often by people who also viewed Econometrics at the university ($s_{j,k}$ is higher than for Econometrics at the University of Amsterdam). However, most of these visitors viewed a large number of programs ($s^w_{j,k}$ is lower than for Econometrics at the University of Amsterdam), indicating that these were perhaps visitors more uncertain about their choice. Moving slightly down the ranks, *Engineering* at Delft University of Technology shows up at the eighteenth place. While not as close of a substitute as the Econometrics and Economics programs, we observe that 6.4% of individuals who visited the program page for Econometrics at Erasmus University Rotterdam, also visit the program page for Engineering at Delft University of Technology, which means they are definitely substitutes to some degree. This makes sense, as both programs require a very high level of mathematical ability to complete. *Medicine* shows up at the twentieth place. Again, not such an obvious substitute, but it is likely that high ability students are interested in both Econometrics – which is regarded as difficult – and Medicine – which is tough to get into without stellar grades.

To further test the credibility of the identified substitutes, Table B2 in the Appendix shows how similar programs' substitutes are based on their rank. We would expect that close substitutes are often (i) at the same university as the study program considered or (ii) the same degree program at a different university. Columns (1) and (2) of Table B2 show that 41.9% and 33.4% of the closest substitutes according to our metric are at the same university and the same university location[7], respectively. This number drops by 6.39 and 4.18 percentage points for rank 2, and continues to drop until rank 20. Still, over 30% of the substitutes at rank 20 are at the same university, and over 25% at the same university location. This makes sense, as students will often have some geographic preference, despite the small country size. Column (3) shows the likelihood of the closest substitute being the same degree program. Again, we see that the likelihood of this being true for the number one ranked substitute is quite high, at 43.6%. It drops when we move down the ranks. The drop is noticeably sharper here, as the number of universities one can study a particular degree program at is often limited. What makes these findings even more striking, is that it is impossible for a study program's substitute to be the same degree program at the same

[7]Some universities offer educational programs at more than one location.

university location, as that is the study program itself. Put differently, Columns (2) and (3) are mutually exclusive. Column (4) combines the two and shows the likelihood of a substitute being either at the same university and location, or the same degree program. The number one ranked substitute of a program is either at the same university and location or the same degree program at a different university in 77.1% of cases. These results clearly show that the substitutes we find make sense. However, while 77.1% of closest substitutes are either at the same university location or the same degree program, neither is the case for 22.9% of study programs; almost a fourth of all programs. The substitutability matrix is particularly valuable for these programs, as it allows us to identify substitutes in an organic manner that does not rely on overlap between topics or locations. Columns (5) and (6) further show the mean value of our substitutability metrics by rank.

4 The Impact of Satisfaction Scores on Enrollment

4.1 Basic fixed effects regressions

Our main goal is to study the relationship between first-year enrollment and (i) a program's own student satisfaction satisfaction score, (ii) the national degree average for that program, and (iii) the satisfaction scores of its substitutes. To this end, we use fixed effects Poisson regressions. The fixed effects take care of the large between program variance in first-year enrollment, as well as account for the time-invariant differences between programs. We use Poisson regressions as our outcome measure is a count variable. Specifically, we estimate the following equation

$$\begin{aligned}\text{First-year enrollment}_{j,t} = \boldsymbol{\beta}\,\textbf{Satisfaction Score Metrics}'_{j,t-1} + \\ \boldsymbol{\gamma}\mathbf{T}' + \zeta_j + \varepsilon_{j,t}.\end{aligned} \tag{3}$$

Here, First-year enrollment$_{j,t}$ indicates the number of students enrolled in the first year of program j at time t. **Satisfaction Score Metrics**$'_{j,t-1}$ is a vector and includes the relevant satisfaction score metrics for program j at time $t-1$. $\mathbf{T}'$ is a vector of time dummies, ζ_j are the program fixed effects and $\varepsilon_{j,t}$ is the idiosyncratic error term.

Table B3 in the Appendix provides a first look into the impact of a program's satisfaction scores, as well as the development of satisfaction scores over time. The first two columns of the Table show the results of a fixed effects Poisson regression of first-year enrollment on a program's satisfaction score (as a continuous variable and individual dummies, respectively) as well as the academic year. The Table shows that a program's satisfaction score has a positive relation to first-year enrollment in the subsequent years. Column (2) shows that the relationship is fairly linear, except for at the tail ends of the satisfaction score distribution. Columns (1) and (2) further show that programs have grown substantially between 2011 and 2019. Column (3) shows that there is a clear positive time trend in satisfaction scores as well, having increased by almost 0.15 in 2019 compared to 2011 (scores published in June of 2018 and June of 2010). It underlines how important it is that we control for the academic year in all of our subsequent analyses.

The results from Table B3 show that a program's satisfaction score is indeed positively correlated with first-year enrollment. However, the model is incomplete, as it does not account for the satisfaction scores of substitutes. Table 2 provides a more complete picture. It shows the results of fixed effects Poisson regressions of first-year enrollment on a program's own satisfaction score, the degree's national average, as well as a number of metrics for the satisfaction score of its closest substitutes. All of the coefficients can be approximately interpreted as the %-impact of a 0.1 increase in the relevant satisfaction score on first-year enrollment. Column (1) confirms the positive relationship between a program's own satisfaction score and first-year enrollment. An increase of 0.1 in the satisfaction score is associated with an increase of approximately 1.9% in first-year enrollment. Column (1) also shows that the national average of all programs that offer the same degree is associated with lower first-year enrollment numbers. As shown in Table B2, the closest substitutes for a program are often the same degree programs at different universities. A higher national average therefore makes the outside option more attractive, leading to lower first-year enrollment in the considered study program. Columns (2) to (7) of Table 2 all add a different metric of the average satisfaction score of a program's substitute. In Columns (2), (3), (4), (5) and (6), we include the average satisfaction score of the closest one, three, five, ten and twenty substitutes, respectively. As we increase the number of substitutes we consider, the impact of the national average decreases, and the impact of the average satisfaction scores of the substitutes becomes more pronounced. In Column (6), where we consider the average satisfaction score of a program's closest twenty substitutes, we see that the impact of the national average is no longer significant. It is likely that the degree-level national average provides a partial picture of the satisfaction scores of a program's closest substitutes. The average satisfaction score of the identified substitutes looks to contain this information, as well as more, crowding out the impact of the national average. It is a testament to the importance of identifying substitutes in the way we do. It may seem counterintuitive that the impact of substitutes' satisfaction scores increases as we include more substitutes; the satisfaction scores of closer substitutes are likely to be more important. However, the measure of the satisfaction score of the closest substitute is much more noisy, potentially leading to attenuation bias. On top of that, the standard deviation decreases as we include more substitutes, as shown in Table 1. Table B4 shows that the increase remains, but is slightly less pronounced when we standardize the program's satisfaction score, national average and the satisfaction score of substitutes. Column (7) of Table 2 shows that when we consider a weighted average of the satisfaction score of all substitutes of a program, there seems to be no effect. This measure of substitutes' satisfaction scores may be inaccurate, however; as we may assign too much weight to very far-away substitutes because we include data from e.g. scraping bots. The negative impact of the national average remains.

