The influence of selection after the gate on graduation rates. Evidence from variation in the size of an academic-dismissal policy

Eline Sneyers & Kristof De Witte
The influence of selection after the gate on graduation rates.
Evidence from variation in the size of an academic-dismissal policy*

Eline Sneyers†† Kristof De Witte§§

October, 2014

Abstract

This paper examines the influence of selection after the gate on student graduation rates. We use the variation in academic dismissal (AD) policies as an instrument for the endogenous student dropout level. Using a parametric 2SLS estimation and a nonparametric IV regression, the instrumented variation in student dropout is used to carve out the impact on graduation rates. The IV outcomes on Dutch program level data suggest that: (i) student dropout increases with higher academic dismissal thresholds, and (ii) programs which are more selective in the first year have higher student graduation rates. The results also indicate that student satisfaction, the percentage of female students in the study program and achieved level of the study program are associated with higher graduation rates.

JEL-Classification: C36; I21

Keywords: Higher education; Academic dismissal policy; Student dropout; Completion rates; Instrumental Variable estimation.

---

*We would like to thank the participants from the 29th annual congress of the European Economic Association, 26e EALE conference, MEBIT 2012-14 Amsterdam, a seminar at the Dutch Ministry of Education, Kees Boele, Wim Groot and Henriëtte Maassen van den Brink for valuable insights and comments.

†Top institute for Evidence Based Education Research, Maastricht University, Kapoenstraat 2, 6200 MD Maastricht, the Netherlands, eline.sneyers@maastrichtuniversity.nl.

‡Corresponding author.

§Faculty of Economics and Business, Leuven Economics of Education Research, University of Leuven (KUL), Naamsestraat 69, 3000 Leuven, Belgium, kristof.dewitte@kuleuven.be; and Top institute for Evidence Based Education Research, Maastricht University, Kapoenstraat 2, 6200 MD Maastricht, the Netherlands, k.dewitte@maastrichtuniversity.nl.
1 Introduction

As higher education is considered as a necessary condition to stimulate employment opportunities, social justice and economic progress, a key priority in policy agendas is to increase attainment levels in higher education. For example, by 2020, 40 percent of the 30-34 years old in the European Union (EU) should have obtained a tertiary degree. To achieve this target, measures are taken to ensure greater access to and graduation of higher education for non-traditional and disadvantaged students. While this democratization of higher education increases the first-year enrollment rates, the target of 40 percent attainment is threatened by high dropout rates (NESET, 2013). On average, 30 percent of the students who enroll in higher education do not obtain a degree (OECD, 2013). Especially students from low social-economic background, ethnic minority students and students with disabilities have a high chance of dropping out. The high student dropout rates also have financial implications. Dropout can be seen as a drain on public finance and a waste of vulnerable resources (NESET, 2013). In this context, governments have introduced performance-based funding (PBF) mechanisms which links prescribed performance measures to funding. Not surprisingly, two central variables in PBF schemes are first-year student dropout and student graduation rates. The increasing popularity of PBF induced two trends in higher education.

First, higher education institutes have become more selective in accepting students. Selection of students falls apart in two practices. There is selection ‘at the gate’ which only accepts students with a high likelihood of obtaining a degree within the nominal study period (Beller, 2001; Salvatori, 2001). Alternatively, institutions can select ‘after the gate’. While all students are allowed to enroll in the first-year of study, there are some strict rules on who can continue after this first-year (Duijndam & Scheepers, 2009; Lindo et al., 2010). Examples of this screening tool include the use of first-year grade point averages (GPA) and first-year credits (ECTs).

Second, there is more focus on student graduation as a measure of institutional effectiveness (Fike & Fike, 2008). Students who do not succeed in obtaining a degree within the nominal study time are costly for both the individual and society. Indeed, students and parents benefit because of lower accumulated tuition fees (Johnstone, 2004). When students graduate within the normal study time, society benefits because of reduced subsidies and earlier (tax) revenues from labour market participation. Due to PBF mechanisms, universities and higher vocation institutions also benefit from graduation within the nominal study time as it leads to higher and sooner funding.

This paper contributes to the literature in four ways. First, it has been argued that higher student dropout is related to lower graduation rates (e.g., Hosch, 2008; Lau, 2003). How-

---

1We will use the term ‘success’, ‘graduation’ and ‘completion’ interchangeably throughout the paper.
ever, we question whether all types of dropout (e.g. due to bad performance, demotivation, dissatisfaction with the institution) have the same influence on graduation rates.

Second, there is an increased attention towards selection after the gate. The literature is silent concerning the relationship between graduation rates and dropout caused by selection after the gate. This paper examines whether dropout caused by selection after the gate leads to a higher graduation rate. If this is the case, we show that different types of dropout have a different influence on graduation rates and, hence, the PBF mechanism should not use dropout as an indicator.

As a third contribution, we use an instrumental variable (IV) approach to investigate the influence of dropout caused by selection after the gate on graduation rates. We cannot use Ordinary Least Squares (OLS) for two reasons. First, this relationship may be biased due to endogeneity issues, more precisely due to unobservable characteristics at the program level (e.g., program quality, student-staff ratio, program attractiveness). The unobserved heterogeneity can also arise from imprecise assignment to the treatment. Students can dropout during the year because they fear that they will not pass the selection after the gate criteria. Other students are forced to leave their study after first-year finals. The instrumental variable approach allows us to identify the influence of student dropout caused by selection after the gate on graduation rates. By using the size of an academic dismissal policy as an instrument, the variation in student dropout that is uncorrelated with the error term is carved out. Next, this exogenous variation in student dropout is used to obtain the impact on graduation rates. We apply both parametric 2SLS estimations, as well as nonparametric IV regressions. The latter are attractive as they do not rely on pre-specified functional forms.

Finally, the application uses a rich Dutch data set at academic degree-level (e.g. biology, applied linguistics). Dropout and graduation rates are interesting to study in the Netherlands as they are measured in a nation-wide standardized way. This limits measurement errors and results in an uniform data set across institutions and academic degrees. In addition, the data contain information on students’ well-being, student-staff ratios and age of teachers.

The remainder of the paper is structured as follows. In section 2, we present a literature review on student success, student dropout and student selection. In section 3 the Dutch setting is briefly described, while section 4 presents the data. Next, we outline the parametric identification strategy to estimate the impact of student dropout on student success. Thereafter, we present the results and nonparametric robustness tests. In section 7, the paper concludes with a discussion and policy implications.

