TIER

Interventions in higher education and their effect on student success: a meta-analysis

Eline Sneyers, Kristof De Witte

TIER WORKING PAPER SERIES
TIER WP 17/04

Universiteit van Amsterdam Maastricht University Leading in Learning! rijksuniversiteit groningen

Interventions in higher education and their effect on student success: a meta-analysis

Eline Sneyers[1] and Kristof De Witte[1,2]

1. Top Institute for Evidence-Based Education Research, Maastricht University, the Netherlands
2. Faculty of Business and Economics, KU Leuven, Belgium

December, 2016

Abstract

This paper provides a meta-analysis on the effect of academic probation, student-faculty mentoring and need-based grants on various student outcomes. Using 25 (quasi-) experimental studies, we find that an academic dismissal policy has a significant negative effect on retention (d=-.17), while it does not have an effect on graduation. Student-faculty mentoring, on the other hand, has a positive significant effect on both retention (d=.15) and graduation (d=.10). Need-based grants are proven to have a positive significant effect on enrollment (d=.05), retention (d=.05) and graduation (d=.05). Based on the general effect sizes of each intervention, student-faculty mentoring has the largest influence on student outcomes. The latter intervention improves retention and graduation of the treatment group by, respectively, 8% and 5% compared to the control group.

JEL-Classification: B4; I21; I23

Key words: Higher education; Academic dismissal policy; Need-based grants; Student-faculty mentoring; Student outcomes; Meta-analysis

1. **Introduction**

Increasing higher educational attainment and student success is high on the agenda of the EU-28 Member States. Universities and higher vocational institutions are trying to reach this goal by implementing various interventions. Common practices include pre-enrollment orientation; individual career counseling services; tutoring programs; residence hall programs; programs for racial/ethnic minority students; freshman seminars; selection practices; financial aid and learning communities

(Habley & McClanahan, 2004). A significant part of the research has focused on evaluating the effectiveness of these interventions. Taster days, for example, increase the intention of students to enroll in higher education (Austin & Hatt, 2005). Intake interviews are controversial, given the mixed results on student success found in the literature (e.g. Morris, 1999). Scrivener and Au (2007) found that career counseling for students at risk increased first-year credits and first-to-second year retention. A study of Angris, Lang, and Oreopoulos (2009) showed that financial aid in combination with peer mentoring leads to more credits, higher grade point average (GPA), and lower levels of academic probation for female students. Student advising also has a significant positive impact on retention (Seidman, 1991). First-year experience programs lead to an increase in student success (Schnell, Louis & Doetkott, 2003). The tracking of students' motivation also seems to be effective. Stratil, Schreiner and Noel (2001) have indicated that, after controlling for GPA, a motivation questionnaire can make a distinction between persisters and non-persisters. The selection of first-year students for the second year based on the number of acquired credits stimulates student success as well. Indeed, Calcagno, Crosta, Bailey and Jenkins (2007) have found a positive relation between first-year credits and the likelihood of graduation. It is clear that there are many options to stimulate student success and educational attainment.

The European Commission has formulated three broad types of measures to increase retention and student success: (I) remove financial barriers to broaden participation; (II) improve guidance and counselling, and help students to choose an appropriate course; and (III) develop skills profiles relevant to the world of work (European Commission, 2015). In this study, we focus on three student success interventions that are in line with these recommendations, i.e. grants, academic dismissal (AD) policies, and student mentoring. In particular, the introduction of grants leads to the removal of financial barriers for students. Mentoring and AD policies, on the other hand, improve the match between students and study programs and/or increase the guidance and counseling. The above interventions often focus on first-year students as the first year is seen as the most critical period for the connection between students and the academic programs (Yorke & Longden, 2004).

Academic dismissal policy[1]. An academic dismissal (AD) policy (or academic probation) is a form of selection 'after the gate'. It is a performance-based selection mechanism to remove students who are making unsatisfactory academic progress from the institution. Performance-based selection systems have a general structure. If a student's performance falls below the minimum standards of the enrolled institution, the student gets a warning which serves as a wake-up call and can lead to escalating penalties. In the US and Canada, academic probation is implemented in most higher education institutions. It is a generally accepted intervention, and is introduced to ensure that enrolled

[1] We also use the term 'academic dismissal policy' and 'academic probation' interchangeably.

students make satisfactory progress. It is based on students' GPA. When students do not reach the GPA norm, they are placed on academic probation. If their grades do not improve by the following evaluation, they face suspension (Lindo et al., 2010). In the Netherlands, a comparable system is used in many universities and vocational institutes, i.e. the AD policy (in Dutch “Bindend Studie Advies”) (Van Heerden, 2013). The main difference between the Dutch AD policy and academic probation relates to the threshold unit. While institutions in the US and Canada base their passing norms on an average performance measure called the GPA, institutions in the Netherlands use norms based on credits. Every course that students follow is linked to a number of credits, and, when students pass the course, they earn the corresponding credits. Students who do not earn enough credits to reach the threshold, and thus make substandard progress after the first year, are dismissed.

