Erasmus University Rotterdam – Erasmus School of Economics

Does career orientation guidance improve retention rates in tertiary education?

Abstract

Dropout and switch of tertiary education students is a long enduring problem in the Netherlands. Roughly 30% of higher secondary education graduates in the Netherlands drops out of tertiary education or switches to a different field of tertiary education. This paper investigates the effect of career orientation guidance at secondary education on retention rates in tertiary education. It uses administrative data of higher secondary education graduates from cohort 2004/2005 until 2013/2014 in the Netherlands. In a difference-in-differences design, it first shows that the introduction of a specific career orientation guidance package, Qompas, does not improve retention outcomes. Subsequently, exploiting shocks in intensity of use of this package, this study shows that providing more extensive career orientation guidance at secondary school significantly improves retention outcomes. The larger the shock in intensity, the larger the effect on retention outcomes. In addition, this study suggests that more extensive career orientation guidance improves enrollment rates in tertiary education among non-western immigrants, Havo students and females.

Peter van den Heuvel

347761

13-10-2016

Supervisor: Prof. dr. H.D. Webbink

Acknowledgments

This thesis would not have been possible without the help and support of a lot of people. First, I would like to express my gratitude to Prof. Dr. Dinand Webbink, for supervising this thesis, but for teaching me in the past year as well.

Further, I would like to thank the Directie Kennis at the Dutch Ministry of Education, Culture and Science for offering me the opportunity to do this research in the past few months. It has been an inspiring environment to finish my studies. Thank you to everyone I played table football with. I had lots of fun and it was helpful to restructure my thoughts.

I would also like to thank everyone at the cluster K2: *Kennisverbindingen en kenniseconomie*, for allowing me to be part of your cluster. It has been interesting to see and hear what subjects came across. Thank you for your curiosity and for being so dedicated. Special thanks to my supervisor at the Ministry, Marc van der Steeg, for all the time you devoted to think along with me and for being so enthusiastic. Due to all your ideas I thought I'd never be done, but your input really improved this thesis and kept me going. You have been a wonderful coach.

I would also like to express my gratitude to Robert Jan van Egmond and Boris Eustatia from Qompas, and to Arrian Rutten and Erik Fleur from DUO for putting together the data I needed and for being so enthusiastic about this study.

Then, I would like to thank Koen, for being an awesome sparring partner. Thank you for the good spirit, motivation and spot-on remarks.

This year could not have been like this without the unconditional support and motivation from my wife, Eva. Thank you for your patience and love when I was tired at night, or writing Stata-code in my head.

Finally, I would like to thank my parents, my sisters, and all friends who supported me in this process. Thank you for being so encouraging.

Table of Contents

1. Introduction 1
2. Literature 5
   2.1. Uncertainty at the moment of choice 5
   2.2. Predictors of study success and the quality of educational choice 9
   2.3. Career orientation guidance 12
   2.4. Other policies to reduce student dropout and switch 15
3. Brief background and data description 17
   3.1. The Dutch schooling system 17
   3.2. Qompas career orientation guidance package 17
   3.3. Data description 18
4. Empirical strategy 21
   4.1. Basic difference-in-differences approach 21
   4.2. Synthetic control method 22
   4.3. Variation in the year of introduction 23
   4.4. Shocks in intensity of use 24
5. Validity of the designs 26
   5.1. Basic difference-in-differences approach 26
   5.2. Variation in the year of introduction 28
   5.3. Shocks in intensity of use 29
6. Main estimates 32
   6.1. Basic difference-in-differences 32
   6.2. Variation in the year of introduction 34
   6.3. Shocks in intensity of use 35
7. Robustness checks 38
8. Discussion and concluding remarks 41

References 44
Appendix 54
Appendix A: Additional estimation results 55
Appendix B: Description of the covariates in the analyses 60
Appendix C: Descriptive statistics of the main models 63
Appendix D: Descriptive statistics of robustness check models 69

1. Introduction

A major problem in tertiary education is that students drop out completely or switch to a different field of tertiary education. This is perceived as a waste of government expenditures on tertiary education (see for instance OECD, 2008). Moreover, this is both discouraging and costly for these students themselves. Figure 1 presents dropout, switch, and retention rates of graduation cohorts 04/05 until 13/14 of Havo and VWO graduates in the Netherlands. This shows that this problem has been there for years and that the dropout (roughly 7%) and switch rates (roughly 23%) in tertiary education are relatively constant.[1] Secondary schools can affect the quality of educational choice of their students. Figure 2 presents the distribution of school fixed effects on the retention outcome of students, controlling for a large set of individual and secondary school characteristics, including cognitive test scores.[2] This suggests that secondary schools have a major impact on how well their students do once they enroll in secondary education. A student at a school in the lower quartile has a roughly 16% lower probability of retention in tertiary education than a student with similar characteristics at a school in the top quartile. A promising policy intervention is career orientation guidance, which aims to stimulate secondary education students to conscientiously make a decision on their field of tertiary education. Although this type of intervention is universally adopted, schools differ widely in the intensity with which this guidance is provided.[3] By guiding students through the transition from secondary to tertiary education, this intervention may contribute to equity[4] and to a better match between supply and demand on the labor market[5] (OECD, 2004).

Connecting data from Qompas, a provider of career orientation guidance teaching material, with administrative educational data, this study investigates the effect of providing career orientation guidance to secondary school students on retention rates in tertiary education. The Qompas package consists of several components that altogether aim to guide students in their tertiary educational choice.

---

[1] Dropout is denoted as 1 if a student is not enrolled in tertiary education one year after the first enrollment in tertiary education. Note that this does not necessarily mean that students stop attaining education. Students may e.g. attend upper secondary vocational education (MBO). Switch is denoted as 1 if a student is enrolled in a different field of tertiary education in the second year this student is enrolled, relative to the first enrollment information.

[2] Table C1, appendix C presents descriptive statistics of schools whose standardized fixed effect significantly predicts retention outcomes and is larger (smaller) than + (-) 1.5 standard deviation.

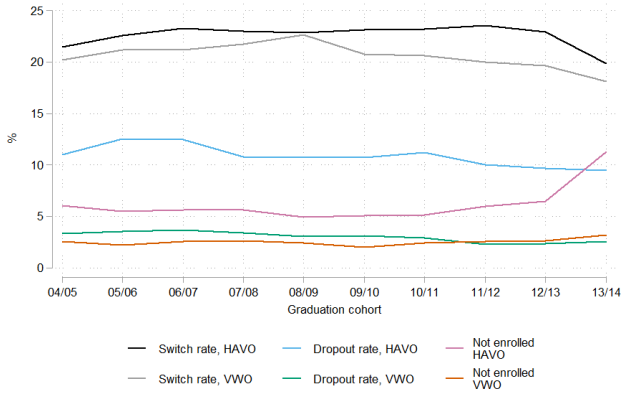
[3] From the Qompas data this study uses, it can be concluded that even within the sample of schools that invest in the Qompas teaching material for career orientation guidance, differences between schools in to what extent they offer guidance are large.

[4] Typically, students who are most aware of the value of career orientation guidance request this guidance the most. Thereby, disadvantaged students, who usually lack guidance from their social environment and therefore need this the most, are overlooked by the counselor. Providing effective guidance may contribute to bridge this gap.

[5] By providing objective information on labor market opportunities to secondary school students, their educational decision may be affected. It might help to have students shift towards choosing a field of study where labor demand is relatively high, thereby improving the match between supply and demand on the labor market on the long term (ROA, 2014).

FIGURE 1

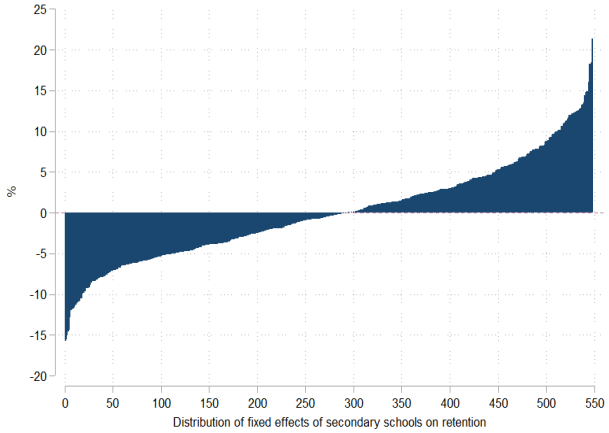
SWITCH, DROPOUT AND NON-ENROLLMENT RATES IN TERTIARY EDUCATION OF HAVO/VWO GRADUATES IN THE NETHERLANDS



Notes: This graph is based on the data this study uses. Dropout denotes students that enroll in freshmen year tertiary education, but do not return for sophomore year. Switch denotes students that enroll in freshmen year of tertiary education, and the subsequent year return for freshmen year of tertiary education in a different specialization. Students that not enroll in tertiary education at all constitute the third group.

FIGURE 2

DISTRIBUTION OF SECONDARY SCHOOL FIXED EFFECTS ON RETENTION OUTCOMES IN TERTIARY EDUCATION IN THE NETHERLANDS



Notes: This graph presents secondary school fixed effects in a regression with the retention outcome as a dependent variable, controlling for individual characteristics and secondary school characteristics exploiting data from cohorts 04/05 until 13/14. For a description of the control variables, see the data section on page 18.

First, the effect of the introduction of the Qompas career orientation guidance package at a secondary school on retention outcomes is estimated in a difference-in-differences framework. This study shows that the introduction of this package, relative to business as usual, does not improve retention rates. This effect is also estimated exploiting variation in the year of introduction of this package. Comparing schools that have introduced the Qompas package with schools that have not introduced the package yet, this design confirms that the introduction of this package, compared to business as usual, does not significantly affect retention outcomes.

Subsequently, this study exploits within-school shocks in intensity of use of the Qompas package to estimate the effect of providing extensive career orientation guidance on retention outcomes. Comparing schools that experience a sudden positive shock in intensity of use with schools that use the same package with a stable intensity in a difference-in-differences framework, this study finds that such a treatment is associated with a roughly 4 percentage point increase in retention rates. The threshold for the magnitude of the shock in intensity to be a major shock is an arbitrary choice. As a robustness check, several thresholds are adopted which suggests that the larger the shock in intensity, the larger is the gain in retention outcomes. In a similar design with the enrollment rate as an outcome variable, it seems that Havo students, females, and non-western immigrants more often enroll in tertiary education.

As career orientation guidance is perceived as a promising policy intervention, the policy debate on this topic has gained attention in recent years.[6] In the Netherlands, schools are not incentivized to improve career orientation guidance they offer to their students. Although the objective of secondary education in the Netherlands is to equip their students for tertiary education, information on how well they do this (in terms of e.g. tertiary graduation rates, retention rates) is hardly available. In addition, the Dutch Inspectorate of Education does not evaluate schools on measures like this. This lack of public information and (vertical) accountability illustrates that schools lack incentives to improve the career orientation guidance they provide. Moreover, tertiary education institutions, that may also affect the quality of educational choice, have an incentive to recruit as many students as possible as they receive funding per student. This may result in tertiary education institutions providing biased information on e.g. labor market opportunities to prospective students (see for instance UWV (2015) or van Nimwegen(2016)).[7]

Altogether, both secondary schools and tertiary education institutions are not incentivized to improve the quality of educational choice. As schools are not forced by law to provide career orientation guidance, it is up to the school board to decide whether to invest in such guidance (Oberon, 2012). In 2008, Kuijpers & Meijers show that career orientation guidance is typically given low priority by school boards. A survey report of Oberon (2012) indicates that career orientation guidance is gaining attention at schools.[8] However, incentives to stimulate this are not yet in place. By contrast, schools do have an incentive to improve the quality of profile choices of their students.[9] Borghans et al. (2008) state that schools in the first three years of secondary education in the Netherlands devote much more time to career orientation guidance than in subsequent years.[10] From Qompas data on their profile choice package (which is similar to the study choice package), it follows that on average, schools use the profile choice package more extensively than they use the study choice package.[11] This confirms the finding that secondary schools offer more guidance to their students when they

[6] See e.g. Onderwijsraad (2013), Ministry of Education, Culture and Science & Ministry of Social Affairs and Employment (2015), Ministry of Education, Culture and Science (2016), or LAKS, ISO, NVS-NVL, PvdA, VVD, CDA (2016).
[7] These studies show that specializations with weak labor market opportunities provide more vague, and sometimes even biased information about the labor market opportunities.
[8] This survey study shows that an increasing share of deans indicate that the school has some written vision on career orientation guidance, which is reflected in financial resources. Furthermore, it shows that a growing rate of schools indicates that its guidance activities are systematically evaluated. However, The Dutch Inspectorate of Education (2016) recognizes that a systematic evaluation of guidance activities at schools is lacking.
[9] In the Netherlands, after three years of secondary education (Havo/Vwo) students have to choose a profile for the rest of their secondary education years. Students typically stay on the same school after choosing a profile. Therefore, if students would lack guidance and fail to make a good profile choice, this may result in e.g. discouraged students, study delay, or angry parents. Accordingly, schools have an incentive to prevent this, e.g. via well-organized career guidance.
[10] In line with this, a survey study examining student satisfaction on career guidance shows that 70% of VWO students and 50% of Havo students is satisfied about the transition from secondary to tertiary education (ResearchNed, 2016).
[11] This can be illustrated by juxtaposing the percentage of students that made a profile choice test (66%) and a study choice test (52%).

need to choose a profile, but that this guidance is provided less often students need to decide on their field of tertiary education.

Unfortunately, little is known about the effectiveness of career orientation guidance. Most literature focuses on enrollment rates, thereby overlooking that better guidance may improve the quality of educational choices which may lead to improved retention rates. This study adds to the existing literature by providing an insight into the effect of career orientation guidance at secondary schools on retention outcomes in tertiary education. A unique aspect of this study is that it matches administrative data on retention outcomes in tertiary education with career guidance data at secondary education. The main finding is that providing more extensive career orientation guidance at secondary school is effective in improving retention outcomes. Being a pupil in a school in the top 20 percent of growth in intensity causes an increase in the retention probability of 4 percentage points, on a base of 68%. This comes down to a 12.5% increase in a student's probability of retention in tertiary education. The larger the growth in intensity, the larger is the gain in retention outcomes.

The rest of this paper is organized as follows. The next section describes the relevant literature on career orientation guidance and tertiary educational choices. Then, some background information on the Dutch schooling system, the Qompas package, and the other data is presented. The fourth section discusses the three empirical designs in detail, as well as the econometric methods that are used. The validity of these designs will then be scrutinized. The subsequent section presents the main estimates, followed by some robustness checks. The last section discusses these results and concludes. Some additional, detailed information can be found in the appendix.

2. Literature

In this section, the relevant literature regarding career orientation guidance and educational choices will be discussed. First, the uncertainty issues high school student face will be discussed. Next, the variables that predict the retention rate will be explored. Furthermore, the literature on the effectiveness of career orientation guidance and related policy interventions will be summarized.

2.1. Uncertainty at the moment of choice

The decision making process of secondary school students regarding a field of education, is surrounded by uncertainties. This section explores the literature on uncertainty regarding objective (e.g. labor market opportunities, costs of college) and subjective information (e.g. what field of study would suit best) high school pupils face at the moment of choice.

2.1.1. Lack of information on pecuniary costs and benefits to education

In the economic literature, the traditional approach to explain choices people make is to explore the costs and benefits of several choice options. These costs and benefits can be individual-specific and depend on the preferences of this individual. Before choosing the field of study, one needs to decide whether or not to enroll in tertiary education. This leads us to the first information imperfection, being the uncertainty on the costs of higher education. Horn, Chen & Chapman (2003) show that secondary school pupils and their parents are largely uninformed about the costs of postsecondary education.[12] This lack of cost-awareness is larger among parents who did not attend tertiary education themselves and among low-income families. Those who report that they did obtain information on this tend to overestimate these costs. Other studies suggest that not only costs of higher education are overestimated by prospective students, but wage benefits of higher education are overestimated as well (Avery & Kane, 2004).[13] The expectation of pecuniary returns to higher education is affected by various factors, including the probability of completing postsecondary education, the expectation of wages, the probability of finding a job and the volatility of the labor market. The literature examining these factors, especially its impact on educational choices, is extensive. Beffy et al. (2012) provide evidence that expected earnings

[12] Horn, Chen and Chapman (2003) scrutinize data from the Parent and Youth Surveys of the 1999 National Household Education Surveys Program in the U.S. They investigate information students in grades 6 to 12, that report they plan to attend postsecondary education, and their parents have on the costs of attending college.

[13] In contrast with other studies, they do not find differences on the perceptions of costs and benefits of college education between socio-economic groups. They suggest that the complex process of applying for financial aid and college admission negatively affects the enrollment of low-income students, while others argue the enrollment gap is the result of differences in these perceptions.

play a small, though significant role in educational choices.[14] Other studies try to explain the observation that enrollment rates among students from high-income families are higher than those of students from low-income families. Betts (1996) finds that the expectations of returns to education of students with relatively poor parents are significantly lower than the expectations of those with richer parents.[15] Streufert (2000) suggests this can be explained by the fact that young people form expectations of returns to education based on the observation of employees in their neighborhood. Therefore, pupils that grow up in a low-income neighborhood might have lower expectations of returns to education, which affects their educational decision. This research suggests that providing better or more information on returns to education may improve educational choices of pupils from lower socioeconomic backgrounds. This hypothesis is confirmed by results from randomized experiments in developing countries, showing that expected earnings play a large role in education decisions (Jensen, 2010[16]; Nguyen, 2008[17]). Students in such countries largely underestimate the expected returns, implying that providing information on the returns to education to these students leads to a large increase in enrollment and makes them less likely to dropout. In line with these studies, Oreopoulos & Dunn (2013) show that providing simplified information to students about post-secondary education makes students more likely to request further information.[18] Their results stress the point that *how* information is provided matters, which is in line with insights from the Behavioral Insights Team (2016). In addition, they show that disadvantaged students update their prior beliefs after obtaining new information, suggesting that closing the information gap may be an effective way to close the enrollment gap. By contrast, Kaufmann (2014) concludes that information gaps are not the main cause of enrollment gaps between rich and poor.[19] She shows that students from low-income families

[14] They identify the effect of expected earnings by exploiting variation in in relative wage returns of various majors at French universities due to business cycle fluctuations.

[15] This paper uses survey data on undergraduates' knowledge of salaries by type of education. The survey was carried out among 1269 undergraduates at the University of California, San Diego.

[16] Jensen (2010) studies a survey among 2,250 boys in the final year of primary school in the Dominican Republic to examine the perceived returns to secondary education. The researcher compares this with survey data of 1,500 households to estimate the actual returns to education. They conduct a randomized experiment in which students are provided information on the returns to schooling.

[17] Nguyen (2008) examines a field experiment in Madagascar in which information about the returns to education is provided at school by either a role model or by providing statistics (or a combination of both). They find that providing statistics is highly effective in updating beliefs about the returns to education, whereas the role model intervention was less effective. The role model intervention did lead to an increase in schooling effort of those children whose family background matches that of the role model. Furthermore, the study shows that providing statistics on the returns to education leads to an increase in schooling attendance and test scores. Program costs are only $2.30 for an additional year of schooling, implying that this is a highly cost-effective intervention.

[18] This follows from a field experiment among five Toronto public schools in disadvantaged neighborhoods. They targeted students who were not sure to enroll in postsecondary education. In the experiment, students would have to fill in a survey asking basic demographic characteristics. After the survey, half of the students are shown a video promoting college and university, informing about financial aid possibilities, and suggesting that many students overestimate costs. Furthermore, they could easily calculate the financial aid they would be eligible for. Three weeks later, this group reported a higher likelihood of college attainment and was more likely to request further information.

[19] This study uses a survey among 23,000 Mexican adolescents that are in their last year of junior high school or are attending high school. It collects information on their subjective expectations of returns to education and labor market opportunities.

are responsive to changes in direct costs, whereas expected returns to college of these students are similar to college attending students. This suggests that differences in enrollment decisions between rich and poor are, for a substantial part, caused by credit constraints. In accordance with this, Booij, Leuven and Oosterbeek (2012) show, in a randomized experiment, that increasing the amount of information on student loans does not lead to more take-up of these loans. This suggests that a lack of information does not necessarily imply a binding information constraint. Another explanation of the enrollment gap is that students from high-income families are less risk averse, as they can bear more risk, than their peers from low-income families. This would be in line with Brown et al. (2011), who find that risk averse students are less willing to take on debt, suggesting they have less willingness to invest in uncertain outcomes.

2.1.2. Risk aversion

The degree to which uncertainties regarding future income are important for prospective students depends on their risk aversion (Dohmen & Non, 2014). First, students face a dropout risk, which increases when the level of difficulty increases. De Paola & Gioia (2011) show that risk averse, high-ability students often choose a field of study with good labor market opportunities, whereas similarly risk averse, but low-ability students often choose a program of study which is less difficult.[20] Furthermore, this suggests that labor market risks, that are affected by the degree of difficulty of education attained, affect secondary school students' decision on a field of study. However, learning about grade performance and academic ability is an important factor in explaining college dropout (Stinebrickner & Stinebrickner, 2012). On average, students are too optimistic at the time of entrance about their ability to graduate. Therefore, learning about academic ability and grade performance might result in students dropping out early because their perception of their own ability turns out to be flawed. This suggests that providing more reliable information about grade performance and academic ability before students enroll may culminate in lower enrollment rates, but in lower dropout rates as well. Moreover, Nielsen & Vissing-Jorgensen (2006), show that students take into account labor market risks when choosing their field of study.[21] By contrast, other studies,

[20] The researchers use data on 3,661 first-year students of the University of Calabria. This data set includes information on cognitive abilities, personality, socioeconomic background, and a large number of individual characteristics. Furthermore, it includes an indication of risk aversion, based on student answers to questions on the willingness to invest in certain assets.

[21] The researchers show, examining a Danish data set on labor incomes and educational choices, that people prefer educations with high mean incomes and low risk. They find that labor income risk significantly affects educational choices.

conducted in Italy, find that parents' and students' risk aversion is no important predictor of educational choices (Leonardi, 2007; Belzil & Leonardi, 2007).[22]

2.1.3. Uncertain quality of education

Another uncertainty prospective tertiary education students face is regarding the quality of schooling. Research suggests that the information students have on the quality of various institutions does play a role in their educational decisions. Monks and Ehrenberg (1999) show that a lower ranking on a publicly available scoreboard reflecting school quality leads to a lower quality of the entering cohort of students. Meredith (2004) adds to this that this effect is only valid at highly ranked private institutions. Griffith and Rask (2007) find consistent results, showing that especially high-ability students are responsive to changes in public rankings of school quality. Obviously, these school rankings are a subjective and imperfect measure of school quality, implying uncertainty regarding this factor in educational choices.

2.1.4 Identity-related uncertainty

Another piece of literature focuses on the non-pecuniary returns to education. This includes utility payoffs in terms of enjoying a field of study more than the other. Secondary school students often have vague ideas of what they like and what they do. Combined with uncertainty of what kind of tasks are involved in various professions, this creates difficulties regarding educational choices. Akerlof & Kranton (2000) offer an economic model of behavior which adopts insights from psychology and sociology on identity. They argue that payoffs of one's choices are based on one's identity. Furthermore, they argue that the choice of identity may be the most important economic decision people make, as it influences one's behavior. However, the choice of identity can be influenced by others.[23] The educational choice of the type of tertiary education reflects a pupil's self-image. It may be that better career orientation guidance leads to a more conscious, and thereby more persistent, choice of identity and self-image. Based on this model, Favara (2011) argues that pupils choose their field of education conforming to not just their expected monetary returns, but also to their payoff based on their identity. Conforming to the social norms of their identity, various educational choices yields various levels of utility for them.[24] Humlum, Kleinjans & Nielsen (2012) study the role of non-pecuniary identity-related returns to education in the tertiary

[22] Both studies use individuals' answers on lottery questions to obtain a degree of risk aversion. They use sample sizes of 867 individuals and 8,135 households respectively.
[23] E.g. parents choosing a certain school in order to influence a child's self-image and identification with a certain socio-economic group.
[24] In addition, this study suggests that this effect is especially important for girls as they follow their talents less, and social stereotypes more, than boys in the educational choices they make. However, both boys and girls are affected by social norms in terms of traditional stereotype male or female professions.

education decision of Danish youth.[25] They find that for females, a higher career factor implies a higher level of education and a higher social factor decreases this level. Furthermore, they find effects of both the social and the career factor for both males and females on the field of education. This suggests that identity-related non-pecuniary returns to education play an important role in education choices

.