2We define dropout as the percentage of students that dropout during the first-year. The graduation rate is defined as the number of students that graduate within one year after the nominal study time conditional on the first-year dropout.
2 Literature

Universities and higher vocational education institutions are increasingly held responsible for their students (Huisman & Currie, 2004). As a result, higher education institutions focus more on student graduation and student dropout. We start by discussing the observed relationship between student dropout and student graduation\(^3\), while subsection 2.2 overviews the literature on student selection 'after the gate'.

2.1 Student dropout and graduation rates

Figure 1 presents the relationship between first-year student dropout and graduation rates for all Dutch bachelor programs between the academic years 2006-07 and 2008-09. We do not observe a clear relationship between the two variables. Some academic programs combine low dropout rates with low graduation rates, while other programs exploit the selective nature of the first-year to achieve high graduation rates. The programs in the green square clearly struggle with obtaining high graduation rates, independently of whether they have high or low student dropout rates. We may not observe a clear relationship because we look at all dropout and because different kinds of dropout may have a different influence on graduation rates.

Earlier studies already indicated that high freshman attrition rates are strongly correlated with the level of student graduation rates (e.g., Hosch, 2008; Lau, 2003; Levitz et al., 1999). Forrest (1982) noted that there exist a wide variation of dropout rates within groups of similar institutions. This suggests that program and institutional characteristics matter for the dropout rates. Not surprisingly, an extensive body of research has emerged that identified the conditions in which students should be placed in order to increase first-to-second year student retention. Based on conceptual models by Tinto (1993) and Bean (1990), academic programs try to adapt various organizational factors in order to promote student retention. For example, some academic programs opt to increase the degree of selectivity 'at the gate'. Indeed, more selective admission standards have been related to higher persistence rates (Cox et al., 2005; Titus, 2004). Alternative interventions increase the student-staff ratio or restrict the number of students enrolled (see Sneyers and De Witte (2013a) for a summary of program and institutional characteristics that influence the ratio of student dropout-graduation rates).

The promotion of student retention might also have negative consequences. For example, as more first-year students are retained, some may not have a good match with their study program such that they linger in their academic program. This, in turn, will lead to student dropout after the first-year and/or to students who graduate after the nominal study time.

\(^3\)Note that the literature review focuses on the total dropout and not only the dropout caused by selection after the gate.
Figure 1: Relationship between first-year dropout and completion rates in study programs at Dutch universities (measured in 2007-08 and 2008-09) and at Dutch higher vocational institutions (measured in 2006-07 and 2007-08) (red square= high student dropout and high completion rates, black square= low student dropout and high completion rates; and green square= low/high student dropout and low completion rates); corr= -.19, p=.000
(Arnold & van den Brink, 2010; Stegers-Jager et al., 2011). It is clear that the relationship between first-year dropout and graduation rates is not unambiguous.

Having discussed the relationship between student dropout and student success, we turn in the next subsection to the literature on student selection mechanisms ‘after the gate’.

2.2 Student selection ‘after the gate’

Selection ‘after the gate’ refers to policies that remove students with unsatisfactory academic progress from the study program. This type of selection systems have a general structure. When a student’s performance falls below the minimum standard set by a study program, the student receives a warning. The warning serves as a wake-up call and can lead to more penalties. As the sophomore year is an important period for the connection between students and their academic programs, selection ‘after the gate’ should ideally happen in the first-year (Yorke & Longden, 2004). Although the introduction of performance standards can motivate some students to improve their performance, others can be discouraged from making an attempt at all (Bénabou and Tirole, 2003). Studies on the effect of negative incentives, such as selection ‘after the gate’, are limited (an exception is Sneyers and De Witte, 2013b).

In the US and Canada, higher education institutions apply a form of selection ‘after the gate’ called academic probation. In this system, students’ grade point average (GPA) is used as the minimum standard. Students are placed on academic probation if they do not reach the GPA norm. When their grades do not improve by the following evaluation, they are suspended from the study program or even from the institution (Lindo et al., 2010). The latter authors came to the conclusion that some students are discouraged from returning to school due to the academic probation. However, students who remained in school despite the academic probation improved their subsequent performance.

A comparable system is used in Dutch study programs (Van Heerden, 2013). This system, called an academic dismissal (AD) policy (in Dutch ‘Bindend Studie Advies’), requires students to make satisfactory progress during the first-year. While the threshold unit in the US and Canada consists of the GPA, the passing norms in the Netherlands are grounded on credits. Depending on the workload, every course in the study program is linked to a number of credits. When students pass the course, they receive the corresponding credits. If students earn a number of credits beneath the AD threshold, and thus make substandard progress in the first-year, they are dismissed. Important to note is that an AD policy also has an signaling function as students who fear that they will not be able to complete a particular study program, will not enter this program. So, the presence of an AD policy in itself leads to a first selection of students. In response to the introduction of an AD policy or to the strengthening of the AD policy threshold students might also react through their effort choice (i.e. learning function).
There exist a number of Dutch studies concerning the efficiency of an AD policy. Duijndam and Scheepers (2009) found that students in a business management program with an AD policy withdrew earlier than students in programs without an AD policy. They concluded that fewer students linger in the program because of the implementation of an AD threshold. By comparing two cohorts of students (one in which an AD policy was implemented and one in which an AD policy was implemented fictitiously), Gijbels et al. (2004) found that students who would have been dismissed, if their program applied a credit threshold, obtained lower numbers of credits in their second year compared to students in programs with an AD policy. Furthermore, a lower first-year graduation rate was found. The authors conclude that an AD policy is a good tool for the selection of well-performing students. Another study found that an increase in credit threshold results in increasing graduation rates (Task force studiesucces, 2009). In line with the above findings, Sneyers and de Witte (2013a) and Arnold (2014) found that the implementation of an AD policy leads to higher first-year dropout rates and higher graduation rates. Sneyers and de Witte (2013a) also proved that student satisfaction decreased due to the introduction of an AD policy. However, a study by Stegers-Jager et al. (2011) claimed that the introduction of an AD policy does not lead to higher graduation rates or early dropout during the first two years of enrollment at medical school. Arnold and Van de Brink (2010) showed that the introduction of an AD policy will not necessarily result in a reduction of the number of students lingering in scientific education. Indeed, students that are dismissed due to the credit threshold get too little help with the selection of a new program and, thus, will perform poorly in their next program. Finally, De Koning et al. (2014) found that the introduction of a credit threshold results in a positive change in study behavior. However, this implementation does not necessarily translate into a higher level of student achievement what my indicate that students try to reach the credit requirement instead of acquiring as much knowledge as possible.