Academic probation is in line with the paradox of institutional commitment (Tinto, 1987). It states that institutions that are willing to encourage students to leave are the intuitions where students are more likely to stay. Tinto (1987) has argued that institutions which are more committed to the education of students are prepared to tell them when it is in their best interest to leave. Those institutions are also more likely to have students more committed to them and are therefore, more likely to persist until graduation (Tinto, 1987). On the basis of the above considerations, we can assume that an AD policy strengthens the selective function of the first year. Students who are not likely to finish the study program can be detected and be dismissed early in the program, allowing programs to select and continue with the most talented and motived students after the first year (De Koning et al., 2014). By eliminating those students who do not master the necessary knowledge and capacities to complete their studies, the learning environment should improve for the remaining students and for teachers. In line of the above statements, we can conclude that the effect of academic probation on dropout can go both ways. Indeed, students who are not likely to finish their studies are dismissed and cause more dropout. Another option is that the most committed students choose which subjects to study[2]. As a result, the dropout rate decreases. In each case, an AD policy should lead to higher graduation rates. There are, however, also a number of disadvantages connected to the use of an AD policy. These can interfere with the mechanism behind academic probation. First of all, Bénabou and Tirole (2003) have argued that the impact of performance standards on students differs by the type of students. Setting performance standards can motivate some students to improve their performance, while others are discouraged from making any attempt at all. The possibility exists that some suitable students do not even enroll in the academic program of their choice due to the AD policy. Next, it is possible that teachers increase the required achievement level for the selected group of students. Based on "grading on a curve", teachers dismiss a fixed number of students, independently of their performance (Sadler, 2005). Finally, the effectiveness of an AD policy as a

[2] However, it is important to note that, from the early literature concerning the AD policy in the Netherlands, it has become clear that students are indifferent to studying at an institution with an AD policy, and do not see it as a threat (Felsö, van Leeuwen & Zijl, 2000).

selection instrument is uncertain. The relationship between first-year progress and graduation is clear (e.g. Gijbels, Van der Rijt & Van der Watering, 2004; Pascarella & Terenzini, 1991). However, when the first year is too difficult and no help is provided (e.g. by tutoring), academic probation can lead to the dismissal of capable students.

Mentoring. Mentoring[3] can be seen as a situation in which a more-experienced member of a higher education institution maintains a relationship with a less-experienced, new member and provides information and guidance. In this context, we refer to the mentor as a faculty member or a professional. The less-experienced individual, or the mentee, is the student. Undergraduate student-faculty mentoring is implemented in various forms in colleges and universities. Some mentoring programs only provide academic help, while others also focus on study skills and social needs (Bettinger & Baker, 2014). Mentors and mentees are recruited by administrative offices and are matched based on criteria such as academic specialty or ethnicity. Often these programs are enhanced with other support services (such as financial aid) with the intention to create a learning and campus environment that stimulates retention and student success (Pascarella & Terenzini, 2005). Not surprisingly, the costs of student-faculty mentoring nationwide may be many millions of euros.

Mentoring addresses the problem of student success in multiple dimensions. First of all, student success and retention can be adversely affected by a lack of information. Indeed, research has shown that many students have little information concerning the course requirements (Goldrick-Rab, 2010). Student mentoring can help close this gap by providing students with the necessary information. Besides academic information, students can be informed about institutional resources. It can be expected that first-generation students and students with a low socioeconomic status especially benefit from this structured advising (Deil-Amen & Rosenbaum, 2003). Next, academic preparation is key for student retention. However, many students do not master the required academic knowledge and are not academically prepared. Mentors can help to bridge this gap. Finally, as Tinto (1975) has stated, feelings of academic or social separation often lead to dropout. Mentoring can help students to integrate into the university and college community.

Need-based grants. Need-based grants are a form of financial aid and are closely linked to family income and economic status. There are a number of measures which determine financial-aid status, such as the family's expected contribution, number of dependent family members, and student status. Besides family income, other factors are used to determine each student's financial-aid formula. Need-based aid can be provided by the government or by higher education institutions. Need-based grants are used in many countries such as the US, France, Canada, and Denmark.