Beyond Table 2, it is interesting to see whether any particular programs or type of students drive these results. Table B5 in the Appendix provides a look into these heterogeneous effects. Columns (2) and (3) split the sample into undergraduate and master's programs. For

Table 2: Relationship between satisfaction score, national average, satisfaction score of substitutes & first-year enrollment

	(1) W/o substitutes	(2) No. 1 subst.	(3) Top 3. subst.	(4) Top 5 subst.	(5) Top 10 subst.	(6) Top 20 subst.	(7) All substitutes (weighted)
Dep. var: **first-year enrollment**							
Satisfaction score of program (×10)	0.0191***	0.0191***	0.0193***	0.0194***	0.0195***	0.0198***	0.0191***
	(0.00330)	(0.00329)	(0.00327)	(0.00328)	(0.00328)	(0.00329)	(0.00329)
National average (×10)	-0.0126**	-0.0130**	-0.0112*	-0.00975	-0.00876	-0.00732	-0.0125**
	(0.00607)	(0.00611)	(0.00605)	(0.00605)	(0.00622)	(0.00612)	(0.00601)
Satisfaction Score Substitute(s) (×10)		0.00115	-0.00465	-0.0104*	-0.0171**	-0.0315***	-0.00169
		(0.00277)	(0.00488)	(0.00595)	(0.00753)	(0.0112)	(0.00790)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12821	12821	12821	12821	12821	12821	12821

Note: Regressions are at study program-year level, estimated by fixed effects Poisson regression. All measures of satisfaction scores are multiplied by 10. This means coefficients can be interpreted as the impact of an increase of 0.1. *Satisfaction score of program* denotes a program's own satisfaction score. *National average* denotes the national average of the same degree program at all universities that offer it. *Satisfaction score of substitutes* indicates the average satisfaction score of all substitutes that are ranked above the number indicated in Column titles (2) to (7). Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

the undergraduate program's we confirm the impact of a program's own satisfaction score on enrollment. While the impact of the satisfaction score of substitutes remains negative, the estimate is statistically insignificant. For master's programs, both coefficients retain the same sign as in Table 2, but lose their significance. We cannot conclude that the effects are driven by either undergraduate or master's programs. Columns (4) and (5) of Table B5 show the impact of the satisfaction score metrics on the number of first-year enrolled men, and the number of enrolled first-year students with a migration background. Both outcomes show similar results to those found in Table 2, indicating that neither students of a specific gender, nor students with (or without) a migration background drive our results.

Another potential concern is that our results are driven by small programs. Since fewer students from these programs fill out the national student survey, these programs naturally have larger variance in satisfaction scores. Moreover, small changes in student numbers may be percentually large. To ensure our results are robust to excluding programs from a certain size, Table B6 in the Appendix shows the results when we exclude different programs based on their average size over the years. Columns (1) to (4) show that the results hold when we only include the largest 95%, 90%, 75% and 50% in our analysis sample. Columns (5) to (8) show that the same is true for when we only include the smallest 50%, 75%, 90% and 95%. It proves that our results are not driven by programs at the tail ends of the size distribution, but are actually present across different program sizes.

All these results point towards a systematic relationship between programs' satisfaction scores and first-year enrollment. However, we cannot interpret these results as causal. It is likely that endogeneity still plagues these estimates. For instance, an increase in a study program's budget is likely to be spent on both facilities and marketing; the former increasing student satisfaction, the latter first-year enrollment. In the next Section, we will therefore attempt to identify the causal impact of these satisfaction scores by exploiting rounding discontinuities.

4.2 Identification of causal impacts

4.2.1 Rounding discontinuities

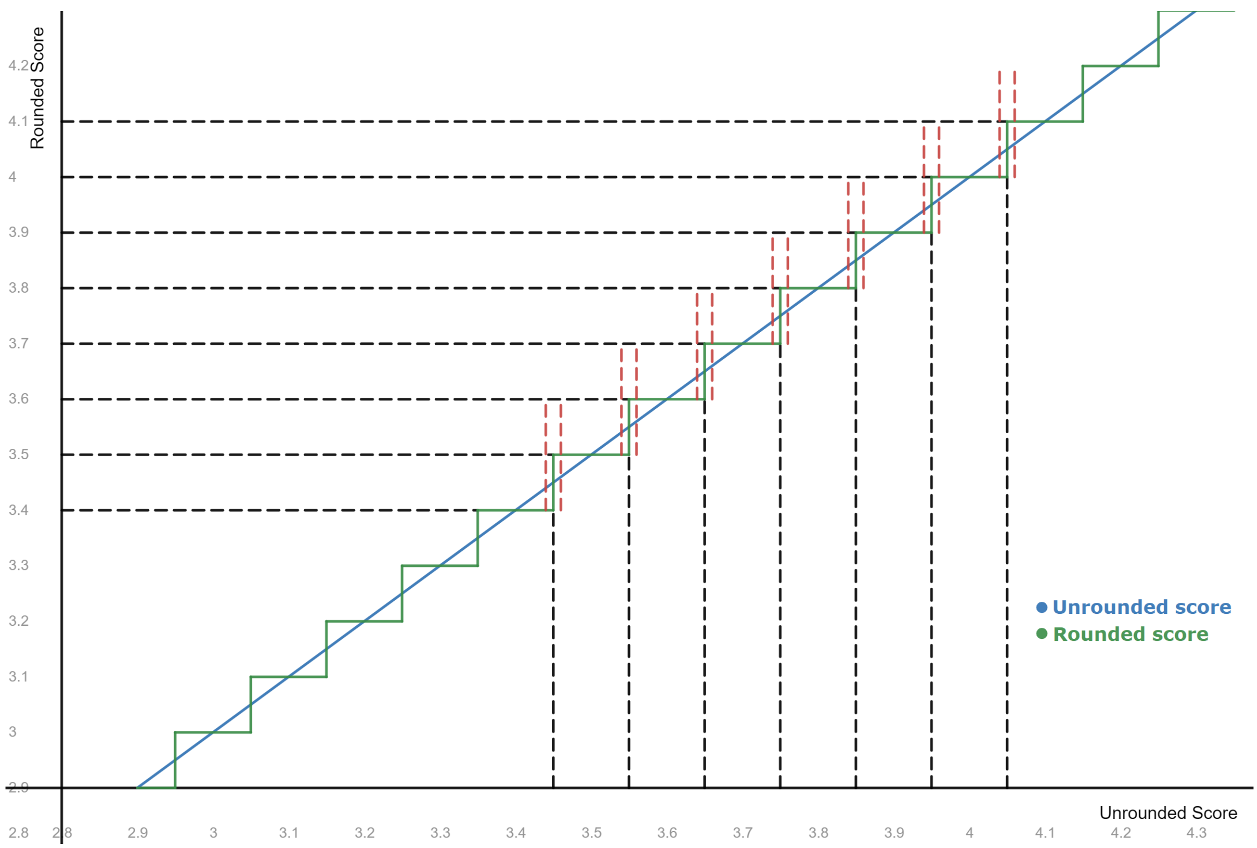
To identify the causal impact of satisfaction scores, we exploit the natural discontinuities that come with the way in which the programs' satisfaction scores are rounded before being published on the Studiekeuze123 website. Satisfaction scores are rounded to the nearest tenth. More formally, the rounded satisfaction score of a program is calculated as follows:

$$\frac{1}{10}\left\lfloor 10 \times (\text{Satisfaction score}) + \frac{1}{2} \right\rfloor, \tag{4}$$

where $\lfloor x \rfloor$ denotes the floor function, returning the largest integer that is smaller than, or equal to x.

The fact that Studiekeuze123 rounds these satisfaction scores before publication means that while the raw satisfaction scores of two programs may be very close, they are a tenth apart on the Studiekeuze123 website. We can exploit this rounding discontinuity to study the effect of having a higher published satisfaction score. We compare programs when they ended up just below to when they ended up just above a particular threshold, again using a fixed effects Poisson regression. Figure 3 provides a visual representation of the analysis we do. The blue line represents the unrounded satisfaction score of a program. The green line represents the rounded satisfaction score. At each $.05^{th}$, the rounded and published satisfaction scores makes a discontinuous jump to the next tenth. We exploit this by comparing programs very close to this discontinuity, indicated by the bandwidths in red. These programs have very similar unrounded satisfaction scores, but their published satisfaction scores differs. We once again estimate our specifications using fixed effects Poisson regressions. For this analysis, we also include dummies for every threshold, as well as polynomials of the distance to the relevant threshold.

Figure 3: Visual representation rounding discontinuity analysis



Note: Visual representation of discontinuity analysis. Straight (blue) line indicates unrounded score. Green (stepwise) line indicates rounded score. Red (dashed) lines indicate bandwidth around the rounding threshold. We consider different bandwidths in our analysis.

Table 3 shows the impact of being just above the rounding threshold, compared to just below, on subsequent enrollment of first year students for multiple bandwidths around the rounding thresholds. Our estimates show a positive impact of being just above the threshold

Table 3: Impact of being just above satisfaction score rounding threshold on first-year enrollment

	B = 0.01	B = 0.02	B = 0.03	B = 0.04	B = 0.05
Dep. Var: **first-year enrollment**					
Rounded up	0.0325	0.0352*	0.0295*	0.0178	0.0170
	(0.0346)	(0.0209)	(0.0159)	(0.0136)	(0.0120)
Year dummies	Yes	Yes	Yes	Yes	Yes
Threshold dummies	Yes	Yes	Yes	Yes	Yes
Distance polynomials	Yes	Yes	Yes	Yes	Yes
Observations	1645	4127	6716	9102	10902
Rounded up	879	2131	3409	4601	5474

Note: regressions estimated by fixed effects Poisson regression. Standard errors are clustered at the study program level. *Rounded up* indicates whether a program was rounded up (i.e., above the rounding threshold). Distance polynomials include a linear and squared term of the distance to the closest rounding threshold. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

of up to 3.52%, but are imprecise. For bandwidths of 0.02 and 0.03 around the threshold, the effect is marginally significant; for larger bandwidths, the estimates decrease and are statistically insignificant.

4.2.2 The degree-level average

The results from Section 4.1 indicate that not just a program's own satisfaction score, but also that of competing programs may be important. As discussed in Section 2, Studiekeuze123 often shows a program's satisfaction score next to the national average for the degree, making it a very salient metric to judge a program by. In this Section, we expand our analysis of the impact of the national average by estimating a more flexible specification compared to that in Table 2. In addition, we again exploit rounding discontinuities to investigate how being just above the national average impacts enrollment.

Column (2) of Table 4 shows the results of the extended specification. We confirm that the national average has a negative impact on enrollment, but do not observe any nonlinearities, or differences in coefficient size depending on whether a program's student satisfaction score is above or below the national average.

Table 4: Impact of being just above national average satisfaction score on first-year enrollment

	(1) Base specification	(2) Extended specification	(3) Rounded up over (B = 0.05)
Dep. Var: **First Year Enrollment**			
Satisfaction score of program (×10)	0.0191***	0.0235***	
	(0.00330)	(0.00688)	
National average (×10)	-0.0126**	-0.0176**	
	(0.00607)	(0.00864)	
Above national average		-0.00398	0.0138
		(0.0134)	(0.0480)
Above national average × Distance to national average		-0.00286	
		(0.0104)	
(Distance to national average)^2		0.00389	
		(0.0120)	
Above national average × (Distance to national average)^2		-0.0133	
		(0.0147)	
Year dummies	Yes	Yes	Yes
Observations	12821	12821	1922
N Above National Average			193

Note: regressions estimated by fixed effects Poisson regression. *Satisfaction score of program* denotes a program's own satisfaction score. *National average* denotes the national average of the same degree program at all universities that offer it. *Above national average* indicates whether program was rounded up over the national average or not. *Distance to national average* is the difference between a program's own satisfaction score and the degree level average. Its singular term is omitted, as it is collinear with the first two terms. *N Above National Average* indicates how many observations were treated. Distance polynomials include a linear and squared term of the distance to the closest rounding threshold. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

To once again exploit the rounding discontinuity, we take a sample of programs whose unrounded satisfaction score is within a certain bandwidth from the unrounded national average. To avoid conflating the effects we find with the effects found in Section 4.2.1, we only consider programs that are above the rounding threshold. The main difference between this analysis and that in Section 4.2.1, is that the likelihood of being 'treated' (i.e., rounded up over the national average) is not $\frac{1}{2}$. Narrow bandwidths cause both the program's satisfaction score as well as the national average to be well over the rounding threshold most of the time, meaning they will often be equal in publication.[8] With a larger bandwidth, the likelihood of a program being rounded up over the national average increases exponentially.