2.2. Predictors of study success and the quality of educational choice

In the academic literature, college enrollment and retention is an important subject. This section explores the variables that typically predict outcomes such as enrollment, retention and graduation. These variables can be divided into two groups: individual characteristics and institutional characteristics. Besides evidence that such variables directly affect study success related measures, both characteristics affect study success indirectly via student satisfaction (Umbach & Porter, 2002).

2.2.1. Individual characteristics

First, the academic literature suggests, although it is not univocal, that gender is an important predictor of study success, in terms of persistence and graduation. Scott et al. (2006), Astin (1996) and Kuh et al. (2008) show that a higher share of female students is positively associated with several measures of student persistence.[26] However, Astin (1996) finds that gender differences decrease over time. Several other studies find no significant impact of gender on student retention (e.g. Reason, 2003; Hu et al., 2000). Mellanby, Martin & O'Doherty (2000) suggests that gender differences in graduation rates can be ascribed to a combination of psychological characteristics and its interaction with the demands of an individual academic assessment system. Another explanation of females being more persistent in college is that females more consciously choose their field of study, as they come out of puberty sooner. Furthermore, traditional, stereotype studies of females are much more homogenous, whereas male students choose from a broader range of fields of studies.

Second, Hu et al. (2000) find a negative relation of age and persistence in higher education. In their sample of high school averages, the older the first year students are, the

[25] The authors combine register data with a follow-up survey which provides information on the planned level and field of education and on the attitude towards education, work life and social issues. From this, they use factor analysis to divide one's identity in a social factor (reflecting one's social and cooperative attitude to education and life in general) and a career factor (reflecting that career and promotion opportunities are an important part of life). These factors should reflect one's identity. Subsequently, they use a logit and a multinomial logit model to estimate the effect of these factors on the choices of field of tertiary education.

[26] Scott et al. (2006) find that graduation rates are 0.12% higher if the share of females is 1 percentage point higher. Kuh et al (2008) find that for females, the odds of persisting to the second year are 1.65 larger than for males.

lower the probability of retention is. This pattern is confirmed by Scott, Bailey and Kienzl (2006) who find a significant negative effect of age on graduation rates.[27]

A third predictor of study success is the socioeconomic background of students. Avery, Howell and Page (2014), in a literature review, point out that those who need college counseling the most, students from low-income families, have the least access to counselors. This view is confirmed by Scott, Bailey and Kienzl (2006), who show that the percentage of minority, foreign and financial aid receiving students is negatively correlated with graduation rates. In line with these findings, Calcagno et al. (2008) find a negative association between the proportion of minority students and educational success. Such results are consistently found in several studies (e.g. Kuh et al., 2008; Titus, 2006). Furthermore, if these students find their way to a college counselor, the focus of this session is generally on college affordability, instead of on exploring which field of education would suit this student best (Bridgeland and Bruce, 2011a). They find that a lack of funding and a lack of time and training of counselors are the most important constraints to an improved accessibility of college counseling. These constraints are most severe at lower-income schools (e.g. Perna et al., 2007). Lapan and Harrington (2010), in a study of Chicago public schools, show that especially the administrative burden and a lack of clarity on the role of a college counselor are time-consuming. This is consistent with a survey study among over 5,300 counselors aimed to identify problems college counselors face (Bridgeland and Bruce, 2011b) and with studies evaluating student satisfaction regarding college counseling (Johnson et al., 2010). However, the most important reason for a lack of college counseling among low-income and minority students is that they lack college going peers. Moreover, quite some evidence is presented in the literature, suggesting that high-achieving, black students underachieve in order to be accepted by the group (e.g. Austen-Smith and Fryer, 2005; Fryer and Torelli, 2010). All this evidence suggests that one's socioeconomic background is an important predictor of one's study success.

Adelman (2006) shows that a high GPA is positively correlated to graduation and is inversely related to dropout. This finding is consistent with several studies showing pre-college GPA is positively related to student persistence (e.g. Hu et al. (2000); Kuh et al. (2008); Cabrera, Burkum & La Nasa (2005)). In line with these findings, Adelman (1999)

[27] The magnitude of this age effect varies substantially. Whereas Hu et al. (2000) find that a one year older student has a 10 percentage point lower probability of persisting to the second year, Scott, Bailey and Kienzl (2006) find that 1 year older lowers expected graduation rate by only 1%.

finds that finishing a course beyond the level of Algebra 2 in secondary school more than doubles the probability of earning a bachelor's degree.

2.2.2. Institutional characteristics

Bailey et al. (2006) and Titus (2004) show that college graduation rates between community colleges vary significantly, suggesting that institutional characteristics matter for the study success of its students. Calcagno et al. (2008) examine the impact institutional characteristics have on the success of community college students in the US. Although they do find that individual characteristics are more important predictors of study success, they show that the size of an institution, measured as the number of students, is inversely related to the study success of its students. This finding is in line with results of Bailey et al. (2006) and Pascarella and Terenzini (2005). By contrast, Titus (2004) finds that institutional size of 4-year colleges in the US is positively related with student persistence, and Scott, Bailey and Kienzl (2006) find a positive relation between the number of undergraduates and graduation rates. However, the magnitude of this effect tends to be small in each of these studies.

Furthermore, Scott et al. (2008) find a significant negative correlation between student/staff ratio and retention rates, using data from 38 Australian universities. Findings from a survey-based, longitudinal study in the US confirm this inverse relationship (Bound, Lovenheim and Turner, 2010). Tinto (2002) offers an explanation for this observation. He argues that an important condition for student retention is that students are socially and academically involved. The more students are integrated in the community, the less likely they are to dropout. Social integration involves conversation with faculty and staff members, implying that a lower student/staff ratio improves the accessibility of staff members, thereby improving social and academic integration.[28] This view is supported by empirical evidence of Scott, Bailey and Kienzl (2006), Chambers (2009) and Hu et al. (2000).[29] In line with this research, Calcagno et al (2008) find that a more professional environment[30] and a more personal approach is positively related to educational success. In addition, Kuh et al. (2008) finds, using data from the annual NSSE survey of undergraduate students measuring participation in educationally purposeful activities, that student engagement has a positive effect on student persistence. In the same vein, Scott, Bailey and Kienzl (2006) show that

[28] This is in line with findings from a randomized controlled trial of an intervention aiming to foster freshmen's sense of social belonging in school. This small intervention improved students' performance, especially for students from minority backgrounds (Walton & Cohen, 2011).

[29] These studies use the percentage of commuters as a proxy of social attachment, and examine the correlation of this percentage with educational outcomes, such as graduation rates. They find an inverse relationship between the two variables.

[30] The proportion of full-time staff is used as a proxy for the professionality of the study environment.

instructional expenditures per student are positively associated with graduation rates, which is consistent with findings from studies such as Titus (2006).

2.3. Career orientation guidance

Several policies have been conducted aiming to reduce dropout and switch rates. This study aims to investigate the effectiveness of providing computer based career orientation guidance to secondary school pupils. In the academic literature, the effect of career orientation guidance is extensively discussed. However, most of this literature discusses its effect on enrollment in postsecondary education only. For instance, Hurwitz and Howell (2013) exploit between-state differences in maximum student/counselor ratios in a regression discontinuity framework. They show that hiring an additional high school counselor leads to a significant increase in four-year college enrollment. This finding is consistent with several other studies suggesting that counseling and other exploratory activities may be effective in increasing enrollment rates (e.g. Bryan et al., 2011; Plank & Jordan, 2001; Belasco, 2013; Bos et al, 2012).

Hoest, Jensen & Nielsen (2012) review the effects of the introduction of student career guidance in Denmark. This reform entailed the introduction of centralized career guidance at lower secondary school (mainly carried out at ages 14-17) in order to improve the enrollment in upper secondary school. They find, using a difference-in-difference strategy with private schools as a control group, that this policy improves enrollment rates of immigrants and that the effect on enrollment rates of native born students is small. This gives rise to the idea that some groups might be affected more than others by such policies. This suggests that a targeted approach towards disadvantaged students, who might be affected most by career guidance policies, may be effective. Seftor, Mamum & Schirm (2008) examine the long term impact on postsecondary outcomes of a program aiming to improve skills and motivation for education beyond high school among randomly assigned disadvantaged students. They find no impact on the enrollment decision, although enrollment does increase among those who, in advance, did not expect to complete a bachelor's degree. Interestingly, the postsecondary graduation rate does increase, especially for those who were in the program for a longer time.[31] In addition, Hoxby & Turner (2013) evaluate data from a randomized controlled trial of an intervention that provides students with semi-customized information on the application process and the costs of college. They show that high-achieving, low-income students gain

[31] The program, the Upward Bound program, is targeted towards students from low-income families or first generation college students (students whose parents did not attend tertiary education) . These students are provided counseling, instruction and tutoring at high school. They estimate the impact on postsecondary outcomes 7 to 9 years after high school graduation.

from this intervention in the sense that they apply for, and are admitted to, both more and better colleges. Without the intervention, high-achieving, low-income students apply for both less and less selective colleges than their higher-income peers. This is in line with Avery (2013), who concludes, using a randomized controlled trial, that a telementoring program is effective in increasing the number of applications among low-income, high-achieving students for more competitive colleges. However, the effect on enrollment in these colleges this study finds is not significant. This might be explained by the fact that a substantial part of the students that are offered the program refused this offer. Other studies evaluating such targeted programs confirm the hypothesis that career orientation guidance is especially effective among disadvantaged students (e.g. Castleman & Goodman, 2014[32]; Oreopoulos, Brown & Lavecchia, 2014[33])

Carrell & Sacerdote (2013) evaluate a program which involves mentoring by a college student and assistance with college transition and applications which is offered to high school students. They find significant positive effects of being in this program on women's enrollment. The marginal women that enroll are as likely to persist in college as the control average. Furthermore, they find that cash incentives do not have an additional effect. Castleman, Page & Schooley (2014) evaluate two randomized trials of interventions in the summer after secondary school graduation, intending to improve enrollment rates among low-income students. College intending graduates are offered summer support, which has a large impact on enrollment rates among this subgroup. A more cost-effective intervention aimed to improve enrollment rates is evaluated in a later study (Castleman & Page, 2015). In this intervention, text-messages are sent to graduated high school students to remind them of key college-related tasks to successfully apply for college. They show, using a randomized controlled trial, that this significantly improves enrollment rates, especially among students with low access to college planning support. Furthermore, they evaluate a peer mentoring intervention, in which peers provide information on the application process and encourage to apply for college. This intervention has a sizable impact on enrollment rates, especially among students with less-defined college plans and less access to college planning support. This confirms findings from Bettinger et al. (2012), who evaluate the H&R block FAFSA

[32] Castleman & Goodman (2014) evaluate a program providing enrollment aid to high school students and, once they enroll in college, persistence aid. Exploiting a discontinuity in the accessibility of this program, they show that enrollment and persistence among low-income students with a GPA close to 2.5 are affected by the program. They enroll more in colleges that are encouraged by the program. Unfortunately, they lack power of data to assess whether they persist in college longer.

[33] Oreopoulos, Brown & Lavecchia (2014) evaluate the 'Pathways to Education' program in Canada, which provides proactive mentoring to students, group activities, career counseling and college transition assistance. In addition to this, the program incentivizes students to bolster a minimum degree of mandatory participation. In a difference-in-differences setup, they show that the program is highly effective in boosting high school completion and postsecondary enrollment rates.

experiment. This is a randomized controlled trial which took place in the US, involving the provision of information among low-income individuals on eligibility and the assistance to complete the application process for a college program. They show that such a program substantially improves college enrollment rates among the targeted group. These results are in line with findings from Dinkelman & Martinez (2011) who investigate a randomized experiment in Chile in which low-income Chilean adolescents are provided information about how to finance higher education. Better information on such issues raised college enrollment. Providing information to parents did not aggravate these effects significantly.

All these studies examine the effect of career orientation guidance on enrollment rates. Unfortunately, a limited number of studies investigates the impact of career orientation guidance on graduation rates and persistence. A few exceptions study the impact on the quality of study choice. One of these exceptions is Borghans et al. (2015), who show that an individual meeting with a study counselor at secondary school leads to a significantly better self-assessed educational choice.[34] As conventional OLS estimates of the effect of individual study counseling suffer from an endogeneity bias[35], they explore variation in the counseling practices between schools in an IV setting. They find large and significant effects of individual counseling at secondary schools on the satisfaction with the educational choice. This effect is especially pronounced for pupils with lower educated parents and for boys. In the same vein, Bettinger & Baker (2014) study the effects of student coaching on graduation rates. They randomly assigned students to be coached. The aim of this coach was to develop a clear vision of students' goals, guide them to connect long term goals to daily activities and to support students in building study-relevant skills. The researchers show that this intervention significantly improves retention and completion rates.[36] However, the coaching in this intervention is provided to university students, and not to secondary school students. Another study that does investigate the effect on college completion is Cunha & Miller (2009), evaluating the Texas GO Center project. This project, which is run by student peers, aims to create a college-going culture at secondary school. In addition, high school students are provided information on college choice and on several practical issues like the application process or financial aid. In a difference in difference setup, the researchers find large effects

[34] This study uses survey data from the Netherlands, which provides information on the self-assessment of students. 18 months after secondary school graduation, students report whether they would, in retrospect, choose the same education as the one they have chosen. Furthermore, students are asked how often they had personal conversations with their study counselor.
[35] The authors argue that more intelligent (uncertain) pupils might seek more help from counselors and make better (worse) educational choices, implying an overestimation (underestimation) of the effect of study counseling using OLS estimation techniques.
[36] The magnitude of the intention to treat effect they find varies between 4.3 and 5.3 percentage points.

on college applications and acceptations. However, they do not find any impact on college completion.[37]

In conclusion, the literature examining the effects of career orientation guidance on the quality of tertiary educational choice is scarce. This study aims to fill this gap in the academic literature by combining administrative data on tertiary education outcomes with administrative data on the intensity of career orientation guidance.

2.4. Other policies to reduce student dropout and switch

Besides the provision of better career orientation guidance, several other policies have been conducted to reduce dropout rates and increase graduation rates. Such policies consist of policies (1) 'at the gate', selecting students that apply for tertiary education before they start their studies; and (2) 'after the gate', selecting students after they start their studies with e.g. performance based selection systems.

2.4.1. At the gate

Some institutions try to improve retention rates by using admission criteria at the gate (e.g. GPA of undergraduate programs, standardized test scores) (Hagedorn and Nora, 1996). In this manner, institutions aim to select those students that have the highest odds of retention and graduation, in order to limit dropout rates. Obviously, this policy requires both knowledge on predictors of study success and a procedure to observe these predictors.

2.4.2. After the gate

Another often used way to increase graduation rates is known as selection after the gate, by introducing performance-based selection systems. Obviously, such policies are not aimed at reducing first-year dropout rates, as they force some students that do not achieve the minimum standards to dropout. Such policies aim to foster study effort of first-year students, create an ambitious study culture and thereby increase study speed (OCW, 2011). This might result into higher graduation rates. However, such extrinsic incentives to perform in the first year might undermine intrinsic motivation (Bénabou & Tirole, 2003).

Arnold (2015), in a study of data from 450 Dutch bachelor programs, shows that an academic dismissal policy can be an effective policy to increase study speed of motivated and

[37] The program did not offer guidance once students enrolled in college. The authors suggest that this is the reason of the absence of college completion effects. Those who are induced by the program to enroll in college might be low-performing with a lack of academic skills. Without guidance once enrolled, they argue, they may be unable to complete college. Another explanation for this might be that the estimate is the lower bound of the treatment effect. The authors do not know whether an individual was treated, but whether a GO Center was installed at the school this individual attended. In line with this, they find larger effects for low-income students and Hispanics, who are more likely to be treated.

talented students and thereby increases four-year completion rates. Furthermore, dropout rates are not decreased, but student dropout takes place earlier in time. In another study, he suggests that institutions should do more to guide dismissed students to other, better suiting, types of education (Arnold & van den Brink, 2010). In addition, Sneyers and De Witte (2015) show that the introduction of an academic dismissal policy leads to a higher first year dropout rate, a higher graduation rate, a reduction in student satisfaction and an increase in student satisfaction on program feasibility. Using a difference-in-differences estimator, they compare programs of study in the Netherlands that introduced an academic dismissal policy with programs which did not. They find that the introduction of this policy results in a higher first year dropout rate and a higher graduation rate.

In contrast, a study comparing cohorts of medical students in the Netherlands with and without an academic dismissal policy finds no effects of this policy on dropout rates, on the timing thereof and on completion rates (Stegers-Jager et al., 2011). This might be explained by the fact that medical schools in the Netherlands use selection at the gate as well. Those who are admitted to medical school are supposed to be a selective group of motivated and talented students. This study suggests that after introducing selection at the gate, selection after the gate generates no additional benefits. De Koning et al. (2014) suggest that the introduction of an academic dismissal policy does not lead to increased learning activities. They argue that students rather try to meet the minimum criteria than learn as much as possible.

Finally, tertiary education institutions can adopt several interventions to improve retention rates of their students. Schippers et al. (2015) show, in a difference-in-differences framework, that an online intervention aiming to boost goal-directed conceptualization and action among college freshman improves both student achievement and retention rates, especially among males and ethnic minority students. Furthermore, Grodner & Rupp (2013) show, in a randomized controlled trial, that introducing compulsory in-class homework sessions improve both retention rates and grades.

3. Brief background and data description

3.1. The Dutch schooling system

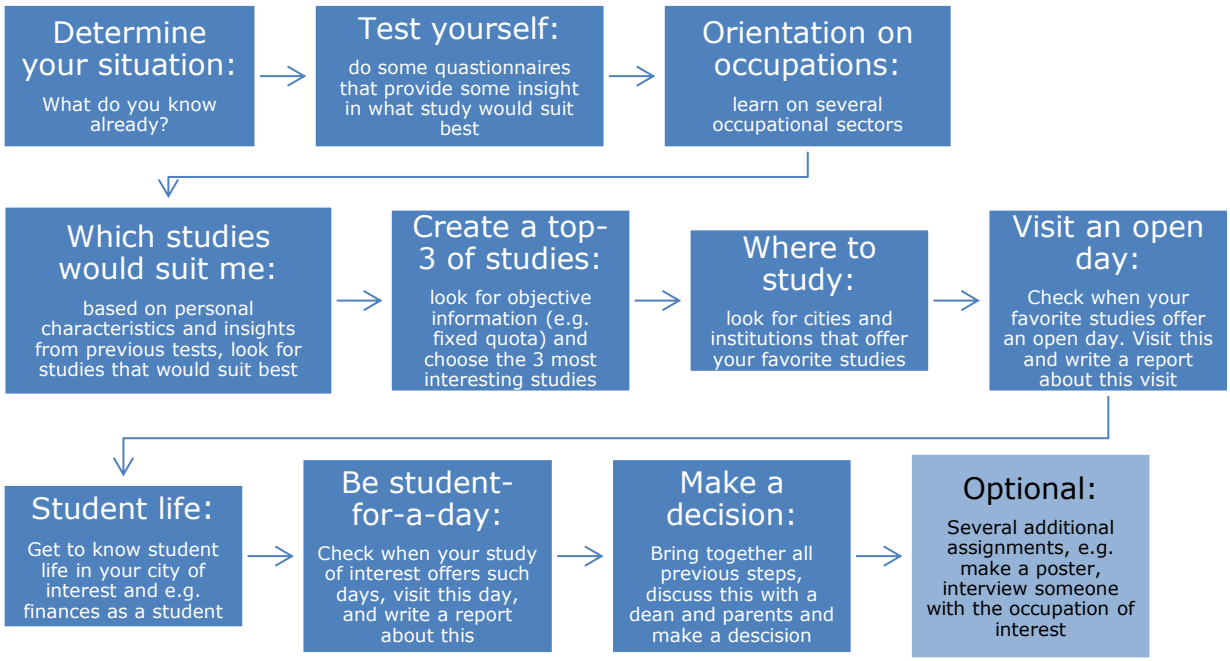
In the Netherlands, students are tracked into 3 different levels of secondary education at age 12. Students may enroll in (1) pre-vocation education (VMBO), lasting four years; (2) general secondary education (HAVO), lasting five years; or (3) pre-university education (VWO), lasting six years. In 2015/2016, respectively 28% and 19% of Dutch 16-year-old students were enrolled in HAVO and VWO (in the remainder, this will be referred to as higher secondary education). At age 17 or 18, higher secondary education graduates have to decide on the field of specialization in tertiary education. A HAVO diploma grants access to University of applied sciences (HBO), whereas a VWO diploma grants access to both Research University (WO) and University of applied sciences. HBO programs focus on learning skills that are relevant on the labor market, whereas a WO program is more research-oriented.

3.2. Qompas career orientation guidance package

Qompas is a for-profit organization offering a computer based package of career orientation guidance. Schools can buy licenses for their students which grants access to this package. With this license, each student can access this online teaching method. Once logged in, students can fill in several questionnaires, which aim to provide insight into occupations and fields of specialization that would suit them, in ones competences and weaknesses, and in ones fields of interest. Furthermore, students can easily find objective information on several fields of study (e.g. which studies have academic dismissal policies or selective admission standards, when tertiary education institutions have open days or when they offer the opportunity to be student for a day). Students can save e.g. fields of studies or open days in their portfolio, to easily access that information afterward. Altogether, the package covers about 11 teaching hours. All these activities are offered in a step-by-step plan, which is visualized in figure 3.

FIGURE 3

VIZUALIZATION OF THE STEP-BY-STEP PLAN OF THE QOMPAS PACKAGE



The dean can monitor the progress of these students. In addition, the package provides additional assignments to help students shape their own ideas about their study choice (e.g. do a game, make a poster, do an interview with someone with the occupation of interest). In response to the results of all these tests and assignments, the package stimulates students to discuss these results with a dean, counsellor or with friends and family.

For graduation cohort 2013/2014, 28% of the schools in the Netherlands use this teaching material to provide career orientation guidance to their students. However, how pervasive this material is used varies widely between schools. Some schools offer the license to their students and from that point, it is the students' responsibility whether they actually use it. If they would like to discuss it, it is up to the student to make an appointment with the dean. By contrast, other schools use this teaching material in-class, in which students are guided through all steps the program provides. In addition, the counsellor monitors the progress of the students and discusses this with them.

3.3. Data description

This study uses administrative data on all higher secondary education graduates in the Netherlands from 2004-2014, which is obtained from the Dutch executive organization responsible for education (DUO). The data includes several individual background

characteristics and secondary school characteristics. In addition, it includes information on their higher education career, including information on enrollment in the freshman year, on retention to the sophomore year, and, for the early cohorts, on graduation.[38] This is complemented with career orientation guidance data from Qompas, which provides information on whether or not a school has purchased the career orientation guidance package. Moreover, for 55% of this data, it provides an indication of the intensity with which these schools adopt this package.[39]

In order to add control variables as a proxy of socioeconomic background, Leefbaarometer data is used. This data, which is openly accessible via the Dutch ministry of the Interior and Kingdom Relations, provides an indication of the livability of a certain neighborhood. The data provides information on several factors contributing to the livability of a neighborhood, being housing stock, public space, amenities, the socioeconomic composition of population, social cohesion and security. Another proxy for socioeconomic background was obtained from the Netherlands Institute for Social Research. They publish data indicating the social status of a neighborhood every four years. This status is derived from the education level, income and labor market position of the neighborhood residents.