3 The setting

3.1 Dutch higher education

Dutch higher education is organized as a binary system which consists of privately funded and governmental funded institutions. We focus on the latter group, which comprises of 18 universities and 39 institutions for higher vocational education. Given the different nature of the Open University, which focusses on distance learning, and of the four religious universities, which prepare students for religious duties (e.g., priest), we ignore them even though they are state funded institutions. Students of higher vocational education institutes are prepared for the practice of a profession. Students of universities are prepared for independent
scientific work in an academic or professional setting. While approximately two-thirds of the higher education students enroll in vocational education institutes, about one-third enroll in universities (Huisman, 2008). Since secondary education, students are tracked for continuing education and prepared for specific disciplines (De Koning et al., 2014). Students who hold a certificate of pre-university education or a first-year certificate of higher vocational education can apply for universities. Admission to higher vocational education institutes is open for students who hold a certificate of senior general education, secondary vocational education or pre-university education. In the paper at hand, we will focus on bachelor programs. In universities a bachelor program normally takes three years, while in vocational institutions it normally takes four years.

Due to the European Credit Transfer System, Dutch higher education institutes offer since 2002 programs of 60 study points per year. Credit points (ECTs) represent the study time or workload required to complete a course or program. The workload comprises of (i) the actual hours spend in class, and (ii) the hours spend on preparing for classes, exams and other assessments. As in most European countries, a credit point represents a workload of 28 hours. Therefore, a program year consists of 1680 hours work.

3.2 Academic dismissal policy

By the end of the first-year of enrollment, Dutch universities and higher vocational institutes are legally obliged to provide each student with an advice concerning the continuation of their studies (Law on Higher Education and Scientific Research, art. 7.8b). This law in mind, many institutions introduced an academic dismissal (AD) policy, i.e. an obligatory (binding) study advice. The number of credit points obtained by the student during the first-year determines whether the binding study advice is negative or positive. The credit-threshold that students need to pass is determined by each institution and can vary by study program (De Koning et al., 2014). The credit-norm is, on average, 41 ECTs at higher vocational education programs and 38 ECTs at academic programs (Inspectie van het onderwijs, 2010). Before students receive a negative AD advice, an academic warning is issued which serves as a wake-up call and gives students the possibility to improve their performance (Rijksoverheid, n.d.).

An AD policy mechanism works in three ways. There is a signaling function which will deter some students from enrolling in the program if they think they will not make the AD policy threshold. Next, an AD policy has a learning feature. Due to the AD policy threshold, students will understand the effort requirements better and will update their expectations about making the threshold. This will cause some students to dropout. Finally, there is a

\[\text{Note that the performance requirements of the AD policy can also be qualitative, such as passing essential courses (Mennen, 2013). However, we will not focus on this type of AD policy as it is difficult to measure and interpret this kind of performance requirements.}\]
dismissal function of an AD policy. This means that students who did not make the threshold are forced to leave. We will focus on the latter two functions of an AD policy.

Two objectives are pursued by introducing an AD policy. First, helping students to determine whether the programs fit them. Second, allowing study programs to select and continue with the most motivated and talented students after the first-year (De Koning et al., 2014).

Although non-obligatory study warnings are a tradition in Dutch higher education, the use of an AD policy increased only recently. Indeed, student guidance became increasingly important. Because the AD policy is in line with the orienting, selecting and referential function of the propaedic phase in higher education, higher education institutions introduced it more frequently (Onderwijsraad, 2008). In the academic year 2007-08, 43 percent of the universities applied such a system, against 98 percent in higher vocational institutes. During the academic year 2007-08, approximately 17 percent of the first-year students in vocational education were obliged to terminate their study program due to a negative AD policy. In academic education, 9 percent of the first-year students received a negative AD advice (Inspectie van het onderwijs, 2010).

4 Data

The data are obtained from the Dutch Ministry of Education (‘Dienst Uitvoering Onderwijs’, DUO). The data contain information at academic degree-level (also denoted by ‘program-level’) concerning the number of first-year students, and the dropout rate for the student cohorts of 2010-11 and 2011-12. It also includes information on the field of study (e.g. law, sciences, engineering and economics) for a longer period: the academic years 2003-04 and 2011-12. Finally, the data have program-level information on the number of students who re-enroll after the first-year, the graduation rates of university programs for the student cohorts 2007-08 and 2008-09, and the graduation rates of higher vocational education programs for the student cohorts 2006-07 and 2007-08. All data concern bachelor students.

Despite its richness, the data have two limitations. First, we lack academic degree level data regarding the number of first-year students and the dropout rate from 2003-04 until 2009-10. This is necessary to compare the same student cohorts. This issue can be solved by combining the data on field of study level of the years 2003-04 until 2009-10 with the program level data of 2010-11 and 2011-12. It requires the assumption that the trend of the indicators at program-level is similar to the trend of the indicators at field of study-level. Because sector organizations use different definitions to measure the above variables, it is difficult to check whether this assumption holds. However, we observe that the final data are fairly similar to reports by other sector organizations (Vereniging Hogescholen, 2014). As a
second limitation, we cannot distinguish between study programs at different locations of an institution. Therefore, programs that did not implement an AD policy simultaneously at all locations are excluded. Specifically, .17 percent of the sample has been removed because of this reason.

The output variables of interest are defined as follows. Student dropout is defined as the percentage of full-time first-year bachelor students that leave the institution. Graduation rates are defined as the share of re-enrolled full-time bachelor students that have graduated at the institution one year after the nominal study time. This indicator is thus measured as the graduation rate conditional on the first-year student dropout (i.e., post-propaedeutic graduation rates). Consequently in order to be included in this latter indicator, students enrolled in vocational education need to graduate within five years, while students enrolled in universities need to graduate after four years.

The data are further enriched by program level data obtained from the annual questionnaire on student satisfaction (‘nationale studenten enquête’, NSE) which is carried out by the organization Studiekeuze123. The NSE data include for each academic degree the students’ opinions about the program they are following. Since 1991, the NSE is executed annually, although a break in the survey questions limits our data set to academic year 2009-10. The NSE data also include institutional and program information. First, the AD threshold represents the threshold of the academic dismissal policy implemented by each academic degree. It can be considered as a proxy of selectivity at program-level. Second, information on the observed quality of a program is obtained by the variable Achieved level of the program. It is based on the accreditations of the Dutch-Flemish Accreditation Organization (NVAO) and is measured as a variable between 0 (insufficient) and 3 (excellent). The indicator is composed by three underlying scores: orientation of the program, level of the program, and domain specific requirements. The latter represents the extent to which the program fulfills the requirements for similar (foreign) programs. Level of the program shows the degree to which the achieved qualifications are in accordance with general and internationally accepted descriptions of qualifications of a Bachelor. Orientation reflects (I) whether bachelor graduates have obtained the qualifications to allow access to the labour market, and (ii) whether the achieved qualifications are distracted from the requirements of international scientific practice, scientific disciplines and relevant practice in the occupational field. In case of a study program at university level, Orientation also reflects whether the obtained degree allows admission to at least one subsequent university course at the Master’s level. Third, we have information on the opinions of students about the program they are taking (i.e., average student satisfaction). This 10-point Likert scale variable ranges from (1) very dissatisfied to (10) very satisfied. Finally, we include institutional characteristics such as the number of students as a proxy of institutional size, percentage of staff older than 50 years,
and student-staff ratio as a proxy for education scale.