[8]To illustrate this, consider the bandwidth to be .01. Only when the program's satisfaction score is between $x.x5$ and $x.x6$ is there a chance that the national average is just below the rounding threshold.

It is therefore valuable to consider a relatively large bandwidth. We use a bandwidth of 0.05, which means we have 193 treated observations; just over 10% of the sample. Column (3) shows no effect of being rounded up over the degree-level national average. While this may be surprising given the saliency of the degree-level national average, it is important to note that the only programs we consider are programs that are average for their degree. Students will thus often have an outside option that provides a better satisfaction score, potentially muting the results.

4.2.3 Rounding discontinuities and close substitutes

Students may not only compare a program's satisfaction score to the national average, but are likely to also compare satisfaction scores between programs. Here, we re-use the methodology described in Section 4.2.2, but compare a program's satisfaction score to that of its closest substitutes instead of the national average. We analyze situations in which a program is just above the rounding threshold, and very close to (a number of) its closest substitute(s). The advantage is that this avoids the issue of all considered programs being very average. We once again take a bandwidth of 0.05 for this analysis, as it suffers from the same issue as our analysis of the national average: while the sample size decreases linearly with the bandwidth, the share of treated observations becomes exponentially smaller when the bandwidth decreases.

Table 5 shows the impact of being rounded up over at least one close substitute. We find that all of our estimates are positive, but tend to become smaller as we include more substitutes. The point estimate of a program being rounded up over its number one ranked substitute in Column (1) is positive, but not statistically significant. However, we have only 1041 observations, of which a mere 139 are treated. Column (2) shows a clearly positive and significant effect of being rounded up over at least one of the top 3 substitutes of 4.37%. A sizeable effect. Column (3) shows a qualitatively similar result, although the point estimate decreases slightly. The point estimates in Columns (4) and (5) are smaller, and not

Table 5: Impact of being rounded up over close substitute on first-year enrollment

	(1) No. 1 subst.	(2) Top 3 subst.	(3) Top 5 subst.	(4) Top 10 subst.	(5) Top 20 subst.
Dep. var: **first-year enrollment**					
Close call win	0.0155	0.0437**	0.0426***	0.0157	0.0148*
	(0.0300)	(0.0175)	(0.0160)	(0.0108)	(0.00876)
Year dummies	Yes	Yes	Yes	Yes	Yes
Threshold dummies	Yes	Yes	Yes	Yes	Yes
Distance polynomials	Yes	Yes	Yes	Yes	Yes
Observations	1041	3090	4552	6496	7840
N at Least one Win	139	476	760	1362	2116

Note: regressions estimated by fixed effects Poisson regression. Only includes programs when they were above the rounding threshold. *Close call win* indicates whether a program was rounded up over at least one close substitute. Columns indicate how many substitutes were considered. Distance polynomials include a linear and squared term of the distance to the closest rounding threshold. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

significant, or only marginally. The reason for this may be that the treatment variable 'close call win' is diluted. Being rounded up over the twentieth substitute is likely to have less of an impact to being rounded up over one of the top three substitutes of any one program.

5 Conclusion & Discussion

In this paper, we analyzed the impact of published student satisfaction scores on first-year enrollment in university programs. We show that satisfaction scores matter for enrollment: both the satisfaction score of the program itself as well as that of its substitutes. Analyses exploiting rounding discontinuities show that a satisfaction score being rounded up to the next tenth increases first-year enrollment in the subsequent year by 1.70% to 3.52%, although the estimates are imprecise. We further find that a program's relative student satisfaction score is of particular importance: conditional on being rounded up, a program that has a (slightly) higher published satisfaction score than at least one of its top substitutes will see an increase in first-year enrollment in the subsequent year of up to 4.37% on average.

Our findings underline the importance of digital information provision to students. However, it should perhaps also serve as a warning sign of how to present information. The fact that we find that substitutes being on different sides of rounding discontinuities has a significant and sizeable impact on the choices between these programs is not necessarily positive. It implies that students sometimes make a decision on which program to enroll in based on a difference that is more of an artefact of publication than a truly large difference in student satisfaction. Informational websites may thus want to emphasize the role of uncertainty in these types of metrics and urge students to interpret all information with care.

References

Alstadsæter, A. (2011). Measuring the consumption value of higher education. *CESifo Economic Studies*, *57*(3), 458–479.

Bean, J. P., & Bradley, R. K. (1986). Untangling the satisfaction-performance relationship for college students. *The Journal of Higher Education*, *57*(4), 393–412.

Bonilla-Mejía, L., Bottan, N. L., & Ham, A. (2019). Information policies and higher education choices experimental evidence from colombia. *Journal of Behavioral and Experimental Economics*, *83*, 101468.

Brewer, D. J., Eide, E. R., & Ehrenberg, R. (1999). Does it pay to attend an elite private college? cross-cohort evidence on the effects of college type on earnings. *Journal of Human Resources*, *34*(1), 104–123.

Britton, J., van der Erve, L., Belfield, C., Vignoles, A., Dickson, M., Zhu, Y., ... Buscha, F. (2021). *How much does degree choice matter?* (Working Paper No. 21/24). IFS.

Chu, A., & Westerheijden, D. F. (2018). Between quality and control: what can we learn from higher education quality assurance policy in the netherlands. *Quality in higher education*, *24*(3), 260–270.

Douglas, J. A., Douglas, A., McClelland, R. J., & Davies, J. (2015). Understanding student satisfaction and dissatisfaction: an interpretive study in the uk higher education context. *Studies in Higher Education*, *40*(2), 329–349.

Gibson, A. (2010). Measuring business student satisfaction: a review and summary of the major predictors. *Journal of Higher Education Policy and Management*, *32*(3), 251–259.

Griffith, A., & Rask, K. (2007). The influence of the us news and world report collegiate rankings on the matriculation decision of high-ability students: 1995-2004. *Economics of Education Review*, *26*, 244–255.