Information on tertiary education institutions, as well as on characteristics of a specific field of education is obtained from the Nationale Studenten Enquête (NSE) and Studiekeuze123. This data includes information on selection policies of institutions and specific fields of study. Furthermore, it provides information of the labor market prospects, at the time students had to decide on their field of tertiary education. These prospects were published on the official website on tertiary education choice of the Dutch government.[40] Furthermore, data from the Dutch Inspectorate of Education was obtained to be able to control for additional characteristics of secondary schools. Means are imputed if data was missing.[41] A more detailed description of the data sources is provided in Appendix B.

The dependent variable in this study is the retention outcome of students. This outcome is defined as 1 if a freshman returns for sophomore year. On the other hand, it is defined as 0 if a student enrolls in the first year of tertiary education in some field of education, but does not return for sophomore year *in the same field of education*. If a student does not return for

[38] More specifically, it includes information on their enrollment in both the first and second year they enrolled in higher tertiary education. If this student enrolled in the same field of education, this student's retention outcome is defined as 1. If this student does enroll in the first year, and in the subsequent either enrolls in another field of higher tertiary education (i.e. the student switches studies) or does not enroll in higher tertiary education at all (i.e. the student drops out), the retention outcome is set to 0.

[39] For the other 45%, this data is not available. In these schools, deans deleted old cohorts from their overview, which had as a consequence that the intensity information for these cohorts was lost.

[40] www.studiekeuze123.nl

[41] If a categorical variable is missing, an extra category is constructed representing the missing data. A dummy variable denoted 1 if that specific observation contained missing data was included in the analysis in order to examine whether this group is a selective group.

sophomore year in the same field of education, this implies that this student has dropped out of higher education or that the student has switched to a different field of education.

This study examines the effect of two different treatments. Which treatment is evaluated in which approach is discussed in section 4. The first treatment consists of being at a school that uses the Qompas package to provide career orientation guidance. This information is obtained from a list of schools that purchased licenses, per graduation cohort, for the career orientation guidance package. The second treatment is the extensive use of this package and uses information on the average intensity of use. This information is available on the school-graduation cohort level. The average intensity of use is measured as the average number of activities per student. As the guidance package consists of several distinct activities (e.g. do a choice-of-study test, save information on a specific field of study or open day), the average number of activities is the most comprehensive measurement of intensity of use.

In addition, several groups of control variables are added to the equation. Firstly, the group of individual characteristics includes age, gender, secondary education type, SE average grade, CE average grade, double degrees, profile, ethnicity, neighborhood status scores and socioeconomic status. Secondly, the group of secondary school characteristics consists of the number of graduation candidates, the average SE grade, the average CE grade, a dummy indicating gymnasium candidates, a dummy indicating if the school is in one of the four, and one of the 37, biggest cities in the Netherlands, the graduation rate, ideology, vision, region, urbanity, lower secondary education efficiency, postponed tracking, and the rate of apcg students. Thirdly, a group of tertiary institution characteristics is added, consisting of institution fixed effects, a HBO/WO dummy, the rate of female entrants, the number of entrants, an academic dismissal policy dummy, a fixed quota dummy, a selective admission standards dummy and the student/personnel ratio. An overview and a more detailed description of these covariates is provided in appendix B.

4. Empirical strategy

In an empirical examination of the effect of this career orientation package, it should be considered that the sample of schools that introduces the Qompas package might be a non-random sample of schools.[42] The introduction of this package may depend on individual characteristics, as well as on school characteristics. An OLS framework addresses this by controlling for several individual χ_{ist} characteristics and school characteristics ψ_{ist}:

$$\boldsymbol{Y_{ist} = \beta_0 + \beta_1 I_{st} + \beta_2 \chi_{ist} + \beta_3 \psi_{st} + \epsilon_{ist}} \quad \textbf{(1)}$$

In this equation, Y_{ist} is the retention outcome of student i from school s at time t. I_{st} is the treatment indicator, which is defined as 1 if the career orientation guidance package is used at school s at time t, and is defined as 0 if otherwise. Therefore, β_1 would be the parameter of interest in this equation. Besides individual (χ_{ist}) and school (ψ_{st}) characteristics, this equation includes an error term ϵ_{ist}.

Now, ϵ_{ist} may consist of unobserved individual or school characteristics. If such characteristics are correlated with I_{st} and Y_{ist}, parameter β_1 is a biased estimator of the effect of the career orientation guidance on retention outcomes. Therefore, this study entails three alternative identification strategies, intending to circumvent this omitted variable bias. In this section, these strategies will be scrutinized.

4.1. Basic difference-in-differences approach

In order to estimate if the Qompas career orientation guidance package affects retention outcomes of students, a difference-in-differences approach is used. In this approach, students from schools that use this package, the treatment group, are compared to students from schools that do not use this package, the control group:

$$\boldsymbol{Y_{ist} = \beta_0 + \beta_1 I_{st} + \beta_2 \chi_{ist} + \beta_3 \psi_{st} + A_s + B_t + \epsilon_{ist}} \quad \textbf{(2)}$$

In this equation, which is similar to equation 1, school fixed effects A_s and graduation cohort fixed effects B_t are included. These school fixed effects capture school-specific characteristics that are constant over time. In addition, the graduation cohort fixed effects capture year-specific characteristics that are the same for all schools. Standard errors are clustered at the school level. The treatment in this approach consists of being at a school that uses the Qompas package to provide career orientation guidance.

[42] E.g. schools may have deans that are passionate about what they do and want to guide their students the best way possible. Otherwise, it might be that these schools observe that their students lack such guidance by their parents and therefore feel responsible for such guidance.

4.2. Synthetic control method

In the two identification strategies below, the control group consists of a pool of schools that are selected according to certain conditions.[43] This study uses the synthetic control method for comparative case studies, which is developed by Abadie and Gareazabal (2003), to construct a synthetic counterfactual from this pool that fits the treatment group properly. This is a data-driven method to construct a proper counterfactual, thereby reducing ambiguity about how control schools are chosen. By using a weighted combination of control schools, this method can provide a control group with a pre-intervention dependent variable trend that matches the pre-intervention trend of the treated group optimally. In addition, it is possible to match on pre-intervention characteristics of the schools as well. The intuition behind using a weighted average of control schools as a control group is that a group of schools should fit the treatment group better than one particular school, both on the dependent variable trend and on other characteristics.

In this section, a formal description of this synthetic control method will be provided. Suppose all secondary schools in the Netherlands $i = 1,\ldots,J+1$ are observed over time $t = 1,\ldots,T$. Of this group, only the first school ($J=1$) receives the treatment in period $T_1 \geq T_0+1$ (so that there are T_0 pre-intervention periods, with $1 \leq T_0 \leq T$), implying that J schools may serve as a control school and collectively represent the so-called donor pool. Furthermore, suppose there are no anticipation effects prior to the first treatment period T_0+1, and interference between schools is absent.

The synthetic control group consists of a weighted average of schools in the donor pool. The synthetic control can be described as a Jx1 vector of weights $W = (w_2,\ldots,w_{J+1})$ so that each school in the donor pool is assigned a weight, with these weights summing up to 1. These weights can be assigned in such a way that the characteristics of the treated group X_1 (a (k*1) vector containing these values) are approximated by the synthetic control characteristics X_0(a (k*1) vector containing these values). The synthetic control, W^* can be selected such that it minimizes the vector $X_1 - X_0W$, containing the differences between the pre-intervention characteristics of the treated and the synthetic control school. However, characteristics may differ in their relative importance as a predictor of the dependent variable. For each characteristic $m = 1,\ldots,k$, X_{1m} and X_{0m} (which is a 1*J vector) reflect the value of the m-th variable for the treated school and of a school in the donor pool. If v_m reflects the

[43] In both strategies, small schools – with on average less than 40 students per cohort – are dropped from the sample, as their retention rates are highly volatile. As a robustness check, the same analysis is performed without this restriction.

relative importance of the m-th variable, W* is chosen such that it minimizes $\sum_{m=1}^{k} v_m(X_{1m} - X_{0m}W)^2$. In this manner, variables with a relatively large predictive power on the outcome variable are assigned large v_m weights as for those variables, it is relatively important that the synthetic control approximates the treated school.

In the context of career orientation guidance, the treatment group does not consist of one single school but of a group of schools. Therefore, a treatment group is constructed as the average of all treated schools and the synthetic control that matches this treatment group as good as possible is constructed as discussed above. Subsequently, the weights W* are assigned to the schools in the donor pool. The pre-intervention characteristics are supposed to be predictors of post-intervention outcomes. In this context, this would imply that characteristics of pre-intervention cohorts would predict the retention outcomes of post-intervention cohorts. As the composition of the cohorts of students may be time-varying, this is not very likely. This may bias the synthetic control estimates, if the effect is the result of a change in observable characteristics. To overcome such biases, a difference-in-differences model is estimated, controlling for these characteristics. The weights that have been assigned to schools by the synthetic control method have been allotted to students i dependent on the respective size of their school. In this manner, the weights of the synthetic control group sum up to 1, as do the weights of the treatment group.

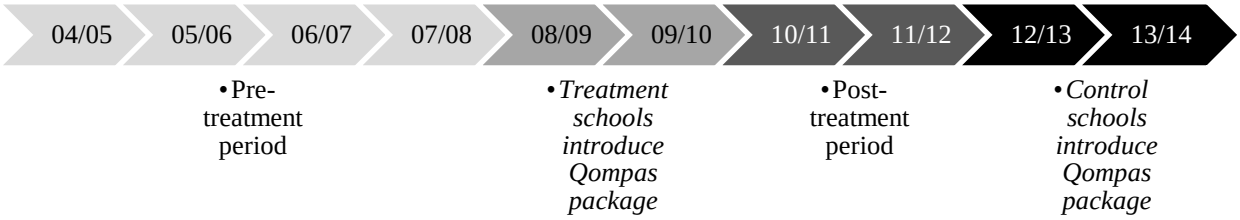
4.3. Variation in the year of introduction

This identification strategy exploits variation in the year of introduction of the Qompas career orientation guidance package. Again, the treatment consists of being at a school that uses the Qompas package to provide career orientation guidance. In a difference-in-differences approach, retention outcomes of cohorts from schools that introduced the package early (graduation cohorts 08/09 or 09/10) are compared with those of students from schools that introduced the package late (graduation cohorts 12/13 or 13/14). The pre-treatment period is defined as the graduation cohorts 04/05 until 07/08, whereas the post-treatment period is defined as the graduation cohorts 10/11 and 11/12. Now, in the post-treatment period, the treatment group has introduced this package, and still uses it, whereas the control group has not introduced this package *yet*.[44] A timeline of this approach is provided in figure 4.

[44] For a similar identification strategy, see Abramitzky & Lavy (2014).

FIGURE 4

TIMELINE OF THE SELECTION OF TREATMENT AND CONTROL GROUPS



Notes: This figure presents an overview of how this identification strategy is set up. It presents, over time, which graduation cohorts are defined as pre- and post-treatment periods. Furthermore, it visualizes when treatment and control schools introduced the Qompas career orientation guidance package.

From this pool of control schools, a synthetic control group is constructed using the synthetic control method. The weights that are the result of this procedure are allotted to students dependent on the respective size of their school, such that both the weights in the control group as the weights in the treatment group sum up to 1. Subsequently, the effect of introducing the Qompas career orientation guidance package is estimated using a weighted probit model:

$$\boldsymbol{Y_{ist} = \beta_0 + \beta_1 I_{st} + \beta_2 \chi_{ist} + \beta_3 \psi_{st} + A_s + B_t + \epsilon_{ist}} \quad \textbf{(3)}$$

4.4. Shocks in intensity of use

A third attempt to circumvent the omitted variable bias exploits shocks in the intensity of use of the Qompas package. The treatment in this setting consists of the extensive use of the Qompas career orientation guidance package at a secondary school. The data indicate that schools vary significantly in how extensively they use this guidance package. The treatment group consists of schools that have introduced Qompas in the pre-treatment period, and suddenly experience a major shock in intensity of use. Such a shock is defined to be the upper quintile of treatment intensity growth.[45] The control group in this strategy consists of schools that use the Qompas package and keep using it at the same level of intensity. Intensity is denoted as stable if intensity growth is within the 4th, 5th, 6th or 7th deciles of treatment intensity growth. In the pre-treatment period, which consists of the 6 cohorts prior to the shock, the intensity of treatment in both treatment and control schools, is stable, as long this is known. The post-treatment period consists of the first and second cohort after the shock. Treatment intensity is stable from the first to the second post-intervention cohort as well. To gauge the magnitude of the shock, treated schools use the package with an average intensity

[45] I.e. the upper 20% of the distribution of the absolute growth in intensity of treatment is denoted as a major shock. This is similar to a minimum absolute growth threshold of about a standard deviation in treatment intensity.

prior to the shock. After these shocks, this group of schools uses the package with a relatively high intensity, which means that from one year to the other, quite some additional time is devoted to provide career orientation guidance. In comparison, control schools are just below average both in the pre- and post-treatment period.

From the pool of control schools, a synthetic control group is constructed using the synthetic control method. The weights that are the result of this procedure are allotted to students dependent on the respective size of their school, such that both the weights in the control group as the weights in the treatment group sum up to 1. Subsequently, the average marginal treatment effect is estimated using a weighted probit model:

$$\boldsymbol{Y_{ist} = \beta_0 + \beta_1 I_{st} + \beta_2 \chi_{ist} + \beta_3 \psi_{st} + A_s + B_t + \epsilon_{ist}} \quad \textbf{(4)}$$

5. Validity of the designs

This section explores the validity of the three designs that have just been discussed.

5.1. Basic difference-in-differences approach

As displayed in table C2, appendix C, students from the control group and the treatment group are significantly different regarding their grades, and the schools in both groups differ regarding graduation rate, size, and urbanity. This indicates that these groups are distinct groups, suggesting a selective group of schools introduced this intervention. Several assumptions need to be made for β_1 to be the unbiased estimator of the average marginal treatment effect of this career orientation guidance package on retention outcomes. The key assumption of a difference-in-differences approach is known as the common trends assumption. This involves that the retention rate growth of the control group represents the counterfactual retention rate growth of the treatment group if there were no treatment. To test this assumption, a constant linear time trend model is estimated, while including an interaction of this time trend with the treatment indicator.[46] The results in panel A, table 1 show that for some treated groups, a linear time trend in the pre-treatment period exists, but that this time trend is the same for treatment and control schools. In addition, a model in which the linear time trend is replaced with year dummies is estimated. Interactions of each of these year dummies with the treatment indicator are included to estimate the difference in time trend in that year between the treatment and control group. The results in panel B of table 1 show that again, the pre-treatment time trend in retention rates is similar for the treatment and control group. These results suggest that the common trends assumption is plausible.

[46] Note that, as the year of introduction of the treatment varies among schools, various models are estimated. For each year of treatment introduction, the pre-treatment time trend of the schools that introduced the treatment in that year is compared with the pre-treatment time trend of control schools – schools that do not use the package in that year -.

TABLE 1

TREATMENT-CONTROL DIFFERENCES IN PRE-TREATMENT TIME TRENDS IN RETENTION RATES PER INTRODUCTION COHORT

	07/08	08/09	09/10	10/11	11/12	12/13	13/14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Linear Trend Model							
Time trend	-0.013***	-0.006***	-0.003**	-0.001	-0.001	0.001	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Treatment *	-0.005	0.003	-0.004	0.003	-0.001	-0.004	0.004
Time trend	(0.011)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)
Treatment	-0.055***	-0.009	0.007	-0.061***	0.065***	0.007	0.006
	(0.014)	(0.009)	(0.009)	(0.006)	(0.006)	(0.011)	(0.012)
B: Cohort Dummies Model							
Treatment	-0.049***	-0.014	0.004	-0.056***	0.065***	0.009	0.014
	(0.016)	(0.011)	(0.012)	(0.009)	(0.009)	(0.017)	(0.020)
Treatment *	-0.020	0.016	0.003	-0.011	0.005	-0.022	-0.014
2005	(0.022)	(0.016)	(0.016)	(0.012)	(0.013)	(0.023)	(0.028)
Treatment *	-0.010	0.011	0.003	0.003	0.001	0.009	0.000
2006	(0.022)	(0.016)	(0.016)	(0.012)	(0.013)	(0.023)	(0.028)
Treatment *		0.013	-0.019	-0.002	-0.007	-0.010	0.011
2007		(0.016)	(0.016)	(0.012)	(0.013)	(0.023)	(0.028)
Treatment *			-0.008	0.014	-0.013	-0.030	0.002
2008			(0.016)	(0.012)	(0.013)	(0.023)	(0.028)
Treatment *				0.005	-0.002	-0.016	0.026
2009				(0.012)	(0.013)	(0.023)	(0.028)
Treatment *					0.003	-0.028	0.006
2010					(0.013)	(0.023)	(0.028)
Treatment *						-0.028	0.020
2011						(0.023)	(0.028)
Treatment *							0.030
2012							(0.028)
F-statistic	0.434	0.412	0.646	0.941	0.479	0.746	0.511
P-value	0.648	0.745	0.630	0.453	0.825	0.633	0.849

Notes: This table presents OLS regression results at the school level, predicting the retention rates of graduation cohorts in the pre-intervention period. All columns present retention rate time trend differences between control schools and a certain graduation cohort. As the year of introduction varies widely, several treatment groups exist. Each column compares pre-treatment period retention rate trends of a certain treatment group (schools that introduce the package for e.g. graduation cohort 08/09, which is column 2) with its control group (schools that have not introduced the package for this graduation cohort). In panel A, outcomes may vary according to a linear time trend that differs in treatment and control schools. Panel B allows differences between treatment and control schools to vary freely for each graduation cohort. Panel B included cohort dummies as well. The F-statistic and P-value test whether the interaction terms in panel B between treatment schools and cohort dummy variables are jointly zero. Standard errors are reported in parentheses.

Furthermore, this strategy assumes that among control schools, the way in which they provide guidance to their students does not change from the pre-treatment to the post-treatment period.[47] Information on the type of guidance schools in the control group provide is not available. Therefore, a threat for this approach is that control schools change their guidance policies over time, which would bias the estimates.

Furthermore, the no interference assumption should hold, which in essence means that the treatment should not affect individuals outside the treatment group in any way. Such interference effects are often thought of as spillover effects. In this context, this would mean

[47] If this assumption would be violated, β_1 reflects not only the effect of the intervention, but also the effect of the change in guidance activities in the control group, which would imply that β_1 is a biased estimator of the true treatment effect.

that students, that receive the intervention, would affect the quality of educational choice on other schools. Another option would be that students that more consciously decided on their specialization of study, are more motivated as a freshman, which would spillover to other students. Both mechanisms seem unlikely and, if it would occur, be of small importance.

A last assumption is that anticipation effects are absent. This entails that students that receive the intervention should not be affected prior to the introduction of the intervention. This seems to be a feasible assumption, as the decision on the introduction of the intervention is generally made by school boards and students are hardly involved in such decisions.

Another threat to the accuracy of the estimates is that if treated schools, besides that they use the Qompas package, provide other kinds of guidance to their students, this would lead to an overestimation of the true treatment effect. However, schools have limited funds to finance the education they provide. As schools are not obliged to provide career orientation guidance, it is unlikely that schools, that purchase the Qompas package, would invest in other guidance methods.

A drawback of this approach is that it is likely that schools, before they introduce the Qompas package, provide other kinds of guidance to their students. This implies that the estimates in this model should be interpreted as the effect of introducing this specific Qompas package, relative to business as usual. Since some guidance is provided in general, this estimate does not (fully) capture the effect of career orientation guidance.

5.2. Variation in the year of introduction

As presented in table C4, appendix C, the treatment group and the synthetic control group in this design are very similar regarding both individual and school characteristics.[48] These results suggest that the approach to compare students from the (selective) pool of schools that ever introduced this intervention leads to reduced differences in observed characteristics between the treatment and control group. Another advantage of this strategy, relative to the previous one, is that the synthetic control group is constructed such, that its pre-intervention trend matches the treatment group trend perfectly. Therefore, the common trend assumption is even more plausible. The application of this synthetic control method assumes that interference and anticipation effects are absent. Regarding these assumptions, the same reasoning as in the previous design holds. As this approach includes a limited sample in the

[48] In the pre-intervention period, only SE grades are significantly different, although the magnitude of this difference is rather small (on a scale of 1 to 10, the difference is .0677). In the post-intervention period, both Se and CE grades are significantly different, but this difference is small as well. Note that the intervention may have affected the grades in this year.

analysis, the model loses some statistical power[49] and the external validity may be limited if the group of schools that introduce Qompas is indeed a selective group of schools. Again, an important drawback of this approach is that it measures the effect of introducing this specific package, relative to business as usual, since it is likely that before the introduction of the Qompas package, some guidance is provided. Furthermore, guidance activities in the control group are unknown, which biases the results if guidance policies in the control group change from the pre-treatment to the post-treatment period.

5.3. Shocks in intensity of use

In this design, the control group no longer consists of schools that do not use the Qompas package in the post-treatment period. All schools in the sample use the Qompas package in the pre-treatment period, as well as in the post-treatment period. Therefore, the assumption of no other guidance activities if the intervention is not provided is no longer problematic. The same reasoning as above holds regarding the no interference and the absence of anticipation effects assumptions. This is assumed by the application of the synthetic control method, which ensures that the common trends assumption is plausible as well. However, as a limited sample of the data is included in the analysis, some power is lost which increases the probability of a type 2 error. Also, it may cause some external validity problems. Furthermore, the origin of this sudden shock in intensity is relevant. Unfortunately, no information on the cause of these shocks is available. Both an exogenous and an endogenous origin is feasible. It may be that these two cohorts that extensively use this program (which results in an observed shock) are two cohorts of motivated students, willing to invest more time in their educational choice.[50] An exogenous explanation of such a shock would be that the school determines that more or better guidance should be provided, such that more teaching hours are devoted to providing career orientation guidance. If the origin of this shock is endogenous, β_1 would be a biased estimator of the true treatment effect. β_1 most likely will be an overestimation of the true treatment effect if the shock is the result of 2 highly motivated cohorts. This explanation only holds if both post-treatment cohorts are really different than the pre-treatment cohorts which is not suggested by the descriptive statistics in

[49] As the sample size has become smaller, the probability that the binary hypothesis test correctly rejects the null hypothesis of no effect, has become smaller. Put differently, the probability of a type two error has increased due to a smaller sample size.

[50] Note that, for this explanation to hold, both post-treatment cohorts should be a selective group, compared to the pre-treatment cohorts of that school. If only one cohort would use the program extensively, but the subsequent cohort would use it much less, this school would not be selected into the treatment group.

table 2.[51] Therefore, this seems to be an unlikely explanation. In case the shock has an exogenous origin (e.g. due to a new dean), the treatment group would not consist of a selective group of students, implying that this does not bias the results. On the other hand, this might raise concerns regarding the assumption that no other guidance related tasks, besides what Qompas offers, is practiced. However, since school budgets are limited, it seems unlikely that schools would use the Qompas intervention extensively and, besides that, arrange several other guidance-related activities. Therefore, the estimates of this design are considered the main estimates of this study, as it has the highest internal validity.

[51] Regarding the relevant background characteristics, the post-treatment cohorts of the treatment group are very similar to their counterparts in the pre-treatment cohorts. Nothing suggests that the post-treatment cohorts are much more motivated.