The final sample includes all bachelor programs in universities between the academic years 2007-09, and all bachelor programs in higher vocational institutions between the academic years 2006-08. The sample is restricted to these years since the graduation rate is only known for these cohorts of students. Programs with graduation rates of 0 or 100 percent are removed from the data as they are atypical rates that are probably linked to programs that have just started or that are going to be terminated. This excluded 67 programs, or 1.86 percent of the sample. The final sample includes 2,003 higher education programs spread across 13 universities (i.e. an average of 249 study programs per year) and 38 higher vocational education institutions (i.e. an average of 753 study programs per year)\(^5\).

Some descriptive statistics at program-level are presented in Table 1. An average bachelor program has a first-year dropout rate of 26.6 percent. There are, however, large differences between the worst performing program (97.2%) and the best performing program (0.5% dropout). It should be noted that our results are robust for eliminating those extreme observations from the sample. Academic programs have, on average, a graduation rate of 62.7%. While some programs achieve a high student graduation rate (i.e. 96.4%), others deal with a low graduation rate (i.e. 8.3%). The average AD policy threshold that academic programs install is 34.9 ECTs\(^6\). Further, we observe that the average program achieves a quality level of 1 (i.e. sufficient). The descriptive statistics also show that the first-year student population of academic programs consist, on average, of 45.8% female students. Academic programs receive on average a score of 6.8 regarding average student satisfaction. This indicates that students are, on average, satisfied with the program they are following. Finally, we observe the average program is part of an institution with 21,462 students, with a third of the employees older than 50 years, and a student-staff ratio of 12.1.

5 Identification strategy

5.1 The instrumental variables estimation

Ordinary Least Squares (OLS) estimates of an independent variable \(X\) on a dependent variable \(Y\) will be biased in case of endogeneity. Endogeneity arises when a regressor is correlated

---

\(^5\)In comparison to Section 3.1 we lose five universities (i.e. the Open University and the four religious universities) and one higher vocational institution (i.e. Design Academy Eindhoven). The religious universities are small and uni-sectoral institutions. This latter exclusion is due to a mismatch between the NSE and DUO dataset.

\(^6\)Note that the average AD policy threshold differs from the one mentioned in Section 3.2. This is the result of a different sample of observations. Indeed, the dataset of the study of Inspectie van het Onderwijs (2010) contained less study programs at university level, but more study programs at higher vocational level.
Table 1: Descriptive statistics at academic degree level

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student dropout</td>
<td>2,003</td>
<td>26.58</td>
<td>11.13</td>
<td>.50</td>
<td>97.20</td>
</tr>
<tr>
<td>Graduation rates</td>
<td>2,003</td>
<td>62.68</td>
<td>16.00</td>
<td>8.30</td>
<td>96.40</td>
</tr>
<tr>
<td>Program characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic dismissal policy</td>
<td>2,003</td>
<td>34.89</td>
<td>13.77</td>
<td>.00</td>
<td>60.00</td>
</tr>
<tr>
<td>Achieved level of the program</td>
<td>1,022</td>
<td>1.01</td>
<td>.70</td>
<td>.00</td>
<td>3.00</td>
</tr>
<tr>
<td>First-year female students (%)</td>
<td>1,943</td>
<td>45.76</td>
<td>28.84</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Average student satisfaction</td>
<td>2,003</td>
<td>6.80</td>
<td>.33</td>
<td>5.52</td>
<td>8.11</td>
</tr>
<tr>
<td>Institutional characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>2,003</td>
<td>21,462.17</td>
<td>10,556.39</td>
<td>313.00</td>
<td>38,551.00</td>
</tr>
<tr>
<td>Staff &gt; 50 years (%)</td>
<td>2,003</td>
<td>36.88</td>
<td>6.29</td>
<td>.00</td>
<td>56.04</td>
</tr>
<tr>
<td>Student-staff ratio</td>
<td>2,003</td>
<td>12.11</td>
<td>4.14</td>
<td>2.06</td>
<td>17.69</td>
</tr>
</tbody>
</table>
with the error term. This violates the zero conditional mean condition (Wooldridge, 2006). An instrumental variables (IV) estimation is a common approach to adequately address endogeneity issues (e.g., Bound et al., 1995). More specifically, IV focuses on the variations in $X$ that are uncorrelated with the error term and excludes the variations in $X$ that create the bias in the OLS coefficients. The use of parametric IV, however, relies on some assumptions: (i) Monotonicity; (ii) Existence of compliers; (iii) Unconfounded type; (iv) Mean exclusion restriction; and (v) Common support (see Frölich, 2007). As a robustness test we will later relax those assumptions in nonparametric IV estimations.

Consider the following 'structural equation', or the regression of interest, with a single endogenous variable $X$:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{li} + \ldots + \beta_{1+r} W_{ri} + u_i$$  \hspace{1cm} (1)

where $Y_i$ is the $i$th observation of the dependent variable; $X_i$ is the $i$th observation of the endogenous explanatory variable; $W_{li}, ..., W_{ri}$ are the $i$th observation of each of the control variables (covariates); and $u_i$ is the i.i.d. error term. Assume that the above equation is biased because the endogenous regressor $X_i$ is correlated with the error term $u_i$. In order to avoid a biased estimation, IV relies on 'instruments' (labelled as $Z_i$) (Basle, 2008). Instrumental variables need to be correlated with the troublesome endogenous regressor ($cov(X, Z \neq 0)$) and uncorrelated with the error term in the structural equation ($cov(Z, u = 0)$) (Larcker & Rusticus, 2010; Murray, 2006).