Hastings, J., Neilson, C. A., & Zimmerman, S. D. (2015). *The effects of earnings disclosure on college enrollment decisions* (Working Paper No. 21300). National Bureau of Economic

Researc. Retrieved from http://www.nber.org/papers/w21300

Hazelkorn, E., Loukkola, T., & Zhang, T. (2014). *Rankings in institutional strategies and processes: Impact of illusion?* (Tech. Rep.). European University Association.

Horstschräer, J. (2012). University rankings in action? the importance of rankings and an excellence competition for university choice of high-ability students. *Economics of Education Review*, *31*, 1162–1176.

Inspectorate of Education. (2022). *Staat van het onderwijs 2022.*

Kerr, S. P., Pekkarinen, T., Sarvimäki, M., & RoopeUusitalo. (2020). Post-secondary education and information on labor market prospects: A randomized field experiment. *Labour Economics*, *33*.

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, *131*(3), 1057–1111.

Lafortune, J. (2013). Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. *American Economic Journal: Applied Economics*, *5*(2), 151–78.

Machin, S., Marie, O., & Vujić, S. (2011). The crime reducing effect of education. *The Economic Journal*, *121*(552), 463–484.

Ministry of Education, Culture and Science. (2010). *Staatsblad 2010, 314.*

Ministry of Education, Culture and Science. (2018). *Staatsblad 2018, 226.*

Mueller, R. E., & Rockerbie, D. (2005). Determining demand for university education in ontario by type of student. *Economics of Education Review*, *24*, 469–483.

Nuffic. (2019). *The education system of the netherlands.* Retrieved 2019-08-04, from https://www.nuffic.nl/documents/459/education-system-the-netherlands.pdf

Oosterbeek, H., & Van Ophem, H. (2000). Schooling choices: Preferences, discount rates, and rates of return. *Empirical Economics*, *25*(1), 15–34.

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling.

Journal of Economic perspectives, *25*(1), 159–84.

Rauhvargers, A. (2013). *Global university rankings and their impact: Report ii* (Tech. Rep.). European University Association.

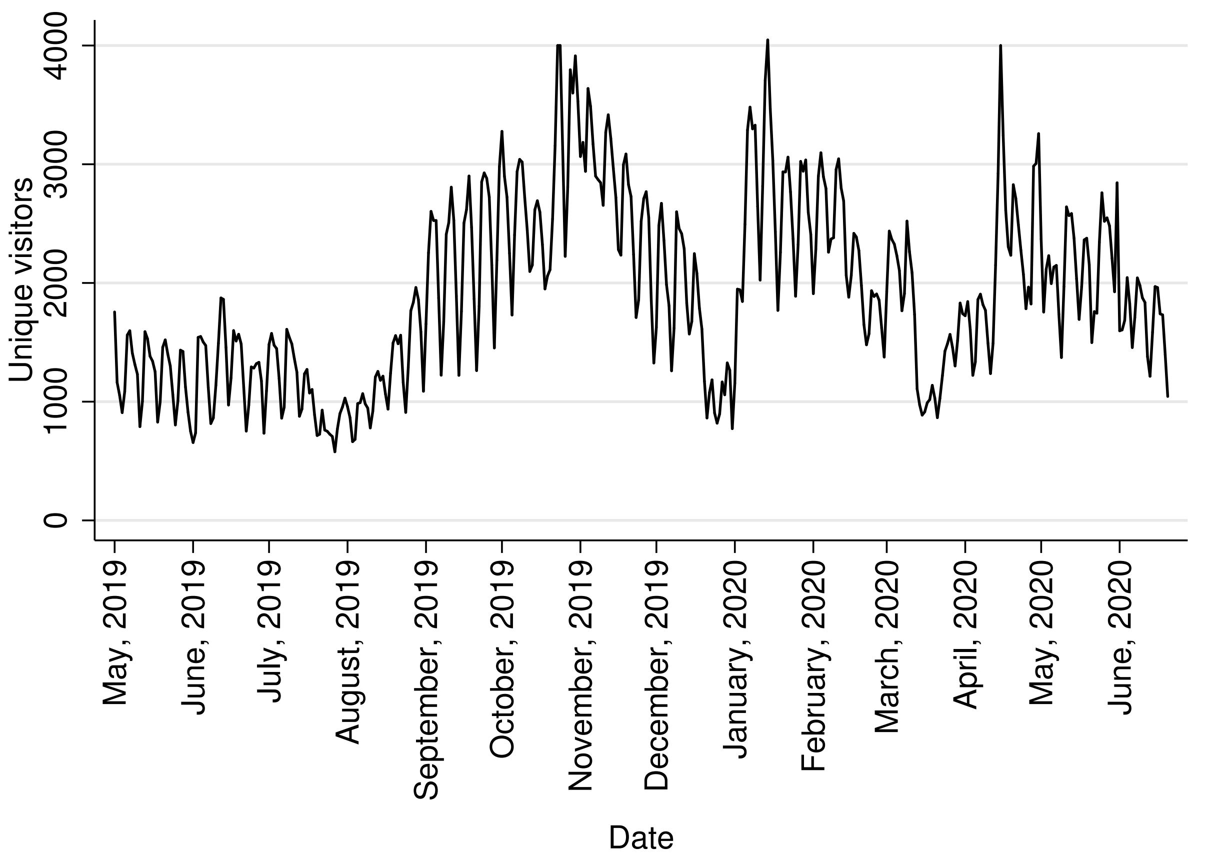
Sa, C., Florax, R. J., & Rietveld, P. (2004). Determinants of the regional demand for higher education in the netherlands: A gravity model approach. *Regional Studies*, *38*(4), 375–392.

Times Higher Education. (2022). *World university rankings.* Retrieved from `https://www.timeshighereducation.com/world-university-rankings/2022/world-ranking`

Veerman, C., Berdahl, R., Bormans, M., Geven, K., Hazelkorn, E., Kan, A. R., … Vossensteyn, J. (2010). *Threefold differentiation: recommendations of the committee on the future sustainability of the dutch higher education system.* Koninklijke Broese en Peereboom.