TABLE 2

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE

	Pre-treatment period			Post-treatment period		
	(1) Treatment	(2) Synthetic control	(3) (1)-(2)	(4) Treatment	(5) Synthetic control	(6) (4)-(5)
Individual characteristics						
Age	16,869	16,751	0,117	16,860	16,769	0,091
	(0,009)	(0,007)	(0,106)	(0,015)	(0,012)	(0,096)
Male	0,461	0,474	-0,013	0,457	0,467	-0,010
	(0,006)	(0,005)	(0,012)	(0,011)	(0,009)	(0,016)
VWO	0,510	0,397	0,113	0,495	0,388	0,107
	(0,006)	(0,005)	(0,178)	(0,011)	(0,009)	(0,178)
SE grades	66,714	65,855	0,859	65,675	65,459	0,216
	(0,070)	(0,054)	(0,649)	(0,121)	(0,102)	(0,665)
CE grades	64,839	63,921	0.919**	65,076	65,110	-0,034
	(0,086)	(0,070)	(0,460)	(0,138)	(0,118)	(0,627)
Native Dutch	0,865	0,871	-0,006	0,863	0,881	-0,018
	(0,004)	(0,003)	(0,025)	(0,007)	(0,005)	(0,025)
Non-western immigrant	0,065	0,071	-0,006	0,075	0,061	0,014
	(0,003)	(0,002)	(0,018)	(0,006)	(0,003)	(0,019)
Socioeconomic status	0,115	0,086	0,029	0,138	0,064	0,074
	(0,010)	(0,008)	(0,152)	(0,019)	(0,014)	(0,185)
Secondary school characteristics						
Number of graduation candidates	98,550	97,973	0,577	103,576	106,926	-3,350
	(0,463)	(0,282)	(11,195)	(0,807)	(0,484)	(12,451)
Graduation rate	0,914	0,897	0.017*	0,889	0,889	0,000
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,014)
Urbanity	2,694	2,708	-0,014	2,525	2,757	-0,232
	(0,014)	(0,012)	(0,371)	(0,028)	(0,021)	(0,451)
Tertiary school characteristics						
Research university students	0,434	0,328	0,106	0,422	0,327	0,095
	(0,006)	(0,005)	(0,142)	(0,011)	(0,008)	(0,146)
Studies with academic dismissal policy	0,762	0,767	-0,005	0,682	0,627	0,054
	(0,005)	(0,004)	(0,048)	(0,010)	(0,008)	(0,066)
Studies with fixed quota	0,145	0,135	0,009	0,217	0,209	0,008
	(0,004)	(0,004)	(0,023)	(0,009)	(0,007)	(0,033)
Studies with selective admission standards	0,059	0,048	0,011	0,147	0,139	0,008
	(0,003)	(0,002)	(0,011)	(0,008)	(0,006)	(0,030)
Tertiary school outcomes						
Retention	0.681	0.677	0.004	0.719	0.680	0.038
	(0.006)	(0.005)	(0.029)	(0.010)	(0.008)	(0.032)
Switch	0.234	0.239	-0.005	0.226	0.233	-0.007
	(0.005)	(0.004)	(0.017)	(0.009)	(0.007)	(0.016)
Dropout	0.084	0.083	0.001	0.055	0.086	-0.031*
	(0.003)	(0.003)	(0.016)	(0.005)	(0.005)	(0.019)
Number of schools	15	42		15	42	
Number of students	6381	20649		2173	7162	

Notes: Columns 1, 2, 4, and 5 present means and standard deviations (in parentheses) of several background characteristics of students and schools in the treatment (1 and 4) and control (2 and 5) group in the pre-treatment (columns 1, 2, and 3) and post-treatment (columns 4, 5, and 6) period. Treatment schools experienced a major shock in intensity of treatment. A major shock in intensity of treatment is defined to be the upper 20% of treatment intensity growth. The control group has a stable treatment intensity (thresholds of stable intensity defined to be the 4th, 5th and 6th deciles in treatment intensity growth). Columns 3 and 6 present the differences between these groups, as well as standard errors that are clustered at the school level in parentheses. These differences are estimates based on a regression of the characteristic as a dependent variable and the treatment indicator as the explanatory variable. The pre-treatment period consists of the 6 cohorts prior to the shock. The post-treatment period consists of the two cohorts after the shock is observed. Treatment intensity is stable from the first to the second post-treatment period for all schools. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

6. Main estimates

This section presents the empirical results of the estimation strategies that have just been discussed.

6.1. Basic difference-in-differences

This section discusses the results of a basic difference-in-differences design, measuring the effect of using the Qompas career orientation guidance package at a secondary school on retention rates, relative to business as usual. The estimates in panel A of table 3 suggest that being in a school that uses the Qompas guidance package does not significantly affect retention rates. An explanation for this may be that schools, before they introduce this specific package, provide other kinds of guidance to their students. Then, replacing old guidance activities for guidance using the Qompas package may not affect retention outcomes, although the guidance itself may have an impact. However, the estimates in panel B shows that non-western immigrants seem to benefit from the availability of this intervention at their secondary school (retention rates improve by roughly 1.5 percentage point). This estimate loses its significance once tertiary education characteristics are included in the model, although this might be the result of included endogenous variables, as discussed before. In panel C, the same model is run, allowing for a differential effect by intensity of use. This model does not find such effects, relative to business as usual. It is likely that before the introduction of this specific package, schools use another package with a similar intensity, which would explain that no such effects are found.

TABLE 3

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF USING THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	-0.006	0.002	0.002	0.001
	(0.004)	(0.002)	(0.002)	(0.002)
B. Heterogeneous effects				
i. Secondary education type				
Average treatment effect		-0.002	-0.002	-0.002
		(0.003)	(0.003)	(0.003)
Average treatment effect * VWO		0.005	0.004	0.004
		(0.005)	(0.005)	(0.005)
ii. Gender				
Average treatment effect		0.003	0.004	0.003
		(0.003)	(0.003)	(0.003)
Average treatment effect * male		-0.003	-0.004	-0.004
		(0.004)	(0.004)	(0.004)
iii. Socioeconomic status				
Average treatment effect		0.001	0.001	0.001
		(0.003)	(0.003)	(0.003)
Average treatment effect * low		0.005	0.003	0.003
socioeconomic status		(0.006)	(0.006)	(0.006)
iv. Ethnicity				
Average treatment effect		-0.000	0.000	0.000
		(0.002)	(0.002)	(0.002)
Average treatment effect * non-western		0.019***	0.013*	0.012
immigrant		(0.007)	(0.008)	(0.007)
C. Intensity of exposure to treatment				
Low	0.009	0.001	0.002	0.002
	(0.008)	(0.005)	(0.005)	(0.005)
Mid	-0.014**	-0.004	-0.004	-0.004
	(0.007)	(0.004)	(0.004)	(0.004)
High	-0.014	0.004	0.006	0.006
	(0.009)	(0.006)	(0.006)	(0.006)
Unknown	-0.003	0.004	0.003	0.003
	(0.005)	(0.003)	(0.003)	(0.003)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	656362	656362	536817	535819

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the introduction of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Column 1 includes no controls and no fixed effects. Columns 2, 3, and 4 include year and school fixed effects. In addition, column 2 adds individual background characteristics to the model, column 3 adds secondary school characteristics to the model, and column 4 adds tertiary school characteristics to the model. Appendix B discusses these groups of covariates in more detail. Panel A presents the estimated effects comparing students from treatment and control schools. Panel B presents a similar model, but adds interaction terms of some characteristic with all explanatory variables. This is reported for VWO students, males, low socioeconomic status (threshold defined to be the lower 20% of all students), and non-western immigrants. Panel C presents the estimated average marginal treatment effects, allowing for varying treatment effects by intensity of exposure to the treatment. High intensity of treatment is defined to be the upper 25% of treatment intensity. Middle intensity is defined to be the 2nd and 3rd quartile and low intensity is defined as the lower 25% of treatment intensity. Schools with unavailable treatment intensity indicators together constitute the unknown intensity group. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

6.2. Variation in the year of introduction

Panel A of table 4 presents estimates comparing early introducers (cohorts 08/09 or 09/10) to a synthetic control group of late introducers (cohorts 12/13 or 13/14). Without adding control variables, the point estimate of 0.8 percentage point – which is not significantly different from zero – is in line with figure 5, presenting average retention rates of the treatment and synthetic control group. Once control variables are added to the equation, the point estimate becomes smaller. These estimates suggest that this treatment does not lead to improved retention outcomes. which presents average retention rates of the treated group and the synthetic control group. Table A1 of appendix A includes interaction effects of males, Vwo students, low socioeconomic status students and non-western immigrants. These estimates suggest that no additional effects for these subgroups exist. In addition, it includes estimates of a model allowing for differential effects by intensity of use of the treatment, suggesting no such differential effects exist. Using different introduction years (treatment: 07/08 or 08/09, control 11/12 or 12/13) leads to similar results, that are displayed in panel B of table 4. Again, the effect of introducing this package, relative to business as usual, is not significantly different from zero. It is likely that schools, before they introduce the Qompas package, provide other kinds of guidance, with a similar intensity.

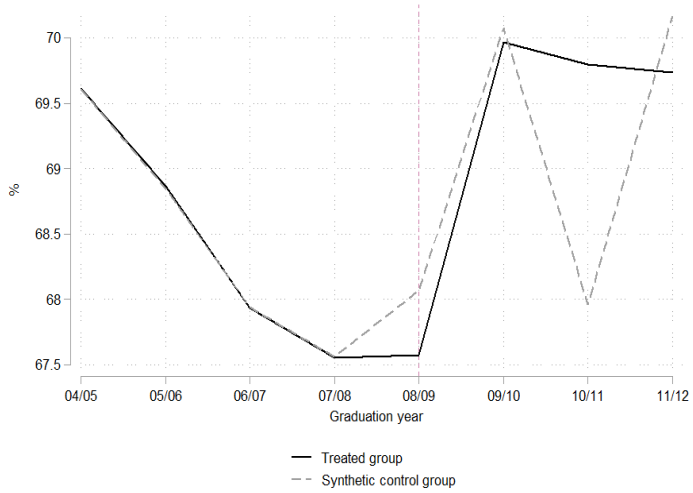
TABLE 4

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF USING THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.008	0.000	-0.001	0.002
	(0.010)	(0.011)	(0.012)	(0.012)
N	63003	62988	43704	43691
B. Difference-in-Differences probit: one year earlier				
Average treatment effect	0.016	0.008	0.010	-0.002
	(0.014)	(0.014)	(0.016)	(0.017)
N	70044	70044	44053	44035
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the introduction of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A presents the estimated effects comparing students from treatment and control schools. Treatment schools introduced the Qompas career orientation guidance package for graduation cohorts 08/09 or 09/10, whereas control schools introduced this package for graduation cohorts 12/13 or 13/14. Panel B presents the estimates of the same model, but both the treatment group and control group introduced this package one year earlier. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

FIGURE 5
RETENTION RATES OF TREATMENT AND SYNTHETIC CONTROL GROUP



Notes: This figure presents average retention rates in the treatment and synthetic group over time. Treatment schools introduced the Qompas career orientation guidance package for graduation cohorts 08/09 or 09/10, whereas control schools introduced this package for graduation cohorts 12/13 or 13/14.

6.3. Shocks in intensity of use

Now, the average marginal treatment effect of extensively using the Qompas package on retention rates is estimated, exploiting within-school shocks in intensity. This approach compares students from schools that experienced a sudden shock in intensity of use of the treatment with students from schools that use the intervention with a relatively stable intensity. Probit estimates of this approach are displayed in table 5. Focusing on the third specification, this treatment has a significant and large impact of roughly 4 percentage point on retention, on a base of 68 percent. The descriptive statistics in table 3 suggest that this effect is mainly driven by a drop in dropout rates. Table A3 in appendix A presents heterogeneous effects for VWO students, males, socioeconomic status and ethnicity. Although none of these interaction effects is significant, the point estimates suggest that the effect is especially present among Havo students. Figure 6 presents average retention rates of the treatment and synthetic control group over time, which confirm the sign of the estimates of the probit model. The magnitude of the effect that is indicated by figure 6 is smaller than the probit estimates, which is the result of adding controls to the model. This suggests that, based on their background characteristics, students in the treatment group are expected to have lower retention rates than students in the control group. Put differently, characteristics that are positively (negatively) correlated with the treatment estimator have, on average, a negative (positive) correlation with retention rates.

TABLE 5

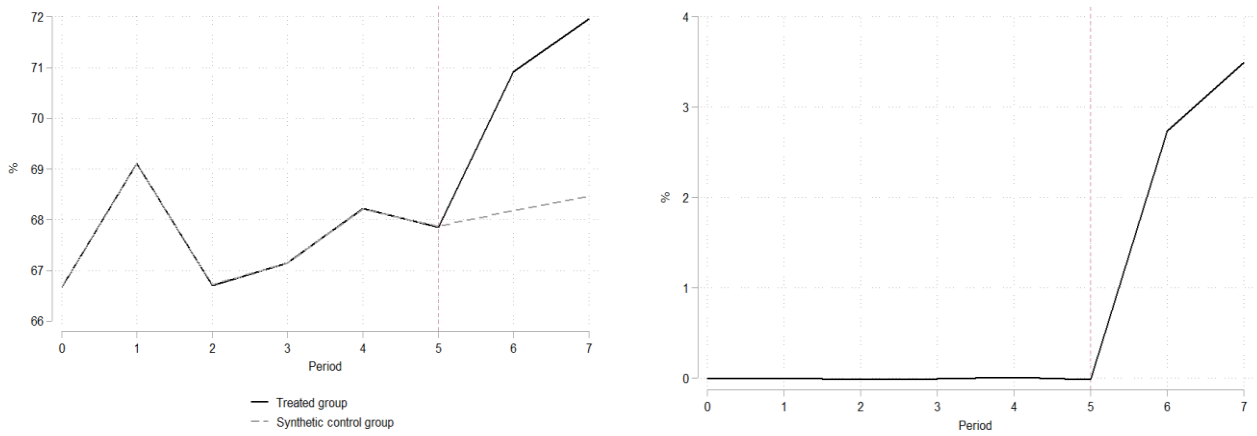
DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF A SHOCK IN INTENSITY OF USE OF THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.035*	0.033**	0.042***	0.042***
	(0.019)	(0.016)	(0.016)	(0.015)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	36365	36365	36365	36354

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the introduction of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A presents the estimated effects comparing students from treatment schools with their peers from control schools. Students from treatment schools experienced a major shock in intensity of treatment. A major shock in intensity of treatment is defined to be the upper 20% of treatment intensity growth. The control group has a stable treatment intensity (thresholds of stable intensity defined to be the 4th, 5th and 6th deciles in treatment intensity growth). The 6 cohorts prior to the observed shock, as well as the 2 cohorts after the shock are included in the model. Treatment intensity growth from the first to the second post-treatment period is close to zero for both treatment and control schools. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

FIGURE 6

AVERAGE RETENTION RATES OF TREATMENT AND SYNTHETIC CONTROL GROUP



Notes: The left plot presents average retention rates in the treatment and synthetic control group over time. The gaps plot on the right shows the difference between these two averages. The treatment and synthetic control group are determined as in table 5.

In addition, the effect of the same treatment is estimated on enrollment rates. Enrollment rates in the Netherlands are typically high[52], implying there is limited opportunity to analyze this. Although non-western Havo/Vwo graduates typically enroll in tertiary education as much as native Dutch students do, the estimates in table A4, appendix A, suggest that this treatment improves enrollment rates among non-western immigrants.[53] Moreover, it seems that Havo students (1.5 percentage point, significant at the 5% level) and females (1.5 percentage point, significant at the 10% level) enroll in tertiary education more often. Intuitively, a group of

[52] Approximately 95% of Havo/Vwo students enroll in tertiary education

[53] The interaction coefficient – which is marginally significant if individual characteristics are included in the model - loses its significance once secondary school characteristics are included in the analysis. Nonetheless, the point estimate remains large (4.5 percentage point).

students that enrolls as a result of this treatment, is likely to be a weaker group of students that has relatively low retention rates. Despite this, the findings in table A3 of appendix A do not suggests that the treatment has a negative effect on retention outcomes among these subgroups.

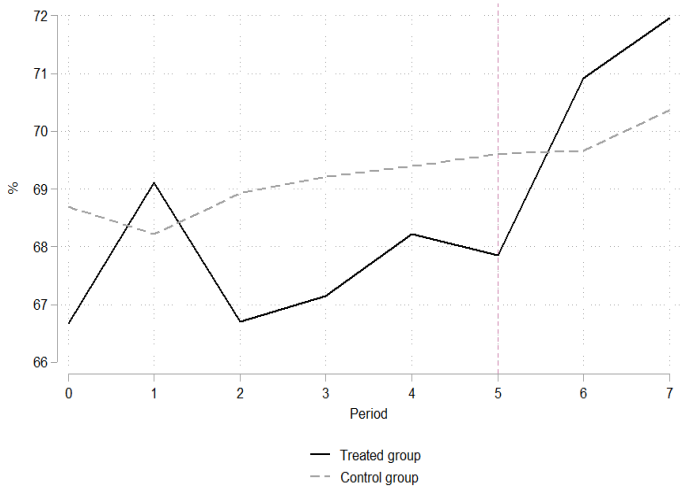
Furthermore, the same model is run with an outcome variable indicating whether the student enrolled in tertiary education immediately after secondary school graduation.[54] 13% of Havo/Vwo graduates that enroll in tertiary education do this after they have taken a break. An often heard reason for this is that they have not decided on their field of tertiary education yet. An effect on this variable may indicate that students start thinking about their educational choice earlier, implying they enroll immediately after secondary school graduation. Results are displayed in table A5, appendix A. The heterogeneous effects by gender indicate that males more often enroll immediately after graduation (roughly 5 percentage points more), although on average, males enroll after a break as much as females.

[54] This variable is defined as 1 if the student, in the year after secondary school graduation, enrolls in tertiary education. It is defined as 0 if this student does enroll in tertiary education 2 or more years after graduation. Students that do not enroll are not included in this analysis.

7. Robustness checks

In this section, the robustness of the findings from the design exploiting shocks in intensity is tested. The average marginal treatment effects following from each of these robustness checks are presented in table 6. As a first check, the same model is estimated, without constructing a synthetic control group that has a similar pre-treatment period. Put differently, all individuals in the control group are designated the same weight, which implies that the similarity of the pre-treatment period retention rate trends deteriorates, which is presented in figure 7. This deteriorates the plausibility of the common trends assumption. Panel A in table 6 shows that the average marginal treatment effect in this setting is approximately 3 percentage points, which is in line with figure 7.

FIGURE 7
AVERAGE RETENTION RATES IN TREATMENT AND CONTROL GROUP



Notes: This figure presents average retention rates in the treatment and control group over time. In contrast with figure 6, this figure presents the average retention outcomes of the control group without using the synthetic control method to construct a synthetic control group.

Second, the results are reproduced using different thresholds for excluding small schools from the sample, which hardly affects the results (see panel B of table 6). Third, the results are reproduced using different thresholds for assignment to the treatment and control group. The magnitude of the effect declines when the threshold of the magnitude of the intensity shock declines. This suggests that the larger the growth in intensity, the larger is the gain in retention outcomes. If conditions for a stable intensity are relaxed, the magnitude of the effect slightly decreases. Fourth, the same model is estimated, but without excluding students that did not enroll in tertiary education at all. Students that did not enroll are now

assigned a retention outcome of 0. Including these students in the analysis does not affect the results.

TABLE 6

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF SEVERAL ROBUSTNESS CHECKS ON THE ESTIMATES OF THE EFFECT OF A SHOCK IN INTENSITY OF USE OF THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES

	(1)	(2)	(3)	(4)
A: Estimates without using the synthetic control method				
Average treatment effect	0.025	0.022	0.031**	0.031**
	(0.018)	(0.016)	(0.016)	(0.015)
B: Estimates with different thresholds for excluding small schools				
Schools with on average less than	0.034*	0.033**	0.041***	0.041***
30 students are excluded	(0.019)	(0.016)	(0.016)	(0.016)
Schools with on average less than	0.035	0.030	0.031*	0.030**
50 students are excluded	(0.024)	(0.019)	(0.016)	(0.015)
No schools are excluded	0.033*	0.032**	0.038**	0.038**
	(0.018)	(0.016)	(0.016)	(0.015)
C: Estimates using different thresholds for treatment and control indicators				
Upper 10% is assigned to treatment	0.089***	0.068***	0.081***	0.078***
group	(0.023)	(0.019)	(0.020)	(0.021)
Upper 15% is assigned to treatment	0.053***	0.049***	0.044**	0.040**
group	(0.019)	(0.018)	(0.017)	(0.018)
Upper 25% is assigned to treatment	0.028	0.034**	0.034**	0.033**
group	(0.020)	(0.016)	(0.016)	(0.015)
Upper 30% is assigned to treatment	0.026	0.037*	0.027	0.025
group	(0.023)	(0.020)	(0.017)	(0.017)
Third – eighth deciles are assigned	0.016	0.019	0.031**	0.030**
to control group	(0.022)	(0.016)	(0.014)	(0.014)
Second and third quartile are	0.031	0.026	0.034**	0.036***
assigned to control group	(0.020)	(0.017)	(0.014)	(0.014)
D: Estimates without excluding students that did not enrol in tertiary education from the sample				
Average treatment effect	0.035*	0.033**	0.042***	
	(0.019)	(0.016)	(0.016)	
E: Estimates excluding the school with the highest Post- period RMSPE/pre-period RMSPE ratio				
Average treatment effect	0.035*	0.033**	0.041***	0.041***
	(0.019)	(0.016)	(0.016)	(0.015)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes

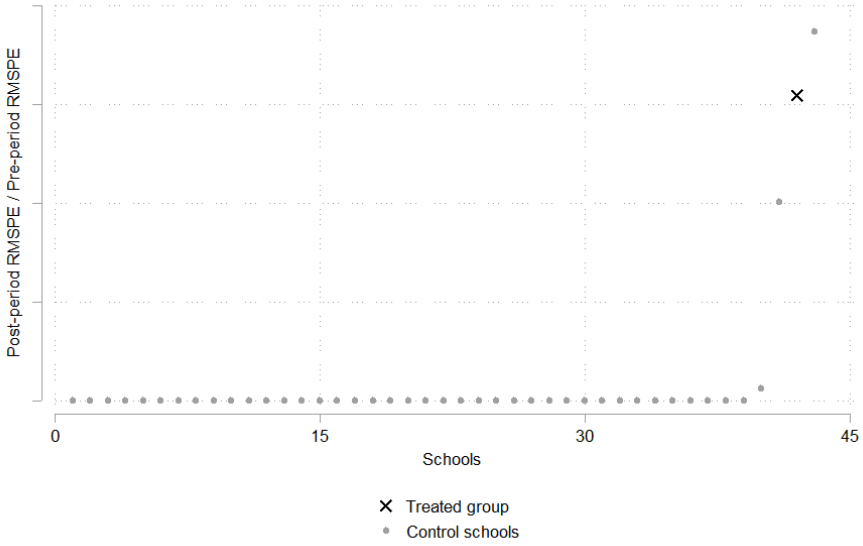
Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the extensive use of the Qompas career orientation guidance package on retention rates in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Each panel presents a certain robustness check, in which the model is slightly different than the baseline model in table 5. Panel A presents the estimates if, instead of using the synthetic control method, each individual in the control group is allotted the same weight. Panel B presents the estimates if the threshold of small schools that are excluded from the sample are set differently. Panel C presents the estimates if the thresholds for assignment to either the treatment of control group are set differently. Panel D presents the estimates if students that did not enroll in tertiary education are included in the sample (retention outcome of these students is set to zero). Finally, panel E presents the estimates if the school with the highest Post-period RMSPE/pre-period RMSPE ratio is excluded from the sample (see figure 8). Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

Fifth, the treatment indicator has been assigned to each school in the pool of control schools. In this manner, synthetic control estimates are obtained for schools that did not experience the treatment. Now, the estimated effect of the treated group is compared to the placebo effects of the control schools. The RMSPE measures the magnitude of the outcome

variable gap between each school and its synthetic control. A large post-treatment RMSPE indicates a large effect of the treatment only if the pre-treatment RMSPE is small.[55] In line with Abadie, Diamond & Hainmueller (2015), figure 8 displays the ratios of the post-treatment period RMSPE to pre-treatment period RMSPE for the treated group and for all other schools in the control group. As shown in figure 8, the treated group has the second highest ratio of this sample. If a school would be randomly selected from this sample, the chance of obtaining ratio as high as this one is $2/43 \approx 0.047$. The school that has a higher ratio than the treated group experiences a drop in retention rate in the post-treatment period. As presented in panel E of table 6, excluding this school from the pool of control schools does not affect the treatment estimator.