[3] It should be noted that mentoring, tutoring, coaching and advising are closely related. Hence, we use these terms interchangeably.

Financial aid can have a large impact on the enrollment decision of low-income students. Students enroll in higher education if the perceived present discounted value of the benefits of a higher education career exceeds the present discounted value of the costs of enrollment. Reducing the cost of going to college should therefore stimulate enrollment. The influence on retention and graduation rates is less clear. Indeed, the grant can act on liquidity constraints. Due to a subsidy, the most disadvantaged students, who have to work to cover maintainance costs, can reduce the time taken by this work and make more time free for studying (Avdic & Gartell, 2015). Next, students who receive the grant can feel that they have better opportunities than others. This can motivate them to study more effectively and with more commitment. However, even though financial aid reduces the hours spent doing part-time paid work, there is no guarantee that these hours are used for studying. Further, since it is less costly to study, the risk of retention decreases, but on the other hand, the time-to-completion increases (e.g. Bettinger, 2004). Aid can even have a negative effect on student success as students with a low probability of completion are induced to enroll due to the lower financial costs they incur for their education.

This meta-analysis investigates the effect of three interventions on student outcomes. The three discussed interventions (i.e. academic dismissal policy; grants; and student-faculty mentoring) are all frequently used in educational institutions to promote student performance. Their effects on student success, however, remain unclear as studies of different design quality are compared with each other. A systematic review of studies with an evidence-based design (i.e. experimental or quasi-experimental) seems imperative. To our knowledge, there are no previous meta-analyses of randomized trials and quasi-experiments on academic probation, need-based grants, and student mentoring and their effects on enrollment, retention, and graduation.

The reminder of the paper is organized as follows. In Section 2, we present the methodology and selection of the literature. Section 3 reviews the results including the literature concerning the interventions. We conclude this study with policy advice.

2. Methodology

2.1 Criteria for inclusion of studies in the review

For this study, only these evaluation studies that have the following characteristics were included. First, included studies needed to assess the impact of: (I) academic dismissal (AD) policy; (II) merit-based grants; or (III) student-faculty mentoring, which included higher education outcomes relevant to the research questions. These studies also needed to use a randomized experiment or a quasi-experiment with evidence of baseline control on a main outcome. Moreover, this meta-analysis included studies that randomly assigned entities (at individual, class, or school level) to intervention or control conditions. The control or comparison condition comprises observations that received no intervention or were exposed to standard practice. We opted to include experiments because they can provide statistically unbiased estimates of the effect of an intervention and can control for both known and unknown covariates (Boruch, 1997). Because randomized trials (RCT) are not always an option (for example, because it is not ethical), we also included quasi-experimental designs that control for baseline differences between the control and intervention group. Quasi-experimental designs (QED) can thus provide causal evidence (e.g. Blundell and Dias, 2009; Van Klaveren & De Wolf, 2015). We included three quasi-experimental designs: (I) regression discontinuity designs; (II) propensity score matching; and (III) difference-in-differences techniques. It is important to note that, although AD is a well-known intervention, little empirical evidence exists. As a result, we allowed the inclusion of cohort studies during the selection of academic probation studies. In cohort studies, an intervention and a control group is included. However, the units of analysis are not randomly assigned to one of the conditions and the equality of (un)observables is not guaranteed. We also run the analyses on a restricted sample which only includes experimental and quasi-experimental studies. Further, included studies need to discuss at least one quantifiable outcome measure of student enrollment, dropout, or graduation. We did not collect data on other intervention impacts, such as those on satisfaction, number of courses completed, or behavior. Also important is that the use of enrollment, dropout, and graduation as main measures may not match with how investigators in the original studies calculated their outcome measure. Fourth, the studies needed to be published or made available after January 2000. We also included both published and unpublished studies (e.g. dissertations, technical reports, and conference papers). Note that the inclusion of unpublished studies is controversial. Indeed, unpublished studies may have data driven or analytical problems. On the other hand, it is known that studies with positive and significant results have a higher likelihood of being published. This meta-analysis presents results that are not subject to upward bias (Van Klaveren & De Wolf, 2015). Finally, we only selected studies that report the effect size (ES) or allowed us to extract the data needed for the calculation of the ES. Unfortunately, not many researchers report the ES of an intervention. Therefore, studies that did not report the ES of the intervention needed to provide us with the necessary data to calculate the ES. This sometimes required us to make assumptions concerning the assignment of control and intervention groups. If these assumptions had not been made, this meta-analysis would have consisted of a very limited number of studies. Still, while various studies did meet all of our criteria, they did not provide the data necessary for the computation of an ES or required us to make too many assumptions in order to calculate the ES (e.g. Bettinger, 2004; Kane, 2003; Schudde & Scott-Clayton, 2014).