While there exist various procedures to estimate IV regressions, we apply the most common IV estimator: the two-stage least squares estimation (2SLS) (Hahn et al., 2004; Murray, 2006). In a first stage, the regressor is regressed on the instrument(s) and the covariate(s). This stage isolates the variation in $X_i$ that is not correlated with $u_i$. Specifically, the first stage regression is estimated as:

$$X_i = \pi_0 + \pi_1 Z_{li} + \ldots + \pi_m Z_{mi} + \pi_{m+1} W_{li} + \ldots + \pi_{m+r} W_{ri} + v_i$$  \hspace{1cm} (2)

Next, in the second stage, $Y_i$ is regressed on the fitted values of $X$, $\hat{X}_1, ..., \hat{X}$, and on the covariates. The second stage regression is estimated as:

$$Y_i = \beta_0 + \beta_1 \hat{X}_i + \beta_2 W_{li} + \ldots + \beta_{1+r} W_{ri} + u_i$$  \hspace{1cm} (3)

where the resulting coefficients of $\beta_0, \beta_1, ..., \beta_{1+r}$ are the 2SLS estimators which measure the causal impact of $X$ on $Y$.

5.2 Choice of the instrument

The first-year dropout rate is clearly an endogenous variable for student graduation rates as they are both influenced by unobserved variables (e.g., quality of the students, student
motivation, teacher opinions). Additionally, because of data constraints (i.e. we do not know which dropout is triggered by the AD policy) we use an IV approach.

An adequate instrument should be correlated to student dropout, while not (or only via dropout) to graduation rates. This is the case for the AD threshold. Due to the introduction of an AD policy students need to obtain a certain amount of credits. Students that do not reach the threshold after the first-year are dismissed. Previous research (e.g., Duijndam & Scheepers, 2009) already showed that due to the introduction of an AD policy fewer students linger in the program. Not surprisingly, the AD policy threshold is strongly correlated with student dropout (corr = .29; p = .000). Furthermore, graduation rates are measured as the share of re-enrolled full-time bachelor students that graduate at the institution one year after the nominal study time. As the AD policy is implemented in the first-year, it cannot have a direct influence on the student graduation rates. The AD policy threshold is therefore not directly correlated with student graduation (although an indirect impact via first-year dropout is possible). Section 6.2.1 tests in a more quantitative way the adequacy of the instrument.

To be an adequate instrument, the AD policy must be set by the management in a rather exogenous way. Annual reports and proposals for the performance agreements between the government and the higher education institutions (www.rcho.nl) show that the introduction and strengthening of an AD policy does not depend on certain institutional or program characteristics. Moreover, it is indicated that programs set an AD policy to: (I) motivate students from the start of the study program, (II) improve the match between student and study program; (III) carefully select the best students, and (IV) improve the quality of the program (e.g., Hanzehogeschool Groningen, n.d.; Hogeschool Rotterdam, 2012; Technische Universiteit Delf, 2010; Universiteit Leiden, 2009). The reports also indicate that some higher education institutions choose to set a very high AD policy threshold (i.e. above 50 ECTS) in order to raise the bar for the students (Avans Hogeschool, 2012). Our believe of an exogenous instrument is strengthened by the observation of Sneyers and De Witte (2013b) that one year before the introduction of an AD policy the dropout rate was decreasing while the graduation rate was increasing.

If the instrument, the AD threshold, fulfills the two critical conditions of the IV estimation (see section 5.1), we can estimate the unbiased local average treatment effect (LATE) (Angrist et al., 1996). More specifically, the LATE is the mean effect on student completion of a change in student dropout for the subpopulation of compliers. The compliers are all study programs whose student dropout would change due to a change in AD policy. Thus, the above practice allows us to observe the influence of dropout caused by selection after the gate on graduation rates.
5.3 Model specifications

We estimate various model specifications which include an increasing number of program and institutional characteristics. A first model specification only includes the instrument: AD threshold. A second model specification includes also year and field of study fixed-effects. The latter capture the clustering at the field of study level which could arise, e.g., from the different nature of the curriculum. Specifically, academic programs can be divided into 11 broad fields of study (e.g., economics and management, health sciences, and education). The former captures the cohort specific variation. A third model specification augments the earlier specifications with the observed quality of the program, the percentage of female students, the average student satisfaction, the number of students in the institution, the percentage of staff older than 50 years and the student-staff ratio. A final model specification replaces the field of study fixed effects for clusters at study code. By clustering at the level of the study code, we allow observations to correlate within groups of programs with the same study code (e.g., Small business and retail management, Fiscal Economics, Chemical, Accountancy) while they need to be independent between groups of programs with the same study code. We observe 232 groups of programs with the same study code in our dataset.

6 Results

6.1 "Naive" OLS regression

Table 2 presents the results of the "naive" OLS regressions. In the first column we regress the total first-year student dropout on graduation rates. In the next two columns year and field of study fixed effects, as well as program and institutional characteristics are added. In the last columns, we remove the field of study fixed effects and cluster the observations at the level of the study code. We observe a negative significant relationship between first-year student dropout and graduation rates in all models. When the dropout rate of a study program increases with one percentage point, the graduation rate declines with .268 percentage point (as indicated by model 1). This is in line with previous studies that showed that high freshman attrition rates go hand in hand with a high percentage of students graduating in the nominal study time (e.g., Hosch, 2008; Lau, 2003; Levitz et al., 1999). Although this influence decreases when additional control variables are added (i.e. Model 2, Model 3 and Model 4), the significance of the estimate prevails.

Important to mention that these analyses are based on all observed dropout and not only the dropout caused by the AD policy. Some students drop out because of demotivation or

---

The study code of a program allows us to identify similar academic degrees at different locations or even at different institutions.
dissatisfaction with the institution, other because of the AD threshold. Also, as earlier noted, the OLS estimates are probably biased due to endogeneity issues. Indeed, student dropout is strongly influenced by observed and unobserved heterogeneity.

6.2 IV: First stage regressions

6.2.1 First stage tests

Table 3 summarizes the statistical tests to assess the reliability and the efficiency of the IV estimation. First, we use the Durbin-Watson-Hausman test. This test evaluates the significance of the IV estimator versus the significance of the OLS estimator. The null hypothesis is that the OLS estimator yields consistent results. The null hypothesis is rejected for all models. This indicates that student dropout is an endogenous regressor, and the use of IV is preferred.

Next, the under-identification test checks whether the equation is identified, i.e. whether the excluded instrument is relevant and thus correlated with the endogenous regressor. The null hypothesis of the Kleibergen-Paap rk LM statistic is that the equation is under-identified. For all models, the null hypothesis is rejected and, thus, all models are identified.

When instruments are weakly correlated to the endogenous regressor, estimators can perform poorly. We use the Kleibergen-Paap F statistic to determine whether the AD policy threshold is strongly correlated with student dropout. The Kleibergen-Paap F statistic generates a statistic to the case of non-i.i.d. errors allowing for heteroskedasticity, autocorrelation and/or cluster robust statistics. Although no critical values are generated for the Kleibergen-Paap F statistic, it is custom to use the Stock-Yogo critical values tabulated for the Cragg-Donald F statistic (Baum et al., 2007). For comparison, we also provide the Cragg-Donald F statistic in table 3. We observe that there is no problem with weak instruments for two reasons: 1) the first-stage F statistic exceeds ten (i.e. the older rule of thumb proposed by Staiger & Stock, 1997); and 2) the Kleibergen-Paap F statistic are for all models well above the highest critical value of 16.38.