FIGURE 8

POST-TREATMENT PERIOD RMSPE / PRE-TREATMENT PERIOD RMSPE RATIOS OF TREATED GROUP AND CONTROL SCHOOLS



Notes: This figure presents the ratios between the post-treatment period RMSPE and the pre-treatment period RMSPE for all schools in the pool of control schools of the design as in table 5. In addition, the same ratio is displayed for the treatment group.

55 If the pre-treatment period RMSPE is large as well, this indicates that the common trends assumption is less plausible, hence a large post-treatment RMSPE is less indicative for a large treatment effect.

8. Discussion and concluding remarks

This paper evaluates a promising policy option, career orientation guidance, to address switch and dropout in tertiary education, which is a long enduring problem in Dutch education. This guidance is broadly provided at Dutch secondary schools, but how extensively this guidance is provided varies widely. First, this paper shows that students' retention outcome correlates with the secondary school they come from, which may suggest that secondary schools can affect the tertiary education retention outcome of their students. It is intuitive to state that secondary schools, by providing solid guidance, can encourage their students to consciously decide on their field of tertiary education.

Firstly, this study, in a difference-in-differences framework, estimates the effect of introducing a specific career orientation guidance package at a secondary school. Both in a design exploiting all available data and exploiting variation in the year of introduction, this estimate is not significantly different from zero. This should be interpreted as the effect of the Qompas package, relative to what schools already did. It is feasible that, before the introduction of this package, schools devoted a similar amount of time to provide guidance to their students.

Secondly, this study estimates the effect of extensively providing career orientation guidance by exploiting within-school shocks in intensity of use in a difference-in-differences framework. In this design, a large effect of 4 percentage point on retention outcomes is estimated. This effect size is based on intensity shocks of *at least* one standard deviation, which is equivalent to the upper 20 percent of the intensity growth distribution. This effect corresponds to a 12.5 percent decrease in the probability of students dropping out or switching in their first year of tertiary education. This estimate is consistent with Borghans, Golsteyn, & Stenberg (2015), who find that a standard deviation more counseling at secondary schools is associated with a 9 percent decrease in the probability of students preferring a different field of education in tertiary education.

Effects of this treatment on enrollment are less ambiguous, although the treatment seems to improve enrollment rates among Havo students and among non-western immigrants. Compared to e.g. Hurwitz and Howell (2013), enrollment effects in this study are small. A reasonable explanation for this is that Dutch enrollment rates of Havo, and especially Vwo graduates are relatively high already. The enrollment effects are more in line with Hoest, Jensen & Nielsen (2012), who find enrollment effects among immigrants, whereas they do not find such effects among native students.

A weakness of this study is that the main estimates are based on a small sample of treated schools as a shock in intensity is observed in only 15 schools, resulting in 2173 students receiving the treatment. More research on a larger sample of treated schools is needed to improve the accuracy of the magnitude of the effect, as well as to provide insight into heterogeneous effects for specific groups of students. Also, long term effects of this intervention (e.g. effects on graduation rates and future earnings) would be an interesting complement to this study. This would be useful to evaluate the cost-effectiveness of this intervention. Furthermore, the findings from this study suggest enrollment effects among some subgroups, on which additional research is needed. Finally, a qualitative study among the group of schools that positively affect their students' retention outcomes (see figure 2) may provide an idea of best practices that improve retention rates.

Another limitation of this study is that the origin of the shock in intensity is unknown. It may be that the shock is the result of two highly motivated cohorts that extensively use the package. Such an endogenous origin of the shock would imply that the estimated treatment effect is biased. However, the descriptive statistics (e.g. grades, graduation rates) in table 2 do not suggest that the post-treatment cohorts consist of better students. Therefore, an exogenous origin of this shock should be considered more plausible.

In conclusion, this study finds that providing extensive career orientation guidance to students in secondary education is associated with large retention rate improvements in tertiary education. Therefore, career orientation guidance is a promising policy option. Policy makers should consider putting in place incentives for secondary schools to improve the guidance they provide. An option may be to make information on enrollment, retention and graduation rates of former students publicly available. Furthermore, the Inspectorate of Education may judge on measures indicating how well students are equipped for tertiary education. Finally, career orientation guidance may be included as a mandatory part in the curriculum. This should be accompanied with certain quality requirements.

References

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *The American Economic Review*, *93*(1), 113-132.

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Abadie, A., Diamond, A., & Hainmueller, J. (2012). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495-510.

Abramitzky, R., & Lavy, V. (2014). How responsive is investment in schooling to changes in redistributive policies and in returns?. *Econometrica*, *82*(4), 1241-1272.

Adelman, C. (1999). Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment. Washington, DC: National Center for Education Statistics.

Adelman, C. (2006). The Toolbox Revisited: Paths to Degree Completion From High School Through College. *US Department of Education*.

Akerlof GA, Kranton RE(2000), Economics and Identity. Quarterly Journal of Economics 115(3), 715–753.

Arnold, I. J. (2015). The effectiveness of academic dismissal policies in Dutch university education: an empirical investigation. *Studies in Higher Education*,*40*(6), 1068-1084.

Arnold, I., & Van den Brink, W. (2010). Naar een effectiever bindend studieadvies [Towards a more effective binding study advice]. Tijdschrift voor Hoger Onderwijs en Management, 5(10), 10-13.

Astin, A. W. (1996). *Degree Attainment Rates at American Colleges and Universities: Effects of Race, Gender, and Institutional Type*. Higher Education Research Institute, Graduate School of Education & Information Studies, 3005 Moore Hall/Mailbox 951521, University of California, Los Angeles, CA 90095-1521; 310-825-1925.

Austen-Smith, D., & Fryer, R. (2005). An Economic Analysis of "Acting White" *The*

Quarterly Journal of Economics, 120(2), 551-583.

Avery, C., & Kane, T. J. (2004). Student perceptions of college opportunities. The Boston COACH program. In *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 355-394). University of Chicago Press.

Avery, C. (2013). "The Amherst Telementoring Program for High-Achieving, Low-Income Students; Results of a Pilot Study with a Randomized Controlled Trial," Harvard Kennedy School Working Paper.

Avery, C., Howell, J. S., & Page, L. (2014). A Review of the Role of College Counseling, Coaching, and Mentoring on Students' Postsecondary Outcomes. Research Brief. *College Board*.

Bailey, T., Calcagno, J. C., Jenkins, D., Leinbach, T., & Kienzl, G. (2006). Is student-right-to-know all you should know? An analysis of community college graduation rates. *Research in Higher Education*, *47*(5), 491-519.

Beffy M, Fougère D, Maurel A. (2012). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? The Review of Economics and Statistics 94(1), 334–347.

Belasco, A.S. (2013). "Creating college opportunity: School counselors and their influence on postsecondary enrollment," *Research in Higher Education*, *54*(7): 781-804.

Belzil, C., & Leonardi, M. (2007). Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics*, *14*(6), 957-970.

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment. *The Quarterly Journal Of Economics*, *127*(3), 1205.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489-520.

Bettinger, E. P., & Baker, R. B. (2013). The Effects of Student Coaching An Evaluation of a Randomized Experiment in Student Advising. *Educational Evaluation and Policy Analysis*, 0162373713500523.

Betts, J. R.. (1996). What Do Students Know about Wages? Evidence from a Survey of Undergraduates. *The Journal of Human Resources*, *31*(1), 27–56.

Booij, A. S., E. Leuven, and H. Oosterbeek. 2012. "The Role of Information in the Take-Up of Student Loans." Economics of Education Review 31 (1): 33 – 44.

Borghans, L., Coenen, J., Golsteyn, B., Huijgen, T., & Sieben, I. (2008). Voorlichting en begeleiding bij de studie-en beroepskeuze en de rol van arbeidsmarktinformatie. *Researchcentrum voor onderwijs en arbeidsmarkt, Maastricht*.

Borghans L, Golsteyn BHH, Stenberg A (2015) Does Expert Advice Improve Educational Choice? LoS ONE 10(12): e0145378. doi:10.1371/journal.pone.0145378

Bos, J. M., Berman, J., Kane, T. J., & Tseng, F. M. (2012). The Impacts of SOURCE: A Program to Support College Enrollment through Near-Peer, Low-Cost Student Advising. In *Association of Public Policy Analysis and Management Annual Conference*.

Bound, J., Lovenheim, M., & Turner, S. (2010). Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources. *American Economic Journal: Applied Economics, 2*(3), 129-157.

The Behavioral Insights Team (2016). Moments of choice. *Published on August 25, 2016*

Bridgeland, J., & Bruce, M. (2011a). "School Counselors Literature and Landscape Review," New York: The College Board.

Bridgeland, J., & Bruce, M. (2011b). 2011 National Survey of School Counselors: Counseling at a Crossroads. *College Board Advocacy & Policy Center*.

Brown, S., Ortiz-Núñez, A., & Taylor, K. (2011). Educational Loans and Attitudes towards Risk. Sheffield Economic Research Paper Series, 2011010

Bryan, J., C. Moore-Thomas, N. Day-Vines, & C. Holcomb-McCoy (2011). "School counselors as social capital: The effects of high school college counseling on college application rates," *Journal of Counseling & Development, 89*: 190–199.

Cabrera, A., Burkum, K., & La Nasa, S. (2005). Pathways to a four-year degree: Determinants of transfer and degree completion. In A. Seidman (Ed.), College student retention: A formula for student success. Westport: ACE/Praeger Series on Higher Education.

Calcagno, J. C., Bailey, T., Jenkins, D., Kienzl, G., & Leinbach, T. (2008). Community college student success: What institutional characteristics make a difference?. *Economics of Education review, 27*(6), 632-645.

Carrell, S. E., & Sacerdote, B. (2013). *Late interventions matter too: The case of college coaching New Hampshire* (No. w19031). Cambridge, MA: National Bureau of Economic Research.

Castleman, B., & J. Goodman (2014). "Intensive College Counseling and the College Enrollment Choices of Low Income Students," Harvard Kennedy School Working Paper, RWP14-031.

Castleman, B. L., L.C. Page, & K. Schooley (2014). "The forgotten summer: Mitigating summer attrition among college-intending, low-income high school graduates," *Journal of Policy Analysis and Management, 33*(2): 320– 344.

Castleman, B. L., & Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates?. *Journal of Economic Behavior & Organization,115*, 144-160.

Chambers, S. R. (2009). *Student engagement: using the NSSE benchmarks to investigate longterm persistence* (Doctoral dissertation, Montana State University-Bozeman, College of Education, Health & Human Development).

Cunha J, Miller T. (2009). Information and the Decision to Attend College: Evidence from the Texas GO Center Project. Unpublished manuscript.

Dinkelman, T., & Martínez, C. A. (2014). Investing in Schooling In Chile: The Role of Information about Financial Aid for Higher Education. *The Review Of Economics And Statistics*, *96*(2), 244.

Dohmen, T., Non, & Arjan. (2014). Hoe maken studenten en ouders keuzes in het onderwijs? In R. Klarus, L. Borghans, & I. Waterreus, *Wat is goed onderwijs?* (pp. 187-204). Den Haag: CPB.

Dutch Inspectorate of Education (2016). De Staat van het Onderwijs: Onderwijsverslag 2014/2015. *Utrecht: April 13, 2016*

Favara M. (2011). The Cost of Acting "Girly": Gender Stereotypes and Educational Choices. Paper presented at the EALE conference, Cyprus September 2011.

Fryer, R. G., & Torelli, P. (2010). An empirical analysis of 'acting white'. *Journal of Public Economics*, *94*(5), 380-396.

Griffith, A. and K. Rask (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 (2), 244–255.

Grodner, A., & Rupp, N. G. (2013). The role of homework in student learning outcomes: Evidence from a field experiment. *The Journal of Economic Education*, *44*(2), 93-109.

Hagedorn, L. S., & Nora, A. (1996). Rethinking admissions criteria in graduate and professional programs. *New Directions for Institutional Research*, *1996*(92), 31-44.

Hoest, A., Jensen, V. M., & Nielsen, L. P. (2013). Increasing the admission rate to upper secondary school: the case of lower secondary school student career guidance. *Education Economics*, *21*(3), 213.

Horn, L. J., X. Chen, and C. Chapman (2003). Getting ready to pay for college: What students and their parents know about the cost of college tuition and what they are doing to find out. National Center for Education Statistics

Hoxby, C., & Turner, S. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, (12-014).

Hu, S., Simmons, A. B., Musoba, G. D., & St John, E. P. (2000). Aptitude vs. merit: What matters in persistence. *The Review of Higher Education*, *24*(2), 131-152.

Hughers, K., & Karp, M. (2004). *School-Based Career Development: A Synthesis of the Literature.* Institute on Education and the Economy Teachers College, Columbia University.

Hughes, K. L., and M. M. Karp. 2004. School-Based Career Development: A Synthesis of the Literature. New York: Columbia University, Institute on Education and the Economy, National Training support Centre.

Humlum M, Kleinjans K, Nielsen H. (2012). An Economic Analysis of identity and Career Choice. Economic Inquiry 50(1), 39–61.

Hurwitz, M., & Howell, J. (2014). Estimating causal impacts of school counselors with regression discontinuity designs. *Journal of Counseling & Development*, *92*(3), 316-327.

Jensen R. (2010). The (Perceived) Returns to Education and the Demand for Schooling. Quarterly Journal of Economics 125(2), 515–548.

Johnson, J., J. Rochkind, A. Ott, & S. DuPont (2010). "Can I get a little advice here? How an overstretched high school guidance system is undermining students' college aspirations,"

Public Agenda and the Bill & Melinda Gates Foundation. Available at http://www.publicagenda.org/files/can-i-get-a-little-advice-here.pdf.

Kaufmann, K. M. (2014). Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns.*Quantitative Economics*, *5*(3), 583-630.

de Koning, B. B., Loyens, S. M., Rikers, R. M., Smeets, G., & van der Molen, H. T. (2014). Impact of binding study advice on study behavior and pre-university education qualification factors in a problem-based psychology bachelor program. *Studies in Higher Education*, *39*(5), 835-847.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education*, *79*(5), 540-563.

Kuijpers, M. & Meijers, F. (2008). Loopbaanleren: onderzoek en praktijk in het onderwijs. *Antwerpen/Apeldoorn: Garant.*

LAKS, ISO, NVS-NVL, PvdA, VVD, CDA (2016). Studiekeuzechecklist: Een aanval op uitval. *Utrecht, June 24, 2016*

Lapan, R., & Harrington, K. (2010). Paving the Road to College: How School Counselors Help Students Succeed. *Center for School Counseling Outcome Research (NJ1)*.

Leonardi, M. (2007). Do parents risk aversion and wealth explain secondary school choice?. *Giornale degli Economisti e Annali di Economia*, 177-206.

Mazza, J., & Hartog, J. (2011). Do they understand the benefits from education? : Evidence on Dutch high school students' rarnings expectations(IZA Discussion Paper, No. 5714; Discussion papers, No. 5714). Bonn: IZA. http://ftp.iza.org/dp5714.pdf

Mellanby, J., Martin, M., & O'Doherty, J. (2000). The 'gender gap'in final examination results at Oxford University. *British Journal of Psychology*, *91*(3), 377-390.

Meredith, M. (2004). Why do universities compete in the ratings game? an empirical analysis of the effects of the us news and world report college rankings. Research in Higher Education 45 (5), 443–461.

Ministry of Education, Culture and Science & Ministry of Social Affairs and Employment (2015). Aanpak Jeugdwerkloosheid. *Den Haag: November 29, 2015*

Ministry of Education, Culture and Science (2016). Betere loopbaanbegeleiding voor leerlingen. *Den Haag: September 28, 2016*

Monks, J. and R. G. Ehrenberg (1999). The impact of us news and world report college rankings on admission outcomes and pricing decisions at selective private institutions. Technical report, National Bureau of Economic Research.

Nguyen T. (2008). Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar. Mimeo MIT.

Nielsen, H. S., & Vissing-Jorgensen, A. (2006). The impact of labor income risk on educational choices: Estimates and implied risk aversion. *Unpublished Manuscript, Department of Economics, University of Aarhus*.

Van Nimwegen, Niels (2016) Universitaire opleidingen zijn te positief over kans op werk. *De Monitor, KRO-NCRV.*

Oberon (2012). Rapportage eindmeting LOB: Directies en decanen VO. September 2012. *Commissioned by VO-Raad. Project Stimulering LOB*

OCW (Ministerie van Onderwijs, Cultuur en Wetenschappen) (2011). Kwaliteit in Verscheidenheid, Strategische Agenda Hoger Onderwijs, Onderzoek en Wetenschap, Ministerie van OCW.

OECD (2004). Career Guidance and Public Policy: Bridging the Gap. *OECD Publications, Paris*

OECD (2008). OECD Reviews of Tertiary Education: Netherlands. *OECD Publication, Paris*

Oliver, L.W., & Spokane, A.R. (1988). Career-intervention outcomes: What contributes to client gain? Journal of Counseling Psychology, 35 (4), 447-462.

Onderwijsraad (2013). Advies overgangen in het onderwijs. *Den Haag: 3 maart 2014*

Oreopoulos, P., & Dunn, R. (2013). Information and College Access: Evidence from a Randomized Field Experiment*. *The Scandinavian Journal of Economics*, *115*(1), 3-26.

Oreopoulos, P., Brown, R. S., & Lavecchia, A. M. (2014). *Pathways to education: An integrated approach to helping at-risk high school students* (No. w20430). National Bureau of

Economic Research.

De Paola, M., & Gioia, F. (2011). Risk aversion and major choice: Evidence from Italian students *Universita della Calabria Working Paper No. 201107.*

Pascarella, E., & Terenzini, P. (2005). How college affects students: A third decade of research, Vol. 2. San Francisco: Jossey-Bass.

Perna, L. W., Li, C., Anderson, R., Thomas, S. L., Rowan-Kenyon, H. T., & Bell, A. (2007). The role of college counseling in shaping college opportunity: Variations across high schools. *The Review of Higher Education, 31*(2), 131-159.

Plank, S.B., & W.J. Jordan (2001). "Effects of information, guidance, and action on postsecondary destinations: A study of talent loss," *American Educational Research Journal, 38(*4*):* 947–979.

Reason, R. D. (2003). Using an ACT-based merit-index to predict between-year retention. *Journal of College Student Retention: Research, Theory & Practice*,*5*(1), 71-87.

ResearchNed (2016). Monitor Beleidsmaatregelen 2015: Studiekeuze, studiegedrag en leengedrag in relatie tot beleidsmaatregelen in het hoger onderwijs, 2006-2015. *Den Haag, ResearchNed.*

ROA (2014).*Evaluatie van het School-Ex 2.0 programma. De rol van studiekeuze-en exitgesprekken in het MBO* (No. 002). Maastricht University, Research Centre for Education and the Labour Market (ROA).

Schippers, M. C., Scheepers, A. W., & Peterson, J. B. (2015). A scalable goal-setting intervention closes both the gender and ethnic minority achievement gap. *Palgrave Communications, 1*.

Scott, M., Bailey, T., & Kienzl, G. (2006). Relative success? Determinants of college graduation rates in public and private colleges in the US. *Research in higher education, 47*(3), 249-279.

Scott, G., Shah, M., Grebennikov, L., & Singh, H. (2008). Improving student retention: A University of Western Sydney case study. *Journal of Institutional Research, 14*(1), 9-23.

Seftor, N. S., Mamun, A., & Schirm, A. (2009). The Impacts of Regular Upward Bound on Postsecondary Outcomes Seven to Nine Years after Scheduled High School Graduation. Final Report. *US Department of Education*.

Sneyers, E., & De Witte, K. (2015). The effect of an academic dismissal policy on dropout, graduation rates and student satisfaction. Evidence from the Netherlands. *Studies In Higher Education*, (10), 1-36. doi:10.1080/03075079.2015.1049143

Stegers-Jager, K. M., Cohen-Schotanus, J., Splinter, T. A., & Themmen, A. P. (2011). Academic dismissal policy for medical students: effect on study progress and help-seeking behaviour. *Medical education*, *45*(10), 987-994.

Stinebrickner TR, Stinebrickner R. (2012). Learning about Academic Ability and the College Dropout Decision. Journal of Labor Economics 30(4), 707–748.

Streufert, P. (2000). The Effect of Underclass Social Isolation on Schooling Choice. *Journal Of Public Economic Theory*, *2*(4), 461-482. doi:10.1111/1097-3923.00046

Tinto, V. (2002). Promoting student retention: Lessons learned from the United States. In *11th Annual Conference of the European Access Network, Prato, Italy* (Vol. 19).

Titus, M. A. (2004). An examination of the influence of institutional context on student persistence at 4-year colleges and universities: A multilevel approach. *Research in higher education*, *45*(7), 673-699.

Titus, M. A. (2006). Understanding college degree completion of students with low socioeconomic status: The influence of the institutional financial context. *Research in Higher Education*, *47*(4), 371-398.

Umbach, P. D., & Porter, S. R. (2002). How do academic departments impact student satisfaction? Understanding the contextual effects of departments. *Research in Higher Education*, *43*(2), 209-234.

UWV (2015). Kiezen voor Kansen: Een voorstel om te komen tot een verbetering van de aansluiting tussen onderwijs en arbeidsmarkt.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*,*331*(6023), 1447-1451.

Whinston S, Sexton TL, Lasoff DL. (1998). Career-Intervention Outcome: a Replication and Extension of Oliver and Spokane. Journal of Counseling Psychology 45(2), 150–165.

Appendix

The appendix provides additional, detailed information and is structured as follows. Appendix A provides tables containing estimation results that were not included in the main text are displayed. Appendix B offers a more detailed description of the control covariates. Subsequently, Appendix C shows descriptive statistics of the main estimates. Lastly, Appendix D presents descriptive statistics of the models that are used to do the robustness checks.