We also excluded a number of studies based on three exclusion criteria. Studies that evaluated the effect of academic probation and focused on an increase of the AD threshold were not selected in this meta-analysis. This meta-analysis only includes studies that investigate the effect of an AD policy by comparing an intervention group with an AD policy and comparison group without the AD policy.

A study by Woelders, Visser and Rijksbaron (2013) which investigated only the increase of the AD threshold was therefore excluded. Studies that focused on grants other than need-based grants (i.e. grants that are awarded based on income-level) were also not included. In addition studies that evaluate (partially) merit-based grants or grants awarded on the family status (e.g. death of father[4]), and studies that only focused on loans were excluded. Examples are the studies by Dynarski (2003), and van der Klauw (2002). Finally, we excluded studies that investigated the effect of mentoring but only focused on peer mentoring, online mentoring, or self-development mentoring. In this study, we are only interested on the effect of student-faculty mentoring (i.e. faculty-student mentoring) or mentoring executed by an external service. We believe that peer, online, and self-guidance mentoring have a different effect on the investigated parameters (such as the development of friendship). For this reason studies by Salinitri (2005) and Sandner (2015) were excluded.

2.2. Search strategies for identification of relevant studies

Our goal for the literature search is to identify relevant papers (published and unpublished). We opted to use three search strategies for the identification of relevant studies. We started with an electronic searcher of databases. Moreover, we searched available resources and databases at the University of Maastricht (through EBSCO host). Examples of the databases are ERIC, EconLit and Business Source Complete. Table 1 contains the complete list of databases and the keywords used can be found in Table 2. Next, we checked the Reference section of the selected papers in order to determine whether possible eligible studies were listed (i.e. snowballing method). We did not check the Reference section of ineligible studies. Lastly, Google and Google Scholar were used for a broad search of the World Wide Web.

2.3. Keyword strategies for bibliographic databases

In order to search the databases, we decided to conduct a relatively broad search. We preferred to have many titles and abstracts to go through with a broad search, rather than potentially miss relevant studies with a too narrow search. We used one search strategy, i.e. the combination of keywords. Therefore, we developed a list of keywords to identify the study eligibility criteria: (1) keywords that were relevant for the interventions; (2) keywords that were relevant for the outcomes of enrollment, retention or graduation; (3) keywords that were relevant for the education level; and (4) keywords that were relevant to the methodology. As we found new studies, we adapted our search terms. Table 2 represents the search words used for the final pass through the databases.

[4] Although there is a correlation between the death of a parent and income, we cannot rule out that some wealthier families also benefit from these kind of grants. As a result, we did not incorporate studies that investigated grants based on loss of a family member.

Table 1: Consulted databases

Databases
CINAHL
EconLit
ERIC
greenFILE
PsycARTICLES
PsycINFO
SocINDEX
Psychology and Behavioral Sciences Collection
MEDLINE
PsycBOOKS

Table 2: Used search words

Intervention	**Effect**	**Context**	**Outcome**
Academic dismissal policy	Experiment	College	Dropout
Academic probation	Quasi-experimental	University	Retention
Bindend studieadvies	Empirical	Higher education	Persistence
Mentoring	Effect	Further education	Graduation
Coaching		Post-secondary education	Completion
Advising		Undergraduate	Performance
Guidance			Enrollment
Need-based grant			
Need-based aid			

The search based on this search strategy identified a large number of studies. Many studies were easily excluded because they were not relevant for the proposed meta-analysis. Each citation was reviewed, and it was determined whether the study could proceed to a second screening. If so, the full text of the eligible study was retrieved. Unfortunately, we could not find a full text for all eligible studies. However, when a full text could be found, the study was read to make sure it fulfilled the inclusion criteria, did not fulfil the exclusion criteria, and included at least one of the relevant outcome measurements. When a study was selected, we used a coding instrument to extract the correct information from each study. The instrument allowed us to identify the following characteristics of each study: the author, the publication (i.e. type of document and year published), the context (i.e. country), the evaluation design (i.e. RCT, QED or cohort study), the treatment group, the control condition, the participants, and the outcomes (see Appendix).