Finally, the overidentification of the instruments can be tested with the Hansen J statistic. However, these tests can only be executed when the number of instruments exceeds the number of endogenous regressors. This is not the case in our models since the equations are always exactly identified.

---

8 We report standard errors and statistics that are robust to the presence of arbitrary heteroskedasticity.
9 The critical values are obtained from Stock and Yogo (2005). We also make the remark that the Kleibergen-Paap Wald statistic is the robust counterparts of the Cragg-Donald Wald statistic.
10 Note that in the case of a single endogenous regressor, the first-stage F statistic is equal to the Kleibergen-Paap F statistic.
Table 2: Estimation of student dropout on graduation rates by 'naive' OLS regression

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
<th>Coef.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>69.815***</td>
<td>.910</td>
<td>63.781***</td>
<td>1.412</td>
<td>41.696***</td>
<td>13.352</td>
<td>40.201**</td>
<td>19.193</td>
</tr>
<tr>
<td>Student dropout</td>
<td>-.268***</td>
<td>.032</td>
<td>-.203***</td>
<td>.031</td>
<td>-.094**</td>
<td>.044</td>
<td>-.127**</td>
<td>.055</td>
</tr>
<tr>
<td>Achieved level of the program</td>
<td>-.294</td>
<td>.701</td>
<td>.184</td>
<td>.873</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-year female students (%)</td>
<td>.145***</td>
<td>.025</td>
<td>.149***</td>
<td>.0288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average student satisfaction</td>
<td>1.760</td>
<td>1.714</td>
<td>.632</td>
<td>2.472</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>-.0002***</td>
<td>.000</td>
<td>-.0002**</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff &gt; 50 years (%)</td>
<td>.353***</td>
<td>.085</td>
<td>.413***</td>
<td>.119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-staff ratio</td>
<td>.733***</td>
<td>.130</td>
<td>.703***</td>
<td>.168</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>YES</td>
<td></td>
<td>YES</td>
<td></td>
<td>YES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector fixed-effects</td>
<td>YES</td>
<td></td>
<td>YES</td>
<td></td>
<td>NO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,003</td>
<td>2,003</td>
<td>1,004</td>
<td>1,004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.0349</td>
<td>.1868</td>
<td>.2202</td>
<td>.1547</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>72.32</td>
<td>35.14</td>
<td>14.62</td>
<td>13.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of clusters</td>
<td>186</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where *** , ** , * denote significance at 1%, 5% and 10%-level; clusters are at study code, which allow us to identify similar academic degrees at different locations or even at different institutions.
Table 3: First stage tests

<table>
<thead>
<tr>
<th>Endogeneity test</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durbin-Wu-Hausman (chi squared) test</td>
<td>106.87***</td>
<td>106.95***</td>
<td>93.70***</td>
<td>49.60***</td>
</tr>
<tr>
<td>Under-identification test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kleibergen-Paap rk LM statistic</td>
<td>142.31***</td>
<td>76.05***</td>
<td>45.68***</td>
<td>26.83***</td>
</tr>
<tr>
<td>Weak identification test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cragg-Donald Wald F statistic</td>
<td>189.24</td>
<td>84.83</td>
<td>50.44</td>
<td>55.38</td>
</tr>
<tr>
<td>Kleibergen-Paap Wald rk F statistic</td>
<td>248.97</td>
<td>97.18</td>
<td>51.54</td>
<td>47.89</td>
</tr>
<tr>
<td>Overidentification test</td>
<td>exactly</td>
<td>exactly</td>
<td>exactly</td>
<td>exactly</td>
</tr>
<tr>
<td>Hansen J statistic</td>
<td>identified</td>
<td>identified</td>
<td>identified</td>
<td>identified</td>
</tr>
</tbody>
</table>

(1) where ***, **, * denote significance at 1%, 5% and 10%-level

(2) Stock-Yogo weak ID test critical values: 10% = 16.38; 15% = 8.96 ; 20% = 6.66; 25% = 5.53

6.2.2 First-stage results

Table 4 presents the results of the first-stage regression. We observe a positive significant relationship between the AD threshold and student dropout in all model specifications. This indicates that with increasing requirements, and thus a more selective first-year, more students fail to make satisfactory progress, and drop out of the program. In Model 2, where year- and field of study-fixed effects are added, we observe that dropout rates differ by year and by field of study.

In Model 3 additional control variables are included. We only discuss the covariates that have a significant influence on student dropout. We observe a significant negative relationship between the achieved level of the program and the student dropout. Assuming that programs that achieve a higher level are also of a higher quality, this finding is in line with studies that mention a positive relationship between quality and student retention (e.g. Peterson et al., 1997). The model indicates a significant negative correlation between student dropout and student satisfaction. Students who do not feel satisfied by their academic program of choice, due to a mismatch between institutional performance and student expectations, have a high risk of dropping out or transferring to another program. Previous studies already showed that student satisfaction has a positive influence on student retention (e.g., Schertzer & Schertzer, 2004). The percentage of staff older than 50 years and the student-staff ratio both have a significant positive correlation. The result regarding student-staff ratio is in line with previous research (e.g., Tinto, 2002). The finding concerning staff older than 50 years (i.e. age can be seen as an indicator of staff quality) is inconsistent with a study by Hoffmann and Oreopoulos (2009) who found that perceived professor quality is negatively related to the likelihood of dropping a course. However, it is important to note that literature about the connection between staff quality and student outcomes at the post-secondary level
Table 4: First-stage regression results - relationship between AD policy threshold and student dropout

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>18.293***</td>
<td>.544</td>
<td>27.273***</td>
<td>1.086</td>
</tr>
<tr>
<td>AD policy threshold</td>
<td>.238***</td>
<td>.015</td>
<td>.168***</td>
<td>.017</td>
</tr>
<tr>
<td>Control Variables</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Field of study fixed-effects</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,003</td>
<td>2,003</td>
<td>1,004</td>
<td>1,004</td>
</tr>
<tr>
<td>Number of clusters</td>
<td></td>
<td></td>
<td></td>
<td>186</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.0864</td>
<td>.1730</td>
<td>.2017</td>
<td>.1823</td>
</tr>
</tbody>
</table>

where ***, **, * denote significance at 1%, 5% and 10%-level; clusters are at study code, which allow us to identify similar academic degrees at different locations or even at different institutions.
is virtually non-existent. Furthermore, some recent evidence at elementary and secondary education levels indicated that teacher-induced learning has low persistence over time (Jacob et al., 2008).