Appendix A: Additional estimation results

TABLE A1

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF USING THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.008	0.000	-0.001	0.002
	(0.010)	(0.011)	(0.012)	(0.012)
B. Heterogeneous effects				
i. Secondary education type				
Average treatment effect		-0.010	-0.010	-0.004
		(0.015)	(0.015)	(0.015)
Average treatment effect * VWO		0.016	0.011	0.004
		(0.021)	(0.026)	(0.025)
ii. Gender				
Average treatment effect		0.001	-0.001	0.002
		(0.015)	(0.017)	(0.017)
Average treatment effect * male		-0.003	-0.001	-0.002
		(0.023)	(0.028)	(0.028)
iii. Socioeconomic status				
Average treatment effect		0.001	-0.000	0.005
		(0.010)	(0.013)	(0.013)
Average treatment effect * low		-0.008	-0.004	-0.019
socioeconomic status		(0.025)	(0.025)	(0.024)
iv. Ethnicity				
Average treatment effect		-0.004	-0.004	-0.000
		(0.011)	(0.013)	(0.012)
Average treatment effect * non-western		0.051	0.045	0.037
immigrant		(0.033)	(0.043)	(0.042)
C. Intensity of exposure to treatment				
Low	-0.006	0.004	0.011	0.015
	(0.027)	(0.018)	(0.018)	(0.017)
Mid	0.005	-0.010	-0.016	-0.013
	(0.015)	(0.013)	(0.015)	(0.015)
High	-0.008	-0.010	-0.003	-0.003
	(0.019)	(0.014)	(0.018)	(0.019)
Unknown	0.028*	0.013	0.008	0.011
	(0.017)	(0.014)	(0.015)	(0.014)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	63003	62988	43704	43691

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the introduction of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A presents the estimated effects comparing students from treatment and control schools. Treatment schools introduced the Qompas career orientation guidance package for graduation cohorts 08/09 or 09/10, whereas control schools introduced this package for graduation cohorts 12/13 or 13/14. Panels B and C present heterogeneous effects and differential effects by intensity as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE A2

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF USING THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION, EXPLOITING ONE YEAR EARLIER INTRODUCTION DATA

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.016	0.008	0.010	-0.002
	(0.014)	(0.014)	(0.016)	(0.017)
B. Heterogeneous effects				
i. Secondary education type				
Average treatment effect		0.005	0.006	-0.013
		(0.019)	(0.016)	(0.017)
Average treatment effect * VWO		0.007	0.004	0.024
		(0.026)	(0.029)	(0.027)
ii. Gender				
Average treatment effect		-0.016	-0.013	-0.027
		(0.022)	(0.022)	(0.024)
Average treatment effect * male		0.049*	0.043	0.051
		(0.028)	(0.032)	(0.033)
iii. Socioeconomic status				
Average treatment effect		0.002	0.004	-0.002
		(0.015)	(0.018)	(0.017)
Average treatment effect * low		0.028	0.030	0.004
socioeconomic status		(0.023)	(0.044)	(0.047)
iv. Ethnicity				
Average treatment effect		0.008	0.016	0.004
		(0.016)	(0.018)	(0.019)
Average treatment effect * non-western		-0.005	-0.029	-0.019
immigrant		(0.042)	(0.049)	(0.048)
C. Intensity of exposure to treatment				
Low	0.021	0.021	0.045**	0.028
	(0.026)	(0.020)	(0.023)	(0.022)
Mid	0.035*	0.012	0.017	0.002
	(0.019)	(0.015)	(0.018)	(0.018)
High	-0.001	-0.016	-0.007	-0.022
	(0.020)	(0.018)	(0.020)	(0.022)
Unknown	0.009	0.009	0.009	-0.002
	(0.017)	(0.016)	(0.018)	(0.020)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	70044	70044	44053	44035

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of the introduction of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A presents the estimated effects comparing students from treatment and control schools. Treatment schools introduced the Qompas career orientation guidance package for graduation cohorts 07/08 or 08/09, whereas control schools introduced this package for graduation cohorts 11/12 or 12/13. Panels B and C present heterogeneous effects and differential effects by intensity as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE A3

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF A SHOCK IN INTENSITY OF USE OF THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON RETENTION RATES

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.035*	0.033**	0.042***	0.042***
	(0.019)	(0.016)	(0.016)	(0.015)
B. Heterogeneous effects				
i. Secondary education type				
Average treatment effect		0.039*	0.058**	0.057**
		(0.024)	(0.025)	(0.026)
Average treatment effect * VWO		-0.015	-0.035	-0.034
		(0.029)	(0.030)	(0.032)
ii. Gender				
Average treatment effect		0.026	0.037	0.041
		(0.025)	(0.026)	(0.025)
Average treatment effect * male		0.011	0.007	0.003
		(0.031)	(0.035)	(0.034)
iii. Socioeconomic status				
Average treatment effect		0.036**	0.046***	0.045***
		(0.017)	(0.017)	(0.016)
Average treatment effect * low		-0.006	-0.004	-0.011
socioeconomic status		(0.035)	(0.043)	(0.045)
iv. Ethnicity				
Average treatment effect		0.035**	0.042***	0.043***
		(0.015)	(0.015)	(0.015)
Average treatment effect * non-		-0.010	0.024	0.000
western immigrant		(0.051)	(0.041)	(0.045)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	36365	36365	36365	36354

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of extensive use of the Qompas career orientation package on the retention rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A is equivalent to table 5. Panel B presents heterogeneous effects as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE A4

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF A SHOCK IN INTENSITY OF USE OF THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON ENROLLMENT

	(1)	(2)	(3)
A. Difference-in-Differences probit			
Average treatment effect	0.004	0.007	0.007
	(0.009)	(0.008)	(0.009)
B. Heterogeneous effects			
i. Secondary education type			
Average treatment effect		0.014*	0.015**
		(0.008)	(0.008)
Average treatment effect * VWO		-0.028	-0.022
		(0.019)	(0.019)
ii. Gender			
Average treatment effect		0.013	0.015*
		(0.009)	(0.009)
Average treatment effect * male		-0.012	-0.013
		(0.011)	(0.012)
iii. Socioeconomic status			
Average treatment effect		0.007	0.008
		(0.008)	(0.008)
Average treatment effect * low socioeconomic		-0.006	-0.018
status		(0.019)	(0.022)
iv. Ethnicity			
Average treatment effect		0.005	0.005
		(0.008)	(0.009)
Average treatment effect * non-western immigrant		0.051*	0.045
		(0.027)	(0.030)
School and year fixed effects	no	yes	yes
Individual characteristics	no	yes	yes
Secondary school characteristics	no	no	yes
Tertiary school characteristics	no	no	no
N	37322	37322	37322

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of extensive use of the Qompas career orientation package on the enrollment rate in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A is equivalent to table 5, with the enrollment outcome as dependent variable. Panel B presents heterogeneous effects as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE A5

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS OF THE EFFECT OF A SHOCK IN INTENSITY OF USE OF THE QOMPAS CAREER ORIENTATION GUIDANCE PACKAGE ON IMMEDIATE ENROLLMENT

	(1)	(2)	(3)	(4)
A. Difference-in-Differences probit				
Average treatment effect	0.037	0.004	0.009	0.007
	(0.030)	(0.019)	(0.019)	(0.021)
B. Heterogeneous effects				
i. Secondary education type				
Average treatment effect		0.005	0.009	0.013
		(0.017)	(0.017)	(0.021)
Average treatment effect * VWO		-0.010	-0.007	-0.019
		(0.023)	(0.025)	(0.027)
ii. Gender				
Average treatment effect		-0.016	-0.016	-0.013
		(0.019)	(0.019)	(0.019)
Average treatment effect * male		0.057*	0.068**	0.059*
		(0.031)	(0.031)	(0.032)
iii. Socioeconomic status				
Average treatment effect		0.010	0.015	0.014
		(0.022)	(0.021)	(0.022)
Average treatment effect * low		-0.033	-0.038	-0.038
socioeconomic status		(0.036)	(0.041)	(0.038)
iv. Ethnicity				
Average treatment effect		0.007	0.012	0.009
		(0.020)	(0.020)	(0.021)
Average treatment effect * non-		-0.053	-0.063	-0.043
western immigrant		(0.041)	(0.041)	(0.043)
School and year fixed effects	no	yes	yes	yes
Individual characteristics	no	yes	yes	yes
Secondary school characteristics	no	no	yes	yes
Tertiary school characteristics	no	no	no	yes
N	36365	34108	34108	33992

Notes: Columns 1, 2, 3, and 4 present difference-in-differences average marginal treatment effects of extensive use of the Qompas career orientation package on immediate enrollment in tertiary education. Clustered standard errors at the school level are reported in parentheses. Specifications of the model are reported as in table 3. Panel A is equivalent to table 5, with the immediate enrollment outcome as dependent variable. Panel B presents heterogeneous effects as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

Appendix B: Description of the covariates in the analyses

The probit regressions in this study have been estimated using several specifications, in which groups of covariates are added to the model one-by-one. This appendix describes the covariates that form these groups. The source of these variables are indicated by the numbers in superscript.

10.2.1. Individual characteristics

Covariate	Description
Age[1]	The age this student had at the 1st of October in his/her graduation year
Gender[1]	Male / Female
Secondary education type[1]	Havo or VWO. An explanation of these education types is provided in the overview of the Dutch education system
SE average grade[1]	The average of grades of non-standardized tests[56]
CE average grade[1]	The average of grades of standardized tests[41]
Double degree[1]	A dummy variable denoted as 1 if the student has more than one secondary education degree – which generally means that after HAVO, the student attained a VWO degree
Profile[1]	A categorical variable denoting the secondary education profile of a student. This may be Culture & Society, Economics & Society, Nature & Health or Nature & Technology.
Ethnicity[1]	A categorical variable, which may be native Dutch, western immigrant or non-western immigrant.
Neighborhood status scores[2]	Scores denoting the livability of the neighborhood of the student in his/her graduation year.
Socioeconomic status[3]	Socioeconomic status scores as developed by SCP.

[56] In Dutch secondary education, students can get two types of grades. During the last 2(Havo) or 3(VWO) years of secondary education, they do several tests that are developed by their teachers (SE grades), which comprehensively determines 50% of the final grade. At the end of their graduation year, students make standardized tests, that are developed by CITO (CE grades), which accounts for the other 50% of the final grade.

10.2.2. Secondary school characteristics

Covariate	Description
Number of graduation candidates[1]	-
Average SE grade[1]	Average non-standardized grade of all graduation candidates
Average CE grade[1]	Average standardized grade of all graduation candidates
Gymnasium[1]	A dummy denoted as 1 if this student attended gymnasium, which is a subtype of VWO. This is similar to VWO, but offers additional courses in Greek and Latin language.
G4	A dummy denoted as 1 if this school is in one of the 4 large cities in the Netherlands (Rotterdam, The Hague, Utrecht or Amsterdam)
G37	A dummy denoted as 1 if this school is in one of 37 other large cities in the Netherlands
Graduation rate[1]	-
Ideology[4]	A categorical variable containing the ideology of this secondary school. (e.g. Protestant, Islamic)
Vision[4]	A categorical variable representing the vision of a school. Most of the schools are regular schools. Some schools in the Netherlands focus more on e.g. self-dependence.
Region[4]	A categorical variable denoting the region (34 regions are distinguished) of the country where the school is located.
Urbanity[4]	A standardized variable denoting the urbanity of the location of the school. This is based on the number of households within a squared kilometer.
Lower secondary education efficiency[4]	The average efficiency of the first three years of secondary education (in terms of e.g. delay). This variable is obtained from data of the Dutch Inspectorate of Education.
postponed tracking[4]	Some schools in the Netherlands track their students 1 or 2 years later than the usual age of 12.
Rate of apcg students[4]	The ratio of apcg students. apcg students are typically students from a very low socioeconomic status and live in a poor neighborhood.

10.2.3. Tertiary education characteristics

Covariate	Description
Institution fixed effects[1]	-
HBO / WO[1]	Research University (WO) or University of Applied Sciences (HBO)
rate of female entrants[5]	The ratio of females of the total number of enrolling students in this specialization.
Number of entrants[5]	-
Academic dismissal policy[5]	Some studies require that freshman obtain a certain number of credits to enroll in sophomore year. This is a dummy variable denoted as 1 if the student enrolls in a specialization with such a policy.
Fixed Quota[5]	A dummy variable denoted as 1 if the student enrolls in a specialization that allows a certain number of students. If more students subscribe for this specialization, a lottery decides which students may enroll.
Selective admission standards[5]	A dummy variable denoted as 1 if the student enrolls in a specialization that has certain requirements to prospective students. This may be a minimum grade. Institutions may also select students using an application process.
Student / personnel ratio[5]	The number of students of the specialization of choice divided by the number of personnel of this specialization.

1. These variables are obtained from DUO, the Dutch executive organization responsible for education (www.duo.nl)
2. The neighborhood status scores (in Dutch: Leefbaarometer), that are obtained from the Dutch Ministry of the Interior and Kingdom Relations, provide scores on 5 aspects of neighborhood livability. First, it includes a score capturing the status of the buildings in the neighborhood, e.g. house prices, home ownership ratios, type of buildings. Second, it scores the quality of the public space, by measuring e.g. noise pollution, air pollution. Third, it scores services in the neighborhood, like supermarkets, schools, and leisure areas. Fourth, it scores the composition of its residents, by measuring e.g. the percentage of unemployed people, ethnicity, income and residential mobility. The last score captures safety in the neighborhood, using e.g. police registrations of crimes (www.leefbaarometer.nl)
3. These scores, that are obtained from the Netherlands Institute for Social Research, capture the socioeconomic , based on the level of education, the level of income, and the labor market position of the residents of a certain neighborhood (www.scp.nl)
4. These variables are obtained from the Dutch Inspectorate of Education (www.onderwijsinspectie.nl)
5. These variables are obtained from NSE and Stichting Studiekeuze123. (studiekeuze123.nl).

Appendix C: Descriptive statistics of the main models

TABLE C1

DESCRIPTIVE STATISTICS OF STUDENTS FROM SCHOOLS WITH SIGNIFICANT POSITIVE FIXED EFFECTS AND FROM SCHOOLS WITH SIGNIFICANT NEGATIVE FIXED EFFECTS ON RETENTION OUTCOMES

	(1)	(2)	(3)
	Schools that do relatively well	Schools that do relatively poor	(1) - (2)
Individual characteristics			
Age	16.817	16.881	-0.064
	(0.003)	(0.005)	(0.064)
Male	0.463	0.474	-0.011
	(0.002)	(0.003)	(0.008)
VWO	0.399	0.441	-0.041
	(0.002)	(0.003)	(0.105)
SE grades	65.602	65.922	-0.320
	(0.023)	(0.034)	(0.397)
CE grades	64.258	63.956	0.302
	(0.028)	(0.042)	(0.303)
Native Dutch	0.815	0.753	0.062**
	(0.002)	(0.003)	(0.029)
Non-western immigrant	0.107	0.161	-0.054**
	(0.001)	(0.002)	(0.025)
Socioeconomic status	0.374	0.070	0.304***
	(0.003)	(0.008)	(0.117)
Secondary school characteristics			
Number of graduation candidates	98.682	99.254	-0.571
	(0.167)	(0.265)	(8.573)
Graduation rate	0.895	0.871	0.023***
	(0.000)	(0.000)	(0.007)
Urbanity	2.315	1.821	0.493***
	(0.004)	(0.006)	(0.172)
Career guidance data			
% of students using the Qompas career	0.201	0.169	0.032
guidance package	(0.002)	(0.002)	(0.041)
Intensity of use	20.722	9.039	11.684**
	(0.221)	(0.299)	(4.629)
Tertiary education outcome			
Retention	0.695	0.647	0.048***
	(0.002)	(0.003)	(0.014)
Switch	0.209	0.258	-0.049***
	(0.002)	(0.003)	(0.007)
Dropout	0.084	0.083	0.001
	(0.001)	(0.002)	(0.010)
Number of schools	46	26	

Notes: Columns 1 and 2 present averages and standard deviations (in parentheses) of several background characteristics of students from schools with a significant positive fixed effect and from schools with a significant negative fixed effect on retention outcomes. These schools are identified based on a regression of the retention outcome on individual and secondary school characteristics, along with secondary school fixed effects. These fixed effects are standardized and subsequently selected to one of these groups if the fixed effect is larger (smaller) than (-)1.5 standard deviation. Subsequently, the averages and standard deviations of these groups are estimated. Column 3 presents the differences between these groups, as well as standard errors that are clustered at the school level in parentheses. Treated schools had a license of the Qompas career orientation package in that year, whereas control schools did not. These differences are estimates based on a regression of the characteristic as a dependent variable and the group indicator as the explanatory variable. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE C2

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS OF QOMPAS USERS (TREATMENT) AND CONTROL SCHOOLS

	(1) Treatment	(2) Control	(3) (1)-(2)
Individual characteristics			
Age	16.814	16.820	-0.006
	(0.002)	(0.001)	(0.015)
Male	0.472	0.471	0.001
	(0.001)	(0.001)	(0.002)
VWO	0.408	0.440	-0.032
	(0.001)	(0.001)	(0.024)
SE grades	65.434	65.984	-0.550***
	(0.016)	(0.007)	(0.091)
CE grades	64.608	64.337	0.271***
	(0.019)	(0.009)	(0.075)
Native Dutch	0.843	0.850	-0.008
	(0.001)	(0.000)	(0.006)
Non-western immigrant	0.093	0.083	0.010*
	(0.001)	(0.000)	(0.006)
Socioeconomic status	0.162	0.130	0.032
	(0.003)	(0.001)	(0.025)
Secondary school characteristics			
Number of graduation candidates	112.896	105.023	7.873***
	(0.128)	(0.057)	(2.074)
Graduation rate	0.882	0.898	-0.017***
	(0.000)	(0.000)	(0.002)
Urbanity	2.355	2.513	-0.158***
	(0.003)	(0.001)	(0.050)
Tertiary school characteristics			
Research university students	0.354	0.376	-0.022
	(0.001)	(0.001)	(0.021)
Studies with academic dismissal policy	0.745	0.739	0.006
	(0.001)	(0.001)	(0.008)
Studies with fixed quota	0.182	0.156	0.026***
	(0.001)	(0.001)	(0.004)
Studies with selective admission standards	0.115	0.072	0.043***
	(0.001)	(0.000)	(0.003)
Tertiary education outcome			
Retention	0.683	0.688	-0.006
	(0.001)	(0.001)	(0.004)
Switch	0.224	0.225	-0.002
	(0.001)	(0.001)	(0.002)
Dropout	0.077	0.077	-0.001
	(0.001)	(0.000)	(0.002)
Number of schools	554	954	
Number of students	116495	539867	

Notes: Columns 1 and 2 present means and standard deviations (in parentheses) of several background characteristics of students and schools in the treatment (1) and control (2) group. Column 3 presents the differences between these groups, as well as standard errors that are clustered at the school level in parentheses. Treated schools had a license of the Qompas career orientation package in that year, whereas control schools did not. These differences are estimates based on a regression of the characteristic as a dependent variable and the treatment indicator as the explanatory variable. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE C3

DESCRIPTIVE STATISTICS OF QOMPAS USERS (TREATED) AND CONTROL GROUP ALLOWING FOR DIFFERENTIAL EFFECTS BY INTENSITY

	(1) Low intensity	(2) Mid intensity	(3) High intensity	(4) Unknown	(5) Not treated	(6) (1) - rest	(7) (2) - rest	(8) (3) - rest	(9) (4) - rest	(10) (5) - rest
Individual characteristics										
Age	16.850	16.826	16.772	16.810	16.820	0.031	0.007	-0.048	-0.010	0.006
	(0.006)	(0.004)	(0.006)	(0.003)	(0.001)	(0.040)	(0.025)	(0.035)	(0.017)	(0.015)
Male	0.467	0.470	0.477	0.472	0.471	-0.004	-0.001	0.006	0.002	-0.001
	(0.004)	(0.003)	(0.004)	(0.002)	(0.001)	(0.006)	(0.004)	(0.006)	(0.003)	(0.002)
VWO	0.514	0.387	0.322	0.414	0.440	0.081	-0.050	-0.115**	-0.022	0.032
	(0.004)	(0.003)	(0.004)	(0.002)	(0.001)	(0.057)	(0.041)	(0.054)	(0.027)	(0.024)
SE grades	65.509	65.348	64.904	65.595	65.984	-0.386*	-0.563***	-1.005***	-0.319***	0.550***
	(0.047)	(0.032)	(0.043)	(0.023)	(0.007)	(0.210)	(0.157)	(0.199)	(0.102)	(0.091)
CE grades	65.145	64.807	64.656	64.355	64.337	0.778***	0.442***	0.277	-0.033	-0.271***
	(0.053)	(0.037)	(0.050)	(0.027)	(0.009)	(0.187)	(0.119)	(0.186)	(0.086)	(0.075)
Native Dutch	0.848	0.848	0.834	0.841	0.850	-0.001	-0.002	-0.015	-0.009	0.008
	(0.003)	(0.002)	(0.003)	(0.002)	(0.000)	(0.012)	(0.007)	(0.019)	(0.007)	(0.006)
Non-western immigrant	0.085	0.088	0.104	0.095	0.083	0.000	0.003	0.019	0.011*	-0.010*
	(0.002)	(0.002)	(0.003)	(0.001)	(0.000)	(0.009)	(0.006)	(0.019)	(0.006)	(0.006)
Socioeconomic status	0.172	0.193	0.183	0.138	0.130	0.038	0.060	0.048	0.002	-0.032
	(0.007)	(0.005)	(0.008)	(0.004)	(0.001)	(0.048)	(0.040)	(0.065)	(0.026)	(0.025)
Secondary school characteristics										
Graduation candidates	126.321	109.893	109.744	111.787	105.023	20.360***	3.636	3.400	5.882**	-7.873***
	(0.413)	(0.229)	(0.341)	(0.182)	(0.057)	(5.505)	(3.206)	(4.396)	(2.521)	(2.074)
Graduation rate	0.886	0.884	0.884	0.879	0.898	-0.010**	-0.012***	-0.012**	-0.018***	0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.004)	(0.005)	(0.003)	(0.002)
Urbanity	2.271	2.303	2.392	2.394	2.513	-0.217**	-0.189**	-0.093	-0.098	0.158***
	(0.008)	(0.005)	(0.008)	(0.004)	(0.001)	(0.110)	(0.074)	(0.113)	(0.065)	(0.050)
Tertiary school characteristics										
Research university students	0.453	0.338	0.278	0.356	0.376	0.083	-0.036	-0.096**	-0.018	0.022
	(0.004)	(0.003)	(0.004)	(0.002)	(0.001)	(0.051)	(0.036)	(0.046)	(0.023)	(0.021)
Academic dismissal policy	0.703	0.699	0.712	0.788	0.739	-0.038**	-0.043***	-0.029	0.052***	-0.006
	(0.004)	(0.003)	(0.004)	(0.002)	(0.001)	(0.017)	(0.015)	(0.020)	(0.008)	(0.008)
Fixed quota	0.226	0.198	0.184	0.163	0.156	0.067***	0.039***	0.023**	0.002	-0.026***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)	(0.010)	(0.007)	(0.011)	(0.005)	(0.004)
Selective admission standards	0.162	0.125	0.118	0.097	0.072	0.084***	0.048***	0.039***	0.019***	-0.043***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.000)	(0.010)	(0.006)	(0.008)	(0.004)	(0.003)
Tertiary education outcome										
Retention	0.698	0.674	0.674	0.685	0.688	0.011	-0.014*	-0.014	-0.002	0.006
	(0.004)	(0.003)	(0.004)	(0.002)	(0.001)	(0.008)	(0.007)	(0.009)	(0.005)	(0.004)
Switch	0.211	0.223	0.224	0.227	0.225	-0.015***	-0.002	-0.001	0.002	0.002
	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)	(0.006)	(0.004)	(0.006)	(0.003)	(0.002)
Dropout	0.066	0.078	0.081	0.078	0.077	-0.012**	0.000	0.003	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.006)	(0.004)	(0.006)	(0.003)	(0.002)
Number of schools	109	180	95	426	954					
Number of students	14801	29461	14742	57491	539867					

Notes: Columns 1, 2, 3, 4, and 5 present means and standard deviations (in parentheses) of several background characteristics of students and schools of the treatment group by intensity (1,2,3, and 4) and the control group(5). Columns 6, 7, 8, 9, and 10 present the differences between a specific group and the other groups, as well as standard errors that are clustered at the school level in parentheses. These differences are estimates based on a regression of the characteristic as a dependent variable and the treatment indicator as the explanatory variable. Treated schools use the Qompas career orientation package to provide guidance, whereas control do not use this package. Treatment intensity is determined as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE C4

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATED AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION

	Graduation years 04/05 - 07/08			Graduation years 10/11 and 11/12		
	(1) Treatment	(2) Synthetic control	(3) (1)-(2)	(4) Treatment	(5) Synthetic control	(6) (4)-(5)
Individual characteristics						
Age	16.883	16.793	0.090	16.889	16.812	0.078
	(0.004)	(0.008)	(0.062)	(0.006)	(0.010)	(0.052)
Male	0.463	0.462	0.001	0.471	0.480	-0.009
	(0.003)	(0.005)	(0.008)	(0.004)	(0.007)	(0.009)
VWO	0.530	0.400	0.130	0.522	0.392	0.130
	(0.003)	(0.005)	(0.096)	(0.004)	(0.007)	(0.095)
SE grades	67.050	66.373	0.677**	65.985	65.033	0.953**
	(0.035)	(0.050)	(0.331)	(0.049)	(0.076)	(0.404)
CE grades	64.211	64.121	0.090	64.568	63.981	0.587**
	(0.045)	(0.069)	(0.300)	(0.056)	(0.087)	(0.284)
Native Dutch	0.837	0.860	-0.023	0.837	0.859	-0.023
	(0.002)	(0.004)	(0.017)	(0.003)	(0.005)	(0.017)
Non-western immigrant	0.087	0.073	0.014	0.094	0.080	0.014
	(0.002)	(0.002)	(0.015)	(0.002)	(0.004)	(0.015)
Socioeconomic status	0.120	0.084	0.036	0.178	0.122	0.056
	(0.006)	(0.009)	(0.075)	(0.008)	(0.013)	(0.087)
Secondary school characteristics						
Number of graduation candidates	88.776	86.953	1.822	101.566	96.379	5.187
	(0.220)	(0.350)	(6.390)	(0.297)	(0.516)	(7.532)
Graduation rate	0.917	0.924	-0.007	0.882	0.874	0.008
	(0.000)	(0.000)	(0.007)	(0.000)	(0.001)	(0.008)
Urbanity	2.298	2.561	-0.263	2.264	2.544	-0.280
	(0.007)	(0.016)	(0.236)	(0.008)	(0.019)	(0.232)
Tertiary school characteristics						
Research university students	0.457	0.339	0.118	0.457	0.333	0.124
	(0.003)	(0.005)	(0.079)	(0.004)	(0.006)	(0.080)
Studies with academic dismissal	0.713	0.772	-0.060	0.725	0.739	-0.014
policy	(0.003)	(0.004)	(0.042)	(0.004)	(0.007)	(0.021)
Studies with fixed quota	0.130	0.113	0.017	0.168	0.152	0.016
	(0.002)	(0.003)	(0.013)	(0.003)	(0.005)	(0.014)
Studies with selective admission	0.030	0.026	0.004	0.096	0.078	0.018*
standards	(0.001)	(0.002)	(0.003)	(0.003)	(0.004)	(0.011)
Tertiary education outcome						
Retention	0.687	0.686	0.002	0.702	0.692	0.010
	(0.003)	(0.005)	(0.012)	(0.004)	(0.007)	(0.015)
Switch	0.229	0.230	-0.001	0.225	0.227	-0.003
	(0.003)	(0.005)	(0.008)	(0.004)	(0.006)	(0.008)
Dropout	0.084	0.084	-0.001	0.073	0.081	-0.007
	(0.002)	(0.003)	(0.008)	(0.002)	(0.004)	(0.010)
Number of schools	90	58		90	58	
Number of students	24031	18227		13775	9709	

Notes: Columns 1, 2, 4, and 5 present means and standard deviations (in parentheses) of several background characteristics of students and schools in the treatment (1 and 4) and control (2 and 5) group. Columns 3 and 6 present the differences between these groups, as well as standard errors that are clustered at the school level in parentheses. These differences are estimates based on a regression of the characteristic as a dependent variable and the treatment indicator as the explanatory variable. Treatment and synthetic control group are determined as in table 4. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE C5

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATED AND SYNTHETIC CONTROL GROUP EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION, ALLOWING FOR DIFFERENTIAL EFFECTS BY INTENSITY

	(1) Low intensity	(2) Mid intensity	(3) High intensity	(4) Unknown	(5) Not treated	(6) (1) - rest	(7) (2) - rest	(8) (3) - rest	(9) (4) - rest	(10) (5) - rest
Individual characteristics										
Age	16.783	16.906	16.840	16.885	16.799	-0.065	0.069	-0.005	0.066	-0.076
	(0.009)	(0.006)	(0.009)	(0.003)	(0.005)	(0.053)	(0.043)	(0.056)	(0.041)	(0.054)
Male	0.458	0.467	0.478	0.467	0.471	-0.011	-0.002	0.010	-0.002	0.003
	(0.006)	(0.005)	(0.006)	(0.002)	(0.003)	(0.009)	(0.007)	(0.009)	(0.006)	(0.007)
VWO	0.413	0.514	0.402	0.552	0.396	-0.058	0.051	-0.070	0.134*	-0.123
	(0.006)	(0.005)	(0.006)	(0.002)	(0.003)	(0.094)	(0.072)	(0.089)	(0.070)	(0.095)
SE grades	65.880	65.731	65.265	66.806	65.690	-0.236	-0.416	-0.884***	1.128***	-0.700**
	(0.075)	(0.051)	(0.070)	(0.027)	(0.034)	(0.482)	(0.261)	(0.330)	(0.250)	(0.339)
CE grades	64.703	64.748	64.589	64.434	64.295	0.285	0.353*	0.164	0.001	-0.233
	(0.087)	(0.058)	(0.081)	(0.034)	(0.043)	(0.468)	(0.192)	(0.310)	(0.229)	(0.236)
Native Dutch	0.835	0.861	0.814	0.834	0.856	-0.010	0.018	-0.032	-0.018	0.019
	(0.005)	(0.003)	(0.005)	(0.002)	(0.002)	(0.022)	(0.012)	(0.038)	(0.013)	(0.017)
Non-western immigrant	0.097	0.073	0.121	0.092	0.079	0.011	-0.015	0.036	0.009	-0.013
	(0.004)	(0.002)	(0.004)	(0.001)	(0.002)	(0.018)	(0.010)	(0.038)	(0.011)	(0.015)
Socioeconomic status	0.124	0.219	0.227	0.141	0.097	-0.012	0.094	0.097	0.009	-0.063
	(0.012)	(0.008)	(0.012)	(0.005)	(0.006)	(0.108)	(0.067)	(0.100)	(0.060)	(0.080)
Secondary school characteristics										
Graduation candidates	117.551	103.021	98.685	93.577	90.877	23.803**	8.935*	3.851	-2.341	-7.053
	(0.567)	(0.303)	(0.425)	(0.173)	(0.223)	(9.264)	(5.200)	(6.297)	(5.319)	(6.864)
Graduation rate	0.874	0.880	0.876	0.907	0.898	-0.024***	-0.019***	-0.022**	0.016***	0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.009)	(0.005)	(0.009)	(0.005)	(0.006)
Urbanity	2.289	2.340	2.320	2.253	2.547	-0.107	-0.056	-0.074	-0.213	0.267
	(0.013)	(0.007)	(0.012)	(0.005)	(0.009)	(0.244)	(0.136)	(0.182)	(0.168)	(0.232)
Tertiary school characteristics										
Research university students	0.379	0.465	0.359	0.476	0.339	-0.030	0.065	-0.051	0.111*	-0.116
	(0.006)	(0.005)	(0.007)	(0.002)	(0.003)	(0.087)	(0.063)	(0.078)	(0.060)	(0.079)
Academic dismissal policy	0.733	0.723	0.747	0.742	0.760	-0.016	-0.027	-0.001	-0.009	0.022
	(0.006)	(0.004)	(0.006)	(0.002)	(0.003)	(0.033)	(0.022)	(0.031)	(0.021)	(0.018)
Fixed quota	0.182	0.213	0.176	0.157	0.153	0.020	0.055***	0.013	-0.010	-0.019
	(0.005)	(0.004)	(0.005)	(0.002)	(0.002)	(0.016)	(0.011)	(0.017)	(0.011)	(0.013)
Selective admission standards	0.104	0.138	0.113	0.067	0.071	0.024*	0.064***	0.034**	-0.022***	-0.016**
	(0.004)	(0.003)	(0.004)	(0.001)	(0.002)	(0.013)	(0.009)	(0.013)	(0.007)	(0.007)
Tertiary education outcome										
Retention	0.692	0.708	0.683	0.697	0.695	-0.004	0.013	-0.014	0.001	-0.002
	(0.006)	(0.004)	(0.006)	(0.002)	(0.003)	(0.019)	(0.011)	(0.015)	(0.010)	(0.012)
Switch	0.225	0.224	0.239	0.227	0.226	-0.001	-0.003	0.013	0.000	-0.001
	(0.006)	(0.004)	(0.006)	(0.002)	(0.003)	(0.011)	(0.007)	(0.009)	(0.005)	(0.006)
Dropout	0.083	0.069	0.078	0.076	0.079	0.006	-0.010	0.001	-0.001	0.004
	(0.004)	(0.002)	(0.004)	(0.001)	(0.002)	(0.010)	(0.007)	(0.009)	(0.006)	(0.008)
Number of schools	43	67	43	90	58					
Number of students	6142	12072	5945	42840	46635					

Notes: This table presents means, standard deviations (in parentheses), and differences as in table C3. Now, treatment and control groups are determined as in table 4. Treatment intensity is determined as in table 3. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE C6

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATED AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD EXPLOITING VARIATION IN THE YEAR OF INTRODUCTION, EXPLOITING ONE YEAR EARLIER INTRODUCTION DATA

	Graduation years 04/05 – 06/07			Graduation years 09/10 and 10/11		
	(1) Treatment	(2) Synthetic control	(3) (1)-(2)	(4) Treatment	(5) Synthetic control	(6) (4)-(5)
Individual characteristics						
Age	16,771	16,759	0,012	16,789	16,749	0,041
	(0,006)	(0,013)	(0,081)	(0,006)	(0,015)	(0,071)
Male	0,454	0,453	0,002	0,469	0,495	-0.025**
	(0,004)	(0,009)	(0,010)	(0,005)	(0,010)	(0,011)
VWO	0,372	0,361	0,011	0,356	0,343	0,013
	(0,004)	(0,008)	(0,117)	(0,004)	(0,009)	(0,116)
SE grades	66,369	66,255	0,114	65,634	65,495	0,139
	(0,042)	(0,085)	(0,469)	(0,048)	(0,098)	(0,355)
CE grades	63,733	63,681	0,053	63,790	62,791	1.000***
	(0,054)	(0,114)	(0,428)	(0,059)	(0,129)	(0,333)
Native Dutch	0,821	0,813	0,008	0,821	0,822	-0,001
	(0,003)	(0,007)	(0,033)	(0,003)	(0,008)	(0,037)
Non-western immigrant	0,107	0,113	-0,006	0,113	0,110	0,003
	(0,002)	(0,006)	(0,027)	(0,003)	(0,006)	(0,032)
Socioeconomic status	0,081	0,061	0,020	0,159	0,039	0,119
	(0,008)	(0,020)	(0,159)	(0,008)	(0,027)	(0,253)
Secondary school characteristics						
Number of graduation candidates	101,210	76,053	25.157***	112,680	90,180	22.500***
	(0,324)	(0,488)	(9,130)	(0,344)	(0,558)	(8,096)
Graduation rate	0,910	0,916	-0,006	0,875	0,862	0,014
	(0,000)	(0,001)	(0,013)	(0,001)	(0,001)	(0,012)
Urbanity	2,233	2,409	-0,176	2,253	2,474	-0,221
	(0,008)	(0,034)	(0,342)	(0,009)	(0,035)	(0,420)
Tertiary school characteristics						
Research university students	0,330	0,306	0,023	0,311	0,297	0,014
	(0,004)	(0,007)	(0,096)	(0,004)	(0,008)	(0,100)
Studies with academic dismissal policy	0,798	0,835	-0,037	0,854	0,851	0,003
	(0,003)	(0,005)	(0,043)	(0,003)	(0,007)	(0,013)
Studies with fixed quota	0,116	0,120	-0,004	0,147	0,149	-0,001
	(0,003)	(0,006)	(0,016)	(0,003)	(0,008)	(0,016)
Studies with selective admission standards	0,027	0,027	0,000	0,056	0,058	-0,002
	(0,001)	(0,003)	(0,006)	(0,002)	(0,005)	(0,010)
Number of schools	72	156		72	156	
Number of students	15647	32346		12171	23447	

Notes: This table is similar to table C4. Now, treated schools introduced the Qompas career orientation package for graduation cohorts 07/08 or 08/09, whereas control schools introduced this package for graduation cohorts 11/12 or 12/13. Columns 1, 2, and 3 present these statistics in the pre-treatment period (graduation cohorts 04/05 – 06/07), whereas columns 4, 5, and 6 present these statistics in the post-treatment period (graduation cohorts 09/10 – 10/11). Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

Appendix D: Descriptive statistics of robustness check models

TABLE D1

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, WITHOUT APPLYING THE SYNTHETIC CONTROL METHOD

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Control	(1)-(2)	Treatment	Control	(4)-(5)
Individual characteristics						
Age	16,869	16,782	0,086	16,860	16,805	0,055
	(0,009)	(0,005)	(0,102)	(0,015)	(0,008)	(0,096)
Male	0,461	0,471	-0,010	0,457	0,470	-0,013
	(0,006)	(0,003)	(0,010)	(0,011)	(0,006)	(0,014)
VWO	0,510	0,444	0,066	0,495	0,452	0,043
	(0,006)	(0,003)	(0,162)	(0,011)	(0,006)	(0,162)
SE grades	66,714	66,141	0,573	65,675	65,596	0,079
	(0,070)	(0,038)	(0,598)	(0,121)	(0,069)	(0,565)
CE grades	64,839	64,271	0,568	65,076	65,617	-0,541
	(0,086)	(0,047)	(0,475)	(0,138)	(0,079)	(0,538)
Native Dutch	0,865	0,849	0,016	0,863	0,856	0,007
	(0,004)	(0,002)	(0,022)	(0,007)	(0,004)	(0,024)
Non-western immigrant	0,065	0,084	-0,019	0,075	0,078	-0,003
	(0,003)	(0,002)	(0,017)	(0,006)	(0,003)	(0,020)
Socioeconomic status	0,115	0,165	-0,050	0,138	0,177	-0,038
	(0,010)	(0,006)	(0,142)	(0,019)	(0,010)	(0,171)
Secondary school characteristics						
Number of graduation candidates	98,550	108,749	-10,200	103,576	117,289	-13,713
	(0,463)	(0,256)	(11,388)	(0,807)	(0,464)	(12,710)
Graduation rate	0,914	0,901	0,013	0,889	0,894	-0,005
	(0,001)	(0,000)	(0,009)	(0,001)	(0,001)	(0,012)
Urbanity	2,694	2,380	0,315	2,525	2,392	0,133
	(0,014)	(0,006)	(0,310)	(0,028)	(0,011)	(0,386)
Tertiary school characteristics						
Research university students	0,434	0,387	0,047	0,422	0,404	0,019
	(0,006)	(0,003)	(0,136)	(0,011)	(0,006)	(0,141)
Studies with academic dismissal	0,762	0,744	0,018	0,682	0,626	0,056
policy	(0,005)	(0,003)	(0,047)	(0,010)	(0,006)	(0,064)
Studies with fixed quota	0,145	0,140	0,004	0,217	0,214	0,003
	(0,004)	(0,002)	(0,023)	(0,009)	(0,005)	(0,031)
Studies with selective admission	0,059	0,050	0,009	0,147	0,142	0,005
standards	(0,003)	(0,002)	(0,011)	(0,008)	(0,004)	(0,028)
Number of schools	14	42		15	42	
Number of students	6381	20649		2173	7162	

Notes: This table is similar to table 2. Now, the synthetic control method is not applied to construct a synthetic control group. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D2

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, EXCLUDING SCHOOLS WITH ON AVERAGE LESS THAN 30 GRADUATES

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,869	16,760	0,109	16,860	16,779	0,081
	(0,009)	(0,007)	(0,107)	(0,015)	(0,011)	(0,097)
Male	0,461	0,473	-0,012	0,457	0,466	-0,009
	(0,006)	(0,005)	(0,012)	(0,011)	(0,009)	(0,016)
VWO	0,510	0,404	0,106	0,495	0,397	0,098
	(0,006)	(0,005)	(0,176)	(0,011)	(0,008)	(0,176)
SE grades	66,714	65,865	0,849	65,675	65,487	0,188
	(0,070)	(0,053)	(0,652)	(0,121)	(0,099)	(0,647)
CE grades	64,839	63,875	0.964**	65,076	65,130	-0,054
	(0,086)	(0,068)	(0,454)	(0,138)	(0,115)	(0,600)
Native Dutch	0,865	0,869	-0,005	0,863	0,878	-0,015
	(0,004)	(0,003)	(0,025)	(0,007)	(0,005)	(0,024)
Non-western immigrant	0,065	0,073	-0,008	0,075	0,063	0,013
	(0,003)	(0,002)	(0,018)	(0,006)	(0,003)	(0,019)
Socioeconomic status	0,115	0,089	0,026	0,138	0,067	0,071
	(0,010)	(0,008)	(0,150)	(0,019)	(0,014)	(0,183)
Secondary school characteristics						
Number of graduation candidates	98,550	97,378	1,171	103,576	106,115	-2,539
	(0,463)	(0,285)	(11,227)	(0,807)	(0,494)	(12,487)
Graduation rate	0,914	0,896	0.018*	0,889	0,889	0,000
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,014)
Urbanity	2,694	2,681	0,013	2,525	2,717	-0,192
	(0,014)	(0,011)	(0,365)	(0,028)	(0,020)	(0,442)
Tertiary school characteristics						
Research university students	0,434	0,334	0,100	0,422	0,336	0,086
	(0,006)	(0,005)	(0,142)	(0,011)	(0,008)	(0,146)
Studies with academic dismissal policy	0,762	0,767	-0,005	0,682	0,628	0,054
	(0,005)	(0,004)	(0,047)	(0,010)	(0,008)	(0,067)
Studies with fixed quota	0,145	0,136	0,009	0,217	0,212	0,005
	(0,004)	(0,004)	(0,023)	(0,009)	(0,007)	(0,033)
Studies with selective admission standards	0,059	0,049	0,011	0,147	0,142	0,005
	(0,003)	(0,002)	(0,011)	(0,008)	(0,006)	(0,029)
Number of schools	15	43		15	43	
Number of students	6381	20819		2173	7231	

Notes: This table is similar to table 2. Compared to the model in table 2, in which schools with on average less than 40 students are excluded from the sample, schools with on average less than 30 students are excluded from the sample. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D3

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, EXCLUDING SCHOOLS WITH ON AVERAGE LESS THAN 50 GRADUATES

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,870	16,837	0,033	16,862	16,844	0,018
	(0,009)	(0,009)	(0,113)	(0,015)	(0,015)	(0,101)
Male	0,458	0,465	-0,007	0,453	0,466	-0,012
	(0,006)	(0,006)	(0,010)	(0,011)	(0,011)	(0,015)
VWO	0,523	0,506	0,017	0,508	0,478	0,030
	(0,006)	(0,006)	(0,194)	(0,011)	(0,011)	(0,196)
SE grades	66,771	66,507	0,264	65,715	66,020	-0,304
	(0,071)	(0,069)	(0,742)	(0,123)	(0,123)	(0,659)
CE grades	64,846	64,081	0,766	65,061	64,798	0,263
	(0,087)	(0,088)	(0,517)	(0,141)	(0,153)	(0,828)
Native Dutch	0,867	0,865	0,002	0,865	0,883	-0,017
	(0,004)	(0,004)	(0,026)	(0,007)	(0,006)	(0,025)
Non-western immigrant	0,065	0,066	-0,001	0,075	0,061	0,014
	(0,003)	(0,003)	(0,017)	(0,006)	(0,005)	(0,020)
Socioeconomic status	0,100	0,106	-0,006	0,118	0,063	0,055
	(0,011)	(0,012)	(0,190)	(0,019)	(0,021)	(0,228)
Secondary school characteristics						
Number of graduation candidates	100,091	93,658	6,433	105,142	103,048	2,094
	(0,458)	(0,283)	(10,964)	(0,802)	(0,479)	(12,182)
Graduation rate	0,915	0,906	0,009	0,890	0,880	0,009
	(0,001)	(0,001)	(0,012)	(0,001)	(0,001)	(0,017)
Urbanity	2,676	2,658	0,018	2,485	2,754	-0,269
	(0,014)	(0,017)	(0,425)	(0,028)	(0,032)	(0,545)
Tertiary school characteristics						
Research university students	0,443	0,425	0,018	0,432	0,413	0,019
	(0,006)	(0,006)	(0,160)	(0,011)	(0,011)	(0,166)
Studies with academic dismissal policy	0,759	0,740	0,019	0,687	0,688	0,000
	(0,005)	(0,005)	(0,056)	(0,010)	(0,009)	(0,066)
Studies with fixed quota	0,146	0,157	-0,011	0,217	0,219	-0,001
	(0,005)	(0,005)	(0,026)	(0,009)	(0,009)	(0,039)
Studies with selective admission standards	0,060	0,054	0,007	0,149	0,146	0,003
	(0,003)	(0,003)	(0,011)	(0,008)	(0,008)	(0,035)
Number of schools	14	41		14	41	
Number of students	6220	20403		2115	7069	

Notes: This table is similar to table 2. Compared to the model in table 2, in which schools with on average less than 40 students are excluded from the sample, schools with on average less than 50 students are excluded from the sample. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D4

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, WITHOUT EXCLUDING SMALL SCHOOLS

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,874	16,768	0,107	16,870	16,789	0,081
	(0,008)	(0,007)	(0,104)	(0,015)	(0,011)	(0,095)
Male	0,460	0,472	-0,012	0,458	0,466	-0,009
	(0,006)	(0,005)	(0,012)	(0,011)	(0,008)	(0,015)
VWO	0,518	0,409	0,108	0,504	0,405	0,100
	(0,006)	(0,005)	(0,170)	(0,011)	(0,008)	(0,170)
SE grades	66,713	65,897	0,816	65,685	65,458	0,228
	(0,069)	(0,050)	(0,635)	(0,120)	(0,093)	(0,615)
CE grades	64,786	63,909	0.878**	65,042	65,150	-0,108
	(0,086)	(0,064)	(0,442)	(0,137)	(0,107)	(0,569)
Native Dutch	0,852	0,865	-0,012	0,850	0,873	-0,023
	(0,004)	(0,003)	(0,027)	(0,008)	(0,005)	(0,027)
Non-western immigrant	0,078	0,076	0,002	0,090	0,067	0,023
	(0,003)	(0,002)	(0,021)	(0,006)	(0,004)	(0,024)
Socioeconomic status	0,086	0,109	-0,023	0,111	0,084	0,027
	(0,011)	(0,007)	(0,144)	(0,020)	(0,014)	(0,177)
Secondary school characteristics						
Number of graduation candidates	97,323	97,030	0,293	102,202	105,095	-2,893
	(0,471)	(0,271)	(11,151)	(0,821)	(0,492)	(12,432)
Graduation rate	0,913	0,896	0.018*	0,887	0,888	-0,002
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,014)
Urbanity	2,667	2,629	0,038	2,497	2,658	-0,161
	(0,014)	(0,010)	(0,345)	(0,028)	(0,018)	(0,418)
Tertiary school characteristics						
Research university students	0,443	0,342	0,101	0,433	0,346	0,087
	(0,006)	(0,004)	(0,139)	(0,011)	(0,008)	(0,142)
Studies with academic dismissal policy	0,757	0,768	-0,011	0,685	0,631	0,054
	(0,005)	(0,004)	(0,047)	(0,010)	(0,008)	(0,064)
Studies with fixed quota	0,145	0,136	0,010	0,220	0,211	0,009
	(0,004)	(0,003)	(0,023)	(0,009)	(0,007)	(0,032)
Studies with selective admission standards	0,060	0,049	0,011	0,151	0,141	0,009
	(0,003)	(0,002)	(0,010)	(0,008)	(0,006)	(0,029)
Number of schools	16	43		16	43	
Number of students	6482	20819		2214	7231	

Notes: This table is similar to table 2. Compared to the model in table 2, in which schools with on average less than 40 students are excluded from the sample, no small schools are excluded from the sample. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D5

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING THE UPPER 10% OF SHOCKS IN INTENTITY OF USE