2.4. Criteria for handling statistical dependencies

We used three criteria to be sure that our findings were independent. This insured that effect sizes were not correlated with each other because, for example, authors had used the same data set or had used multiple designs in one study. First, we considered an evaluation study as distinct if it uses a unique data set. Sometimes investigators publish multiple articles or reports on the same intervention and on the same sample (for example, when investigating the long-term effects of a study, Campbell & Campbell, 1997, 2007). In this case, we opted to include the most recent study in our meta-analysis. Some authors opted to copy the intervention and the evaluation method of another researcher (for example, Casey, Cline, Ost & Qureshi, 2015; Lindo et al., 2010). When there were two separate samples, we considered these studies as distinct and use them both in this review. Second, we chose to use the study design that was the "most rigorous", or the analysis that provided the most controls. In some studies, multiple designs were used on the same sample. We then included the outcomes of the strongest design in our review. In some cases, researchers reported the results of multiple estimation models. We chose to include the model that includes the most control variables. Because regression-adjusted estimates for the control versus experimental groups theoretically reduce statistical noise, we relied on these estimates and used them in the analyses. Statistical noise can occur due to chance fluctuations, violations of the randomization (in the case of experiments), or uncontrolled variables (in the case of quasi-experiments). We have allowed one exception in this study, i.e. when we used the percentage of observations in the control or intervention groups and/or the mean of the dependent variable to calculate the ES (see Section 2.4.). This information is often based on the total number of observations without taking the missing values of the control variables into account. To ensure that we based our calculations of the ES on the correct data, we used the regression estimate of the model with the same number of observations that were used to calculate the two affected metrics. Unfortunately, this is almost never the regression-adjusted estimate with control variables (due to missing values). Finally, a number of studies only reported outcomes by specific subgroups such as gender (male/female) or status policy change (for example, a winner/loser due to changes in the grant system). Even though the outcomes for these subgroups can be very interesting for policy makers, they are difficult to handle in our meta-analysis. We decided to average the effect over the subpopulation and to use the overall effect in our meta-analyses.

2.5. Statistical procedures

In this study, we used the standardized mean difference (Cohen's d) as effect size metric for all outcome variables. We opted for the standardized mean difference (Cohen's d) because it is a flexible effect size metric, and many formulae are available to estimate the effect size from information reported in the selected studies (e.g. regression coefficients, probability levels). All effect sizes are coded in such a way that positive effect sizes represent better outcomes (i.e. higher enrollment, retention, and graduation). Standardized mean difference effect sizes were calculated as:

$$d = \frac{\bar{X}_{TG} - \bar{X}_{CG}}{SD_{Pooled}}, \tag{1}$$

where the numerator is the difference in group means from the intervention group and the control group. The dominator consists of the pooled standard deviation of those groups. Since we often used the unstandardized regression coefficient to calculate the standardized mean difference, this formula is also mentioned:

$$d = \frac{\beta_{unstandardized}}{SD_{DV}}, \tag{2}$$

where the numerator is the unstandardized regression coefficient (in the case of a binary intervention variable, this coefficient represents the difference between the mean of the intervention and that of the control group) and the dominator represents the standard deviation from the dependent variable (this information can be calculated from the mean of the dependent variable, since all outcomes in this study are binary).

The variance of the standardized mean difference was calculated using the following formula:

$$V_d = \frac{n_{TG} + n_{cG}}{n_{TG} * n_{cG}} + \frac{d^2}{2(n_{TG} + n_{cG})}. \tag{3}$$

A transformation procedure, commonly used in this study, is the logit transformation of cells frequencies to standardized mean differences. In this calculation the odds ratios for these data is first computed. Moreover, the *d* is calculated as follows:

$$d = \left[ln\left(\frac{A * D}{B * C}\right)\right] * \frac{\sqrt{3}}{\pi}, \tag{4}$$

where *A* and *B* are the counts of the "successes" and "failures" in the treatment group, and *C* and *D* are "successes" and "failures" in the comparison group.

It should also be noted that, because of presumed heterogeneity in the true effects across samples and countries, we reported the results of the random effects models in our analyses. Under the fixed-effects model, the assumption is that there is one true effect size which underlies all studies in the analysis, and that all observed differences are due to sampling error (Borenstein, Hedges & Rothstein, 2009). Under the random-effects model, the true effect size differs between studies since mixes of participants and the implementation of interventions also differ between studies. When we moved from the fixed-effects model to the random-effects model, the relative weights became more balanced. Consequently, extreme studies lost influence if they were large, and gained influence if they were small. We reported overall effects for all interventions on the corresponding outcomes (see Table 3).