In a fourth and final model, we cluster the results at the level of the study code. This prevents us from including field of study specific effects. Compared to Model 3, we observe that the significance and direction of the effects of the achieved level of the program and average student satisfaction remained. The significance of the variables staff > 50 years and student-staff ratio vanishes. This may due to the fact that we cluster at study code-level, while the latter two variables are at institutional level.

6.3 IV: Second stage regressions

The second stage results are presented in Table 5. Several interesting observations can be made. First, student dropout caused by selection after the gate has a significant positive impact on student graduation rates across all models. This indicates that programs with high first-year student dropout caused by selection after the gate have also higher student graduation rates. The coefficient reveals that the observed influence on student graduation is large. Indeed, as indicated by Model 1, a 1% increase in first-year student dropout caused by the AD policy leads to an increase in student graduation of one percent. By adding control variables the observed impact further increases. Although the relative increase in the influence of student dropout caused by selection after the gate on student graduation is most distinct in the transition from Model 1 to Model 2, the majority of the year and field of study effects disappears once the other covariates are added in Model 3 and 4. This shows that program and institutional characteristics have a large influence on students' performance. Finally, we observe that the results of the OLS estimation (see section 6.1) and the IV estimation differ considerably. Most importantly, the influence of student dropout is significantly negative in the OLS estimation, while in the IV estimation this impact is significantly positive (see Figure 2). We also observe that the influence, in comparison to the OLS estimation, is much larger. We can conclude that our concerns regarding the type of student dropout and the endogeneity issues are justified.

As indicated by the results of Model 3, three program characteristics significantly correlate to graduation rates. The level of the program has a positive influence on student graduation. This is in line with research that assumes that study programs with high graduation rates are considered to provide higher quality than study programs with low graduation rates (OCW, 2011a, b; Kokkelenberg et al., 2008). We further observe a positive relationship between graduation rates and the percentage of first-year female students. This is consistent with previous research (e.g., Porter, 2006; Scott et al., 2006). These gender differences in degree performance may, for example, be due to variations (i) in psychological and/or bio-
Table 5: Two-stage least squares results - relationship between student dropout and student graduation rates (estimated with a continuous instrument)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>35.946***</td>
<td>3.632</td>
<td>-.855</td>
<td>8.776</td>
</tr>
<tr>
<td>Student dropout</td>
<td>1.006***</td>
<td>.137</td>
<td>1.687***</td>
<td>.250</td>
</tr>
<tr>
<td>Control variables</td>
<td>NO</td>
<td></td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>YES</td>
<td></td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Field of study fixed-effects</td>
<td>YES</td>
<td></td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,003</td>
<td></td>
<td>2,002</td>
<td></td>
</tr>
<tr>
<td>Number of clusters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where ***, **, * denote significance at 1%, 5% and 10%-level; clusters are at study code, which allow us to identify similar academic degrees at different locations or even at different institutions.
logical factors, and/or (ii) in characteristics that are correlated with attainment (e.g. family background) (McNabb, 2002; Mellonby et al., 2000). Student satisfaction also plays a major role in the performance of students. This may be due to a sense of belonging to and loyalty to the institution that derives from student satisfaction (Tinto, 1993). Previous research already showed that student satisfaction is positively correlated with academic performance (Bean and Vesper, 1994). The number of students enrolled in the institution has a significant negative impact on student graduation. This finding is in line with earlier studies (e.g., Calcagno et al., 2008) and may due to the fact that the amount of students increases faster than the number of facilities offered to them. As a result, academic and social support suffers (Chickering and Reisser, 1993).

In Model 4, we cluster programs at study code-level. Not surprisingly, the significant result for the variable 'students enrolled' vanishes. However, all control variables on program-level remain significant.

Figure 2: Relationship between first-year dropout and completion rates in study programs at Dutch universities and at Dutch higher vocational institutions. The figure shows the fitted values of student dropout (estimated by regressing the AD policy threshold and the control variables of Model 4 on student dropout) and the fitted values of the completion rates (estimated by regressing the fitted values of student dropout and the control variables of model 4 on the completion rates); corr=.63, p=.000
6.4 Robustness tests

6.4.1 An alternative instrument

To test the robustness of the results, we reproduce the above findings by using an ordinal instrument. While the instrument above was continuous with values between 0 and 60, we now make an ordinal instrument which varies between 0 and 2. The instrument obtains a value of 0 for programs without or with a low AD threshold (i.e., between 0 and 19). Programs with an average threshold (i.e., between 20 and 39) are assigned a value of 1. Programs which have high AD thresholds, obtain a value of 2.\(^{11}\)

The 2SLS results with this alternative specification of the instrument are provided in Table 6.\(^{12}\) Results are in general very robust. In model 1, we observe that there exist a significant positive relationship between student dropout caused by an AD policy and student graduation. Consequently, the more students drop out in the first-year due to an AD policy, the higher the graduation rates for the remaining group. Moreover, we observe that when the student dropout rate increases with one percent due to selection after the gate, the student graduation rate increases with 1.44 percent. Again, adding control variables increases this impact. In Model 2, we see that year and field of study have a significant influence on student success. But once we add the program specific and institutional specific covariates, the majority of their significance disappears. Similar to section 6.3, some institutional and program characteristics have a significant influence on student success: (i) the number of students enrolled, (ii) percentage of first-year female students, (iii) student satisfaction, and (iv) level of the program. The direction of the effects is also similar to before although the significant effect of the number of students enrolled does not disappear once we cluster on study code-level (Model 4). Note that we observe for all models a similar size of the influence of student dropout caused by an AD policy on graduation rates in comparison to earlier results (see Table 5).

6.4.2 Non-parametric IV

The parametric IV methods rely on some strong assumptions on the specification of the functional form (see section 5.1). In order to relax these assumptions, we estimate a nonparametric IV specification as suggested by Frölich (2007). It has been demonstrated that this estimator is efficient and asymptotically normal. For brevity, we skip further discussions concerning the specifications of this estimator and, instead, refer to Frölich (2007). Note that in order to use this non-parametric IV estimator, we need to transform the original instrument into a binary variable. Programs with and without an AD policy are assigned a different

\(^{11}\)It should be noted that a dummy specification of the instrument (i.e., 0 and 1 for programs without and with an AD policy, respectively) yields similar results.