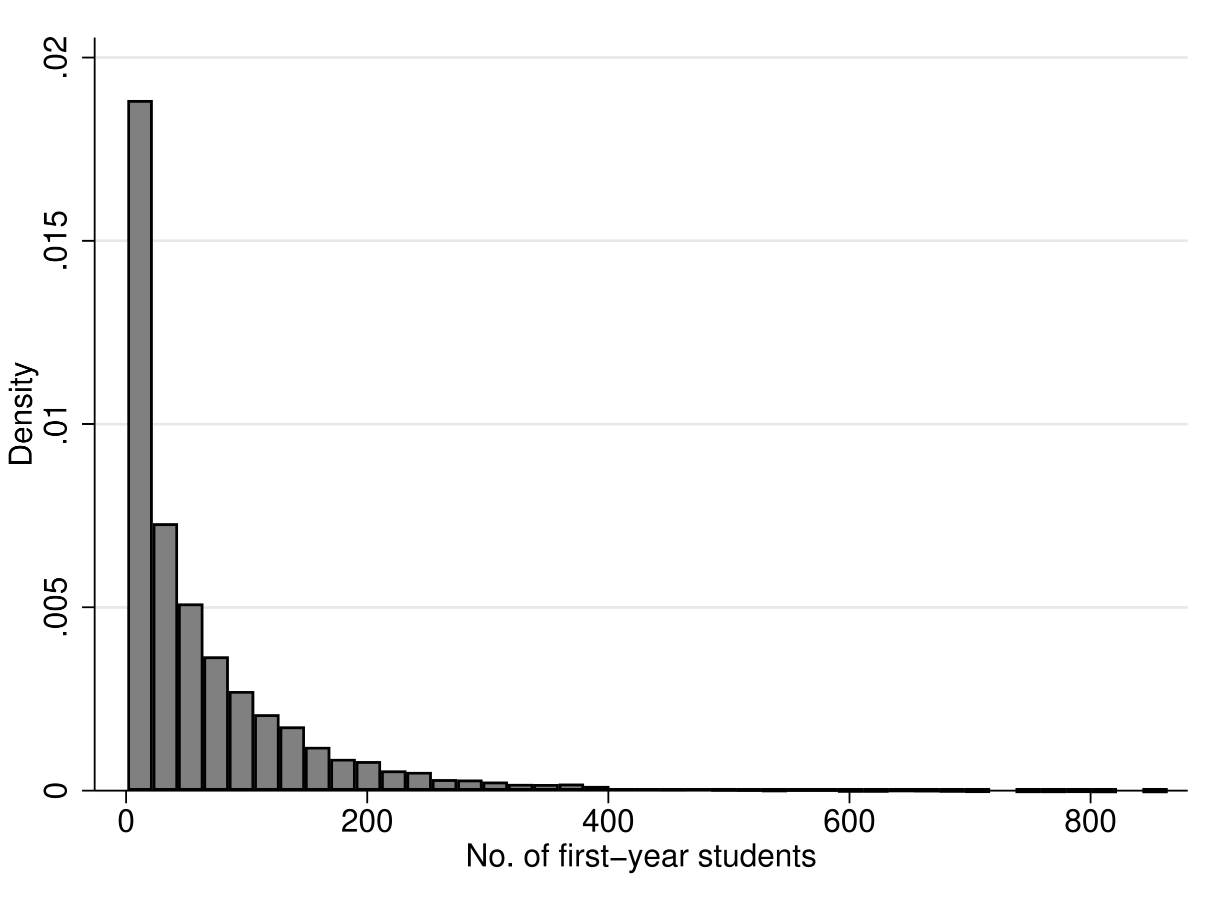
Appendix A: Additional Figures

Figure A1: Daily visitors to Studiekeuze123.nl



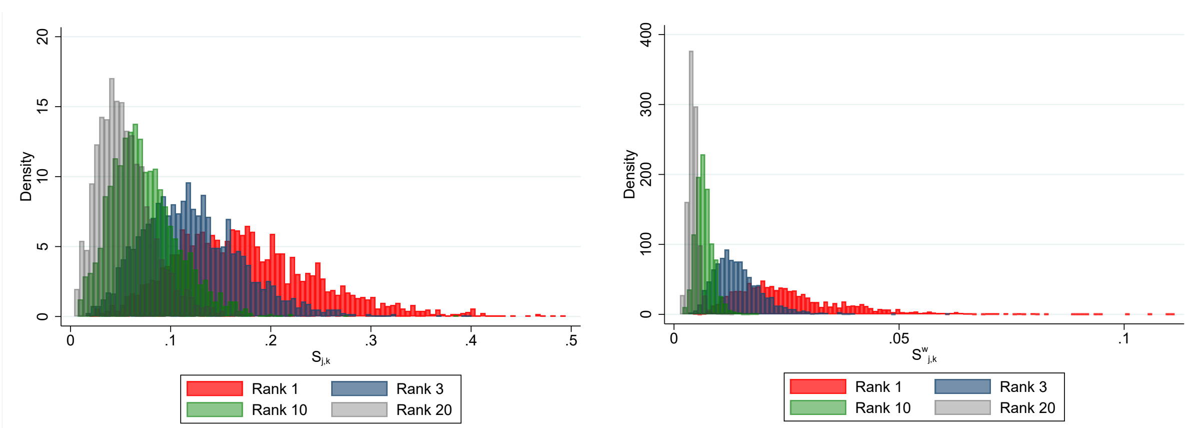
Note: Figure shows number of daily visitors to Studiekeuze123.nl between May of 2019 and June of 2020. Four values were (much) larger than 4,000. They have been winsorized to 4,000 to make the graph more easily readable. Original values: 14^{th} of January, 2020: 4,047 visitors; 24^{th} of October, 2019: 5,404 visitors; 15^{th} of April, 2020: 7,548 visitors; 23^{rd} of October, 2019: 18,111 visitors.

Figure A2: Distribution of no. of first-year enrolled students



Note: Figure shows distribution of no. of first-year students enrolled in programs. Data at program-year level.

Figure A3: Distribution of $S_{j,k}$ and $S^{w}_{j,k}$



(a) Distribution of $S_{j,k}$ by rank

(b) Distribution of $S^{w}_{j,k}$ by rank

Note: Figure shows distribution of substitutability metrics $S_{j,k}$ and $S^{w}_{j,k}$ as explained in Section 3.3 by rank of substitute (based on $S^{w}_{j,k}$).