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,842	16,641	0,201	16,810	16,670	0,140
	(0,016)	(0,012)	(0,217)	(0,028)	(0,020)	(0,199)
Male	0,458	0,518	-0.060**	0,475	0,466	0,008
	(0,012)	(0,009)	(0,029)	(0,021)	(0,016)	(0,028)
VWO	0,550	0,270	0,280	0,559	0,264	0,295
	(0,012)	(0,007)	(0,293)	(0,020)	(0,012)	(0,292)
SE grades	66,385	65,504	0,881	65,598	64,723	0,875
	(0,118)	(0,087)	(0,672)	(0,220)	(0,158)	(0,929)
CE grades	64,530	63,643	0.887**	64,601	64,650	-0,049
	(0,158)	(0,118)	(0,410)	(0,246)	(0,180)	(0,653)
Native Dutch	0,906	0,892	0,014	0,912	0,901	0,012
	(0,007)	(0,006)	(0,033)	(0,012)	(0,009)	(0,031)
Non-western immigrant	0,046	0,060	-0,014	0,045	0,044	0,002
	(0,005)	(0,005)	(0,024)	(0,009)	(0,006)	(0,018)
Socioeconomic status	0,359	0,362	-0,004	0,354	0,410	-0,056
	(0,012)	(0,010)	(0,146)	(0,022)	(0,016)	(0,169)
Secondary school characteristics						
Number of graduation candidates	86,338	115,691	-29,353	89,594	120,974	-31,380
	(0,842)	(0,996)	(31,242)	(1,555)	(1,998)	(35,119)
Graduation rate	0,925	0,885	0.040***	0,891	0,901	-0,010
	(0,001)	(0,001)	(0,014)	(0,001)	(0,002)	(0,022)
Urbanity	3,415	3,017	0,398	3,487	3,033	0,454
	(0,018)	(0,017)	(0,477)	(0,029)	(0,028)	(0,447)
Tertiary school characteristics						
Research university students	0,435	0,216	0,219	0,441	0,216	0,225
	(0,012)	(0,006)	(0,222)	(0,020)	(0,010)	(0,228)
Studies with academic dismissal policy	0,823	0,835	-0,012	0,758	0,558	0.199**
	(0,009)	(0,006)	(0,053)	(0,018)	(0,016)	(0,093)
Studies with fixed quota	0,138	0,121	0,018	0,185	0,185	0,000
	(0,008)	(0,006)	(0,029)	(0,016)	(0,012)	(0,027)
Studies with selective admission standards	0,048	0,036	0,012	0,115	0,113	0,003
	(0,005)	(0,003)	(0,013)	(0,013)	(0,009)	(0,027)
Number of schools	5	46		5	46	
Number of students	1818	22194		594	7676	

Notes: This table is similar to table 2. Compared to the model in table 2, only the upper 10% (instead of 20%) in the intensity growth distribution is defined as a major shock. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D6

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING THE UPPER 15% OF SHOCKS IN INTENTITY OF USE

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,997	16,713	0,284	16,968	16,726	0,242
	(0,010)	(0,011)	(0,190)	(0,018)	(0,020)	(0,188)
Male	0,449	0,469	-0,020	0,448	0,459	-0,010
	(0,008)	(0,007)	(0,017)	(0,014)	(0,013)	(0,024)
VWO	0,799	0,447	0,352	0,804	0,435	0.369*
	(0,006)	(0,007)	(0,217)	(0,011)	(0,013)	(0,214)
SE grades	67,687	65,839	1.847**	66,518	65,141	1.377**
	(0,092)	(0,080)	(0,758)	(0,164)	(0,136)	(0,678)
CE grades	65,351	63,985	1.366***	65,705	65,123	0,582
	(0,113)	(0,097)	(0,512)	(0,185)	(0,158)	(0,636)
Native Dutch	0,885	0,894	-0,009	0,886	0,901	-0,014
	(0,005)	(0,004)	(0,033)	(0,009)	(0,007)	(0,033)
Non-western immigrant	0,051	0,061	-0,009	0,052	0,052	0,000
	(0,003)	(0,003)	(0,020)	(0,006)	(0,005)	(0,020)
Socioeconomic status	0,294	0,145	0,149	0,344	0,083	0.261**
	(0,011)	(0,010)	(0,106)	(0,020)	(0,019)	(0,132)
Secondary school characteristics						
Number of graduation candidates	89,697	101,093	-11,397	92,402	105,058	-12,655
	(0,506)	(0,613)	(16,378)	(0,853)	(1,071)	(16,960)
Graduation rate	0,923	0,899	0.024**	0,908	0,891	0,017
	(0,001)	(0,001)	(0,011)	(0,001)	(0,002)	(0,016)
Urbanity	2,777	2,626	0,151	2,716	2,624	0,091
	(0,015)	(0,011)	(0,395)	(0,030)	(0,019)	(0,436)
Tertiary school characteristics						
Research university students	0,675	0,379	0,295	0,686	0,377	0.309*
	(0,007)	(0,007)	(0,181)	(0,013)	(0,012)	(0,183)
Studies with academic dismissal policy	0,700	0,737	-0,036	0,746	0,622	0,124
	(0,007)	(0,007)	(0,065)	(0,012)	(0,013)	(0,076)
Studies with fixed quota	0,185	0,127	0.057*	0,268	0,182	0,085
	(0,006)	(0,005)	(0,032)	(0,012)	(0,009)	(0,053)
Studies with selective admission	0,077	0,043	0.033**	0,187	0,130	0,057
standards	(0,004)	(0,003)	(0,013)	(0,011)	(0,008)	(0,045)
Number of schools	10	44		10	44	
Number of students	4072	21211		1337	7342	

Notes: This table is similar to table 2. Compared to the model in table 2, only the upper 15% (instead of 20%) in the intensity growth distribution is defined as a major shock. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D7

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING THE UPPER 25% OF SHOCKS IN INTENTITY OF USE

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,864	16,771	0,093	16,882	16,789	0,093
	(0,008)	(0,008)	(0,108)	(0,014)	(0,013)	(0,094)
Male	0,462	0,474	-0,012	0,468	0,469	-0,001
	(0,006)	(0,006)	(0,012)	(0,010)	(0,010)	(0,019)
VWO	0,513	0,445	0,068	0,489	0,432	0,057
	(0,006)	(0,006)	(0,183)	(0,010)	(0,011)	(0,184)
SE grades	66,700	66,054	0,646	65,528	65,714	-0,186
	(0,064)	(0,065)	(0,669)	(0,111)	(0,123)	(0,729)
CE grades	64,626	63,978	0,648	65,047	65,114	-0,068
	(0,079)	(0,084)	(0,460)	(0,127)	(0,143)	(0,693)
Native Dutch	0,865	0,880	-0,015	0,864	0,890	-0,026
	(0,004)	(0,003)	(0,026)	(0,007)	(0,005)	(0,024)
Non-western immigrant	0,064	0,064	0,000	0,074	0,055	0,019
	(0,003)	(0,002)	(0,017)	(0,005)	(0,003)	(0,018)
Socioeconomic status	0,173	0,101	0,072	0,209	0,082	0,127
	(0,009)	(0,008)	(0,147)	(0,017)	(0,016)	(0,179)
Secondary school characteristics						
Number of graduation candidates	96,092	96,957	-0,865	102,152	105,910	-3,758
	(0,426)	(0,316)	(10,790)	(0,753)	(0,533)	(12,034)
Graduation rate	0,912	0,900	0,011	0,891	0,891	0,000
	(0,001)	(0,001)	(0,009)	(0,001)	(0,001)	(0,015)
Urbanity	2,740	2,774	-0,034	2,597	2,844	-0,247
	(0,012)	(0,015)	(0,383)	(0,024)	(0,026)	(0,459)
Tertiary school characteristics						
Research university students	0,435	0,361	0,073	0,414	0,360	0,054
	(0,006)	(0,006)	(0,142)	(0,010)	(0,010)	(0,147)
Studies with academic dismissal policy	0,759	0,760	-0,001	0,684	0,640	0,044
	(0,005)	(0,005)	(0,050)	(0,009)	(0,010)	(0,064)
Studies with fixed quota	0,143	0,142	0,001	0,210	0,216	-0,006
	(0,004)	(0,004)	(0,022)	(0,008)	(0,009)	(0,035)
Studies with selective admission	0,056	0,050	0,006	0,143	0,145	-0,003
standards	(0,003)	(0,003)	(0,010)	(0,007)	(0,008)	(0,030)
Number of schools	18	40		18	40	
Number of students	7450	19639		2566	6811	

Notes: This table is similar to table 2. Compared to the model in table 2, the upper 25% (instead of 20%) in the intensity growth distribution is defined as a major shock. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D8

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING THE UPPER 30% OF SHOCKS IN INTENTITY OF USE

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,819	16,724	0,095	16,837	16,753	0,084
	(0,007)	(0,011)	(0,115)	(0,012)	(0,019)	(0,092)
Male	0,468	0,464	0,004	0,473	0,469	0,004
	(0,005)	(0,008)	(0,011)	(0,009)	(0,013)	(0,022)
VWO	0,386	0,366	0,020	0,379	0,343	0,037
	(0,005)	(0,008)	(0,191)	(0,008)	(0,013)	(0,191)
SE grades	66,402	65,673	0,729	65,379	65,633	-0,254
	(0,054)	(0,081)	(0,717)	(0,095)	(0,152)	(0,734)
CE grades	64,465	63,638	0,827	65,036	64,657	0,379
	(0,068)	(0,106)	(0,510)	(0,110)	(0,186)	(0,885)
Native Dutch	0,870	0,889	-0,019	0,871	0,899	-0,028
	(0,003)	(0,004)	(0,028)	(0,006)	(0,007)	(0,023)
Non-western immigrant	0,061	0,056	0,005	0,067	0,049	0,018
	(0,002)	(0,003)	(0,019)	(0,004)	(0,005)	(0,017)
Socioeconomic status	0,129	-0,063	0,193	0,130	-0,109	0,239
	(0,008)	(0,013)	(0,191)	(0,015)	(0,022)	(0,214)
Secondary school characteristics						
Number of graduation candidates	94,657	98,024	-3,368	97,782	110,367	-12,585
	(0,341)	(0,420)	(10,131)	(0,620)	(0,513)	(10,661)
Graduation rate	0,903	0,902	0,001	0,886	0,881	0,004
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,018)
Urbanity	2,869	2,991	-0,122	2,747	3,091	-0,344
	(0,010)	(0,021)	(0,467)	(0,020)	(0,037)	(0,546)
Tertiary school characteristics						
Research university students	0,329	0,290	0,039	0,323	0,278	0,045
	(0,005)	(0,007)	(0,144)	(0,008)	(0,011)	(0,149)
Studies with academic dismissal policy	0,793	0,772	0,022	0,676	0,624	0,052
	(0,004)	(0,007)	(0,051)	(0,008)	(0,013)	(0,068)
Studies with fixed quota	0,131	0,137	-0,006	0,188	0,210	-0,022
	(0,003)	(0,006)	(0,022)	(0,007)	(0,011)	(0,036)
Studies with selective admission	0,049	0,050	-0,001	0,123	0,138	-0,015
standards	(0,002)	(0,004)	(0,008)	(0,006)	(0,009)	(0,032)
Number of schools	24	38		24	38	
Number of students	9901	18772		3306	6498	

Notes: This table is similar to table 2. Compared to the model in table 2, the upper 30% (instead of 20%) in the intensity growth distribution is defined as a major shock. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D9

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, WITH A WIDER RANGE OF STABLE INTENSITY (25% - 75% OF INTENSITY GROWTH DISTRIBUTION)

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,869	16,805	0,063	16,860	16,820	0,040
	(0,009)	(0,007)	(0,104)	(0,015)	(0,013)	(0,099)
Male	0,461	0,471	-0,010	0,457	0,473	-0,016
	(0,006)	(0,005)	(0,011)	(0,011)	(0,009)	(0,015)
VWO	0,510	0,438	0,072	0,495	0,428	0,066
	(0,006)	(0,005)	(0,173)	(0,011)	(0,008)	(0,172)
SE grades	66,714	66,196	0,518	65,675	65,611	0,064
	(0,070)	(0,052)	(0,649)	(0,121)	(0,091)	(0,529)
CE grades	64,839	63,889	0.950**	65,076	65,004	0,072
	(0,086)	(0,065)	(0,449)	(0,138)	(0,112)	(0,572)
Native Dutch	0,865	0,865	0,000	0,863	0,875	-0,012
	(0,004)	(0,003)	(0,020)	(0,007)	(0,005)	(0,021)
Non-western immigrant	0,065	0,071	-0,006	0,075	0,066	0,010
	(0,003)	(0,002)	(0,014)	(0,006)	(0,004)	(0,018)
Socioeconomic status	0,115	0,133	-0,018	0,138	0,074	0,064
	(0,010)	(0,008)	(0,147)	(0,019)	(0,015)	(0,182)
Secondary school characteristics						
Number of graduation candidates	98,550	95,294	3,255	103,576	101,407	2,169
	(0,463)	(0,231)	(10,590)	(0,807)	(0,400)	(11,861)
Graduation rate	0,914	0,898	0,016	0,889	0,884	0,005
	(0,001)	(0,001)	(0,012)	(0,001)	(0,001)	(0,013)
Urbanity	2,694	2,679	0,015	2,525	2,727	-0,202
	(0,014)	(0,009)	(0,323)	(0,028)	(0,016)	(0,415)
Tertiary school characteristics						
Research university students	0,434	0,372	0,062	0,422	0,376	0,047
	(0,006)	(0,005)	(0,146)	(0,011)	(0,008)	(0,149)
Studies with academic dismissal policy	0,762	0,775	-0,013	0,682	0,674	0,008
	(0,005)	(0,004)	(0,050)	(0,010)	(0,008)	(0,062)
Studies with fixed quota	0,145	0,149	-0,004	0,217	0,204	0,013
	(0,004)	(0,004)	(0,023)	(0,009)	(0,007)	(0,034)
Studies with selective admission standards	0,059	0,051	0,009	0,147	0,133	0,013
	(0,003)	(0,002)	(0,011)	(0,008)	(0,005)	(0,031)
Number of schools	15	47		15	47	
Number of students	6381	22528		2173	7694	

Notes: This table is similar to table 2. Compared to the model in table 2, the 2nd and 3rd quartiles (compared to the 4th, 5th, 6th and 7th quintiles) in treatment intensity growth distribution is defined as a stable intensity. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D10

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, WITH A WIDER RANGE OF STABLE INTENSITY (20% - 80% OF INTENSITY GROWTH DISTRIBUTION)

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,869	16,827	0,042	16,860	16,792	0,068
	(0,009)	(0,008)	(0,109)	(0,015)	(0,014)	(0,105)
Male	0,461	0,471	-0,010	0,457	0,470	-0,013
	(0,006)	(0,006)	(0,011)	(0,011)	(0,010)	(0,014)
VWO	0,510	0,488	0,022	0,495	0,477	0,018
	(0,006)	(0,006)	(0,194)	(0,011)	(0,010)	(0,194)
SE grades	66,714	66,345	0,369	65,675	66,078	-0,402
	(0,070)	(0,068)	(0,740)	(0,121)	(0,122)	(0,647)
CE grades	64,839	64,766	0,073	65,076	65,983	-0,907
	(0,086)	(0,085)	(0,728)	(0,138)	(0,146)	(0,792)
Native Dutch	0,865	0,862	0,003	0,863	0,876	-0,013
	(0,004)	(0,004)	(0,024)	(0,007)	(0,006)	(0,025)
Non-western immigrant	0,065	0,067	-0,002	0,075	0,063	0,012
	(0,003)	(0,003)	(0,014)	(0,006)	(0,005)	(0,018)
Socioeconomic status	0,115	0,119	-0,004	0,138	0,111	0,028
	(0,010)	(0,010)	(0,139)	(0,019)	(0,017)	(0,177)
Secondary school characteristics						
Number of graduation candidates	98,550	102,202	-3,652	103,576	107,550	-3,973
	(0,463)	(0,392)	(12,953)	(0,807)	(0,532)	(13,052)
Graduation rate	0,914	0,908	0,006	0,889	0,903	-0,013
	(0,001)	(0,001)	(0,012)	(0,001)	(0,001)	(0,014)
Urbanity	2,694	2,420	0,274	2,525	2,477	0,048
	(0,014)	(0,012)	(0,368)	(0,028)	(0,022)	(0,463)
Tertiary school characteristics						
Research university students	0,434	0,423	0,011	0,422	0,424	-0,001
	(0,006)	(0,006)	(0,165)	(0,011)	(0,010)	(0,170)
Studies with academic dismissal policy	0,762	0,736	0,026	0,682	0,675	0,007
	(0,005)	(0,005)	(0,060)	(0,010)	(0,009)	(0,063)
Studies with fixed quota	0,145	0,150	-0,006	0,217	0,213	0,004
	(0,004)	(0,004)	(0,023)	(0,009)	(0,008)	(0,033)
Studies with selective admission standards	0,059	0,055	0,005	0,147	0,143	0,004
	(0,003)	(0,003)	(0,011)	(0,008)	(0,007)	(0,031)
Number of schools	15	50		15	50	
Number of students	6381	24060		2173	8253	

Notes: This table is similar to table 2. Compared to the model in table 2, the 3rd, 4th, 5th, 6th, 7th , and 8th quintiles (compared to the 4th, 5th, 6th and 7th quintiles) in treatment intensity growth distribution is defined as a stable intensity. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D11

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, WITHOUT EXLUDING STUDENTS THAT DO NOT ENROLL IN TERTIARY EDUCATION

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,869	16,751	0,117	16,860	16,769	0,091
	(0,009)	(0,007)	(0,106)	(0,015)	(0,012)	(0,096)
Male	0,461	0,474	-0,013	0,457	0,467	-0,010
	(0,006)	(0,005)	(0,012)	(0,011)	(0,009)	(0,016)
VWO	0,510	0,397	0,113	0,495	0,388	0,107
	(0,006)	(0,005)	(0,178)	(0,011)	(0,009)	(0,178)
SE grades	66,714	65,855	0,859	65,675	65,459	0,216
	(0,070)	(0,054)	(0,649)	(0,121)	(0,102)	(0,665)
CE grades	64,839	63,921	0.919**	65,076	65,110	-0,034
	(0,086)	(0,070)	(0,460)	(0,138)	(0,118)	(0,627)
Native Dutch	0,865	0,871	-0,006	0,863	0,881	-0,018
	(0,004)	(0,003)	(0,025)	(0,007)	(0,005)	(0,025)
Non-western immigrant	0,065	0,071	-0,006	0,075	0,061	0,014
	(0,003)	(0,002)	(0,018)	(0,006)	(0,003)	(0,019)
Socioeconomic status	0,115	0,086	0,029	0,138	0,064	0,074
	(0,010)	(0,008)	(0,152)	(0,019)	(0,014)	(0,185)
Secondary school characteristics						
Number of graduation candidates	98,550	97,973	0,577	103,576	106,926	-3,350
	(0,463)	(0,282)	(11,195)	(0,807)	(0,484)	(12,451)
Graduation rate	0,914	0,897	0.017*	0,889	0,889	0,000
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,014)
Urbanity	2,694	2,708	-0,014	2,525	2,757	-0,232
	(0,014)	(0,012)	(0,371)	(0,028)	(0,021)	(0,451)
Tertiary school characteristics						
Research university students	0,434	0,328	0,106	0,422	0,327	0,095
	(0,006)	(0,005)	(0,142)	(0,011)	(0,008)	(0,146)
Studies with academic dismissal policy	0,762	0,767	-0,005	0,682	0,627	0,054
	(0,005)	(0,004)	(0,048)	(0,010)	(0,008)	(0,066)
Studies with fixed quota	0,145	0,135	0,009	0,217	0,209	0,008
	(0,004)	(0,004)	(0,023)	(0,009)	(0,007)	(0,033)
Studies with selective admission standards	0,059	0,048	0,011	0,147	0,139	0,008
	(0,003)	(0,002)	(0,011)	(0,008)	(0,006)	(0,030)
Number of schools	15	42		15	42	
Number of students	6381	20649		2173	7162	

Notes: This table is similar to table 2. Compared to the model in table 2, students that did not enroll in tertiary education at all are not excluded from the sample, but given a retention outcome of 0 – similar to students that switch or dropout. Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).

TABLE D12

DESCRIPTIVE STATISTICS OF STUDENT AND SCHOOL CHARACTERISTICS IN TREATMENT AND SYNTHETIC CONTROL GROUP IN THE PRE- AND POST-TREATMENT PERIOD, EXPLOITING SHOCKS IN INTENTITY OF USE, EXCLUDING POSSIBLE OUTLIER

	Pre-treatment period			Post-treatment period		
	(1)	(2)	(3)	(4)	(5)	(6)
	Treatment	Synthetic control	(1)-(2)	Treatment	Synthetic control	(4)-(5)
Individual characteristics						
Age	16,869	16,756	0,113	16,860	16,771	0,089
	(0,009)	(0,007)	(0,107)	(0,015)	(0,012)	(0,097)
Male	0,461	0,474	-0,013	0,457	0,466	-0,009
	(0,006)	(0,005)	(0,012)	(0,011)	(0,009)	(0,016)
VWO	0,510	0,406	0,104	0,495	0,397	0,098
	(0,006)	(0,005)	(0,179)	(0,011)	(0,009)	(0,180)
SE grades	66,714	65,900	0,814	65,675	65,506	0,169
	(0,070)	(0,056)	(0,651)	(0,121)	(0,105)	(0,677)
CE grades	64,839	63,935	0.905*	65,076	65,056	0,020
	(0,086)	(0,072)	(0,466)	(0,138)	(0,122)	(0,646)
Native Dutch	0,865	0,871	-0,007	0,863	0,881	-0,018
	(0,004)	(0,003)	(0,026)	(0,007)	(0,005)	(0,025)
Non-western immigrant	0,065	0,072	-0,007	0,075	0,061	0,014
	(0,003)	(0,002)	(0,018)	(0,006)	(0,003)	(0,020)
Socioeconomic status	0,115	0,070	0,045	0,138	0,050	0,088
	(0,010)	(0,008)	(0,154)	(0,019)	(0,015)	(0,188)
Secondary school characteristics						
Number of graduation candidates	98,550	97,292	1,258	103,576	106,523	-2,947
	(0,463)	(0,283)	(11,173)	(0,807)	(0,488)	(12,451)
Graduation rate	0,914	0,897	0.017*	0,889	0,889	0,000
	(0,001)	(0,001)	(0,010)	(0,001)	(0,001)	(0,015)
Urbanity	2,694	2,702	-0,008	2,525	2,757	-0,232
	(0,014)	(0,013)	(0,377)	(0,028)	(0,023)	(0,460)
Tertiary school characteristics						
Research university students	0,434	0,335	0,099	0,422	0,334	0,088
	(0,006)	(0,005)	(0,143)	(0,011)	(0,008)	(0,148)
Studies with academic dismissal policy	0,762	0,765	-0,003	0,682	0,632	0,049
	(0,005)	(0,004)	(0,049)	(0,010)	(0,009)	(0,066)
Studies with fixed quota	0,145	0,137	0,008	0,217	0,210	0,007
	(0,004)	(0,004)	(0,023)	(0,009)	(0,008)	(0,034)
Studies with selective admission standards	0,059	0,049	0,011	0,147	0,141	0,006
	(0,003)	(0,002)	(0,011)	(0,008)	(0,006)	(0,030)
Number of schools	15	41		15	41	
Number of students	6381	20108		2173	6963	

Notes: This table is similar to table 2. Compared to the model in table 2, the school with the highest post-treatment period RMSPE / pre-treatment period RMSPE – following from a placebo study - is excluded from the pool of control schools (see figure 8). Stars indicate the level of significance of the estimates (* p<0.10, ** p<0.05, *** p<0.01).