\(^{12}\)In order to save some space, the first stage regressions are not presented. The first stage tests and results are comparable to the findings in section 6.2
Table 6: Two-stage least squares results - relationship between student dropout and student success (estimated with an ordinal instrument)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
<td>SE</td>
</tr>
<tr>
<td>Intercept</td>
<td>24.447***</td>
<td>5.419</td>
<td>-16.810</td>
<td>13.172</td>
</tr>
<tr>
<td>Student dropout</td>
<td>1.438***</td>
<td>.204</td>
<td>2.153***</td>
<td>.378</td>
</tr>
<tr>
<td>Achieved level of the program</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-year female students (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average student satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>-.0002*</td>
<td>.0001</td>
<td>-.0002*</td>
<td>.0001</td>
</tr>
<tr>
<td>Staff &gt; 50 years (%)</td>
<td>.016</td>
<td>.161</td>
<td>.194</td>
<td>.168</td>
</tr>
<tr>
<td>Student-staff ratio</td>
<td>.153</td>
<td>.248</td>
<td>.223</td>
<td>.298</td>
</tr>
<tr>
<td>Year fixed-effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>field of study fixed-effects</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2,003</td>
<td>2,003</td>
<td>1,003</td>
<td>1,004</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>186</td>
<td>186</td>
<td>186</td>
<td>186</td>
</tr>
</tbody>
</table>

where *** , ** , * denote significance at 1% , 5% and 10%-level
value (i.e., no AD policy=0; AD policy=1).

The results of the non-parametric IV estimation are presented in Table 7. The results are comparable to the results observed in Section 6.3. In Model 1, we observe a positive significant impact of student dropout caused by an AD policy on graduation rates. Moreover, an increase of student dropout by one percent due to selection after the gate leads to an increase in graduation rates of about one percent. Adding covariates increases the influence of student dropout on student graduation. Model 2 shows an 1.3 percent increase in graduation rates due to a one percent increase in dropout. Model 3 even suggests an increase in student graduation of 1.9%.

Again the size of the impact of student dropout caused by selection after the gate on student graduation rates of all models is in line with earlier results (see Table 5 and Table 6).

Table 7: Non-parametric IV (estimated with a binary instrument)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>Coef.</td>
</tr>
<tr>
<td>Student dropout</td>
<td>.966***</td>
<td>.062</td>
<td>1.307***</td>
</tr>
<tr>
<td>N</td>
<td>2003</td>
<td></td>
<td>2003</td>
</tr>
</tbody>
</table>

where *** , ** , * denote significance at 1%, 5% and 10%-level

7 Discussion and policy implications

Higher education institutions have an increased attention towards student graduation. Student graduation can be increased by exploiting the selective nature of the first-year, for example by implementing an academic dismissal (AD) policy, which is a form of selection after the gate. Previous research argued that there might exist a positive relationship between student retention and student graduation (e.g., Lau, 2003). However, this relationship may be biased due to unobserved characteristics. This paper examines the relationship between student dropout caused by selection after the gate and graduation rates in Dutch academic programs. We use an instrumental variables (IV) estimation. By using the AD policy as an instrument, the variation in student dropout that is uncorrelated with the error term is carved out and, hence, the impact of student dropout caused by an AD policy on student graduation is obtained.

Our results show that the size of an AD threshold has a positive influence on student dropout. The higher the AD threshold, the higher the level of student dropout in the program. The second stage results show that student dropout caused by selection after the gate and

\[13\] Note that we remove the sector fixed-effects in Model 3. We also do not include Model 4 in our analysis. Both interventions are necessary because we do not obtain precise estimates due to a small population, a small proportion of compliers and the inclusion of an unordered variable.
graduation rates (which are conditional on ‘surviving’ the first-year) are positively correlated. The more selective the first year of higher education, the higher the graduation rates. We also found that some program characteristics have a significant influence on the dropout rate. In particular, the higher the achieved level of the program and the more students are satisfied, the lower the dropout rate. Further, we observed a significant influence of some program characteristics on the graduation rate of students. The more students are satisfied, the higher the graduation rate. The percentage of first-year female students has a positive influence on student graduation. Finally, the quality of the program is positively correlated with student performance. The results are robust for alternative specifications of the instrument, as well as for a nonparametric IV estimation (instead of the parametric 2SLS).

Two final remarks are in place at this point. First, the paper at hand does not advocate to stimulate student dropout. Given the social and individual costs, student dropout should be kept to a minimum. This can be achieved by a good matching practice at the gate or by student guidance practices during the first-year. Despite these practices there are always some students that did not enroll into the correct program and thus have a high probability of dropping out. Selection after the gate can be a solution for this phenomenon. Our results indicate that the selective nature of the first-year needs to be strengthened. Indeed, a good selection mechanism, such as an AD policy, ensures high student graduation rates. Our study shows that first-year student dropout, due to selection ‘after the gate’, has a positive influence on graduation rates. Consequently, although selection after the gate leads to more dropout in the first-year, it will lead to a reduced dropout in the following years of the program. This, in turn, will lead to a substantial cost saving. Indeed, in the Netherlands, 37,164 first-year students enrolled in an university and 86,781 in a higher vocational institution during the academic year 2004-05 (Centraal Bureau voor de Statistiek; 2014a). In the first year 15% of the higher vocational students dropped out against 8% of the university students. Five years after the first enrolment the dropout rate rises to 13% for the university students and 22% for the higher vocational students (CBS, 2014b, 2014c). The government costs per student are, on average, €6000/year. If we succeed in preventing dropout after the first-year by selection after the gate, we achieve a cost saving of 25 million euros for the university sector and a cost saving of 57 million for higher vocational education (Rijksoverheid, 2013).

Second, it is possible that in time less students will dropout due to the AD policy. Research showed that students base their learning effort on grading standards (e.g., De Paola & Scoppa, 2007). In the Dutch context, this means that students will increase their learning effort in relation to the size of the AD policy threshold (i.e. the learning function, see section 3). Thus, in the medium run, an AD policy will trigger endogenous student responses. On the other

---

14We base these calculations on student data of the academic year 2004-05 because the use of selection after the gate was not so common.
hand, education institutions will also implement intended and unintended changes based on the presence of an AD policy. It is possible, for example, that secondary institutions will start preparing their pupils for selection after the gate (i.e. the AD policy). Teachers at higher education institutions may also change their curriculum or teaching style in such a way that more students will pass the AD policy threshold. Given that this paper considered the first years after the introduction of AD policies in the Netherlands, it is unlikely that this endogenous outcome prevailed yet.

References


het bindend studieadvies in het hoger onderwijs”. Utrecht, the Netherlands: Inspectie van het Onderwijs.