Appendix B: Additional Tables

Table B1: Substitutes for Econometrics at Erasmus University Rotterdam (3,528 views)

Substitute	**Substitute rank**	**Times also viewed**	$S_{j,k}$	**Weighted views**	$S^w_{j,k}$
Econometrics - Free University Amsterdam	1	683	0.194	86.02	0.024
Econometrics - University of Amsterdam	2	428	0.121	65.59	0.019
Economics - Erasmus University Rotterdam	4	553	0.157	61.06	0.017
Engineering - Delft University of Technology	18	225	0.064	14.86	0.004
Medicine - Erasmus University Rotterdam	20	150	0.043	13.50	0.004

Note: *Times also viewed* indicates the number of times an individual visitor visited both the program page for Econometrics at Erasmus University Rotterdam and that of its substitute. $S_{j,k}$ displays the share of individuals who visit the substitute's program page, conditional on visiting the program page for Econometrics at Erasmus University Rotterdam. *Weighted views* penalizes a visitor's view by $\frac{1}{Views_i}$ (1 over the number of pages the visitor viewed in total). $S^w_{j,k}$ displays the weighted share of individuals who visit the substitute's program page, conditional on visiting the program page for Econometrics at Erasmus University Rotterdam. Rank is determined by $S^w_{j,k}$

Table B2: Similarity of programs and their substitutes

	Same Uni	Same Uni & Loc.	Same Degree	Same Uni & Loc. **or** Same Degree	$S_{j,k}$	$S^w_{j,k}$
Rank 1 (Constant)	0.420***	0.334***	0.436***	0.771***	0.176***	0.0260***
	(0.00999)	(0.00955)	(0.0100)	(0.00851)	(0.00147)	(0.000265)
Rank 2	-0.0639***	-0.0418**	-0.0516***	-0.0934***	-0.0346***	-0.00810***
	(0.0139)	(0.0133)	(0.0141)	(0.0127)	(0.00188)	(0.000308)
Rank 3	-0.0680***	-0.0340*	-0.118***	-0.152***	-0.0542***	-0.0117***
	(0.0139)	(0.0133)	(0.0138)	(0.0130)	(0.00179)	(0.000288)
Rank 4	-0.0717***	-0.0307*	-0.151***	-0.182***	-0.0680***	-0.0140***
	(0.0139)	(0.0133)	(0.0136)	(0.0131)	(0.00174)	(0.000280)
Rank 5	-0.0537***	-0.00451	-0.215***	-0.220***	-0.0755***	-0.0155***
	(0.0140)	(0.0135)	(0.0131)	(0.0132)	(0.00170)	(0.000275)
Rank 6	-0.0627***	-0.0209	-0.224***	-0.245***	-0.0839***	-0.0166***
	(0.0139)	(0.0134)	(0.0130)	(0.0132)	(0.00168)	(0.000272)
Rank 7	-0.0627***	-0.0164	-0.257***	-0.274***	-0.0879***	-0.0175***
	(0.0139)	(0.0134)	(0.0127)	(0.0132)	(0.00167)	(0.000270)
Rank 8	-0.0627***	-0.0164	-0.289***	-0.305***	-0.0938***	-0.0182***
	(0.0139)	(0.0134)	(0.0124)	(0.0132)	(0.00164)	(0.000269)
Rank 9	-0.0652***	-0.0160	-0.305***	-0.320***	-0.0979***	-0.0188***
	(0.0139)	(0.0134)	(0.0122)	(0.0132)	(0.00164)	(0.000268)
Rank 10	-0.0516***	-0.00943	-0.307***	-0.316***	-0.102***	-0.0193***
	(0.0140)	(0.0135)	(0.0121)	(0.0132)	(0.00163)	(0.000268)
Rank 11	-0.0664***	-0.0119	-0.333***	-0.345***	-0.104***	-0.0197***
	(0.0139)	(0.0134)	(0.0118)	(0.0131)	(0.00162)	(0.000267)
Rank 12	-0.0754***	-0.0254	-0.331***	-0.356***	-0.107***	-0.0200***
	(0.0139)	(0.0134)	(0.0118)	(0.0131)	(0.00161)	(0.000267)
Rank 13	-0.0656***	-0.0193	-0.345***	-0.364***	-0.110***	-0.0203***
	(0.0139)	(0.0134)	(0.0116)	(0.0131)	(0.00160)	(0.000267)
Rank 14	-0.0783***	-0.0336*	-0.359***	-0.392***	-0.112***	-0.0206***
	(0.0139)	(0.0133)	(0.0114)	(0.0130)	(0.00161)	(0.000266)
Rank 15	-0.0902***	-0.0398**	-0.359***	-0.399***	-0.114***	-0.0209***
	(0.0138)	(0.0133)	(0.0114)	(0.0130)	(0.00159)	(0.000266)
Rank 16	-0.0996***	-0.0533***	-0.359***	-0.413***	-0.115***	-0.0211***
	(0.0138)	(0.0132)	(0.0114)	(0.0129)	(0.00159)	(0.000266)
Rank 17	-0.107***	-0.0631***	-0.362***	-0.425***	-0.117***	-0.0213***
	(0.0137)	(0.0131)	(0.0114)	(0.0129)	(0.00158)	(0.000266)
Rank 18	-0.109***	-0.0631***	-0.370***	-0.434***	-0.119***	-0.0215***
	(0.0137)	(0.0131)	(0.0112)	(0.0128)	(0.00158)	(0.000266)
Rank 19	-0.0959***	-0.0537***	-0.372***	-0.426***	-0.121***	-0.0216***
	(0.0138)	(0.0132)	(0.0112)	(0.0128)	(0.00157)	(0.000266)
Rank 20	-0.117***	-0.0734***	-0.371***	-0.444***	-0.122***	-0.0218***
	(0.0137)	(0.0131)	(0.0112)	(0.0128)	(0.00157)	(0.000266)
Observations	48800	48800	48800	48800	48800	48800

Note: regressions estimated by ordinary least squares using the full set of identified substitutes. Columns display the dependent variables. The outcome variable Columns (1) to (4) is a dummy indicating whether the substitute at a certain rank is at the same university (1), at the same university and university location (2), the same degree program (3) and either at the same university and location or the same degree program (4). The dependent variables Columns (5) and (6) are the substitutability metrics, as explained in Section 3.3. Substitutes are ranked based on $S^w_{j,k}$. The number one ranked substitute is used as the baseline. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B3: Satisfaction scores & first-year enrollment over time

	(1) Enrollment	(2) Enrollment	(3) Satisfaction score
Satisfaction score of program (×10)	0.0159***		
	(0.00311)		
2012	0.00156	0.00360	-0.201***
	(0.00988)	(0.0103)	(0.0708)
2013	0.0751***	0.0729***	0.207***
	(0.0117)	(0.0122)	(0.0706)
2014	0.0298**	0.0361**	0.235***
	(0.0133)	(0.0140)	(0.0698)
2015	-0.00461	0.000143	0.528***
	(0.0163)	(0.0167)	(0.0700)
2016	0.0570***	0.0609***	1.179***
	(0.0162)	(0.0166)	(0.0687)
2017	0.112***	0.116***	1.465***
	(0.0172)	(0.0175)	(0.0684)
2018	0.140***	0.145***	1.552***
	(0.0180)	(0.0185)	(0.0676)
2019	0.151***	0.156***	1.259***
	(0.0174)	(0.0180)	(0.0675)
Satisfaction score: 3.6		-0.00447	
		(0.0222)	
Satisfaction score: 3.7		0.00994	
		(0.0219)	
Satisfaction score: 3.8		0.0191	
		(0.0231)	
Satisfaction score: 3.9		0.0392*	
		(0.0230)	
Satisfaction score: 4.0		0.0486*	
		(0.0248)	
Satisfaction score: 4.1		0.0790***	
		(0.0265)	
Satisfaction score: 4.2		0.116***	
		(0.0297)	
Satisfaction score: 4.3		0.146***	
		(0.0329)	
Satisfaction score: 4.4		0.133***	
		(0.0363)	
Satisfaction score: 4.5		0.228***	
		(0.0667)	
Constant			38.57***
			(0.0527)
Observations	12969	12241	13073

Note: regressions are at study program-year level. Columns (1) and (2) are estimated by fixed effects Poisson regressions. Column (3) by ordinary least squares. *Satisfaction score of program* denotes a program's own satisfaction score. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B4: Standardized satisfaction score, national average, satisfaction score of substitutes & first-year enrollment

	(1) W/o substitutes	(2) No. 1 subst.	(3) Top 3. subst.	(4) Top 5 subst.	(5) Top 10 subst.	(6) Top 20 subst.	(7) All substitutes (weighted)
Dep. var: **first-year enrollment**							
Satisfaction score of program (standardized)	0.0532***	0.0531***	0.0537***	0.0539***	0.0544***	0.0552***	0.0532***
	(0.00918)	(0.00916)	(0.00911)	(0.00913)	(0.00913)	(0.00915)	(0.00915)
National average (standardized)	-0.0375**	-0.0386**	-0.0332*	-0.0290	-0.0260	-0.0217	-0.0372**
	(0.0180)	(0.0181)	(0.0180)	(0.0180)	(0.0185)	(0.0182)	(0.0179)
Satisfaction Score Substitute(s) (standardized)		0.00296	-0.00883	-0.0174*	-0.0251**	-0.0413***	-0.00246
		(0.00711)	(0.00927)	(0.00998)	(0.0110)	(0.0147)	(0.0115)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12821	12821	12821	12821	12821	12821	12821

Note: regressions are at study program-year level, estimated by fixed effects Poisson regression. All measures of satisfaction scores are standardized. This means coefficients can be interpreted as the impact of an increase of one standard deviation. *Satisfaction score of program* denotes a program's own satisfaction score. *National average* denotes the national average of the same degree program at all universities that offer it. *Satisfaction score of substitutes* indicates the average satisfaction score of all substitutes that are ranked above the number indicated in Column titles (2) to (7). Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B5: Heterogeneity of relationship between satisfaction score, national average, satisfaction score of substitutes & first-year enrollment

	(1) Base result	(2) Undergraduate	(3) Master's	(4) Men	(5) Migration background
Dep. var: **first-year enrollment**					
Satisfaction score of program (×10)	0.0198***	0.0231***	0.00448	0.0220***	0.0151***
	(0.00329)	(0.00372)	(0.00514)	(0.00385)	(0.00440)
National average (×10)	-0.00732	-0.00640	-0.00303	-0.00209	-0.0114
	(0.00612)	(0.00699)	(0.0103)	(0.00667)	(0.00964)
Satisfaction Score Substitute(s) (×10)	-0.0315***	-0.0196	-0.0317	-0.0329**	-0.0815***
	(0.0112)	(0.0123)	(0.0274)	(0.0131)	(0.0171)
Year dummies	Yes	Yes	Yes	Yes	Yes
Observations	12821	8919	3902	12783	12793

Note: regressions are at study program-year level, estimated by fixed effects Poisson regression. All measures of satisfaction scores are multiplied by 10. This means coefficients can be interpreted as an increase of 0.1. *Satisfaction score of program* denotes a program's own satisfaction score. *National average* denotes the national average of the same degree program at all universities that offer it. *Satisfaction score of substitutes* indicates the average satisfaction score of the top 20 substitutes. Columns (1) and (2) show results for undergraduate and master's programs, respectively. Columns (3) and (4) take the number of first-year students who are men, and the number of first-year students who have a migration background as the dependent variable, respectively. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B6: Robustness check for size; relationship between satisfaction score, national average, satisfaction score of substitutes & first-year enrollment

	(1) Largest 95%	(2) Largest 90%	(3) Largest 75%	(4) Largest 50%	(5) Smallest 50%	(6) Smallest 75%	(7) Smallest 90%	(8) Smallest 95%
Dep. var: **first-year enrollment**								
Satisfaction score of program (×10)	0.0199***	0.0200***	0.0206***	0.0220***	0.0117***	0.0124***	0.0159***	0.0171***
	(0.00330)	(0.00332)	(0.00348)	(0.00410)	(0.00374)	(0.00336)	(0.00346)	(0.00331)
National average (×10)	-0.00737	-0.00735	-0.00646	-0.00538	-0.0142*	-0.00377	0.00146	-0.00347
	(0.00613)	(0.00622)	(0.00651)	(0.00758)	(0.00787)	(0.00717)	(0.00670)	(0.00641)
Satisfaction Score Substitute(s) (×10)	-0.0314***	-0.0311***	-0.0304***	-0.0295**	-0.0521***	-0.0618***	-0.0477***	-0.0386***
	(0.0112)	(0.0112)	(0.0115)	(0.0127)	(0.0177)	(0.0133)	(0.0117)	(0.0111)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12373	11832	10020	6757	6046	9420	11460	12133

Note: regressions are at study program-year level, estimated by fixed effects Poisson regression. All measures of satisfaction scores are multiplied by 10. This means coefficients can be interpreted as an increase of 0.1. *Satisfaction score of program* denotes a program's own satisfaction score. *National average* denotes the national average of the same degree program at all universities that offer it. *Satisfaction score of substitutes* indicates the average satisfaction score of the top 20 substitutes. Columns indicate which programs are included in the analysis. Standard errors are clustered at the study program level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.