Table 3: Organization of outcomes

Interventions		
Academic probation	**Faculty-student mentoring**	**Need-based grants**
<u>First-year retention</u>	<u>First-year retention</u>	<u>Enrollment</u>
A student persists from the first to the second year	A student persists from the first to the second year	A student enrolls in a study program in higher education
<u>Graduation</u>	<u>Graduation</u>	<u>First-year retention</u>
A student completes the study program within the nominal time or within 6 years	A student completes the study program	A student persists from the first to the second year
		<u>Graduation</u>
		A student completes the study program

3. Results of the meta-analysise

3.1 Description of found studies for each intervention

3.1. Pipeline of studies

524 studies were retrieved from our first search. As a second step, we determined whether a study seemed eligible based on the abstract and title. We also included extra studies through snowballing. Unfortunately, his left us with 71 studies. As a final step we retrieved the full-text reports and checked whether: (I) the studies matched with our research questions; (II) there was evidence of baseline control; and (III) the necessary information for the effect size calculation was present in each study. We maintained 25 studies which we coded for study characteristics and effect sizes. The pipeline of studies is illustrated in Figure 1[5].

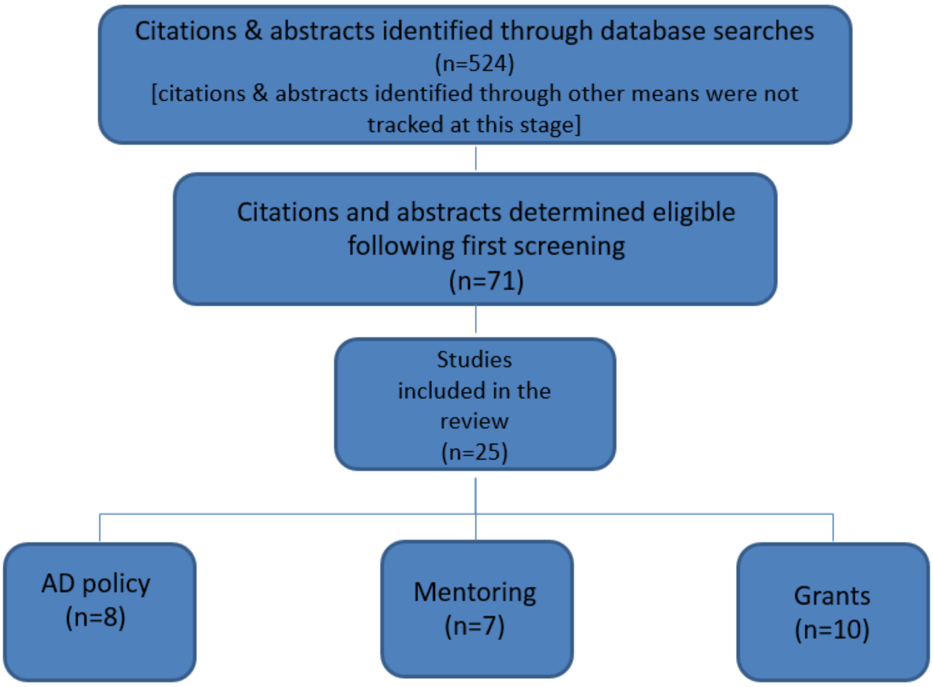

Figure 1: Pipeline for review sample

3.2. Descriptive statistics

Three broad interventions were included in this review. Eight studies evaluate the AD policy; seven are about student-faculty mentoring; and 10 are about need-based grants (see Figure 1 and the Appendix for a detailed description of each study). Most of these programs target universities (68%). Only 12% of the studies focused on colleges, and 20% focused on both types of institutions. The sample of studies was diverse concerning the nations of origin. Studies were conducted in seven different nations, with the US (n=14), the Netherlands (n=4), Canada (n=2) and Denmark (n=2) as the most common. In this review 24% of the studies used randomization while 60% made use of a quasi-

[5] Note that, when a study provides us with multiple effect sizes (e.g. when there are multiple independent samples present), we count this study as one in the descriptive statistics.

experimental design. When randomization was in place, we found one study that assigned classes to the control or treatment condition instead of to the individual student level. Four studies (16%) made use of cohort comparison. All of the studies were published between 2002 and 2015. Figure 2.2 shows that most of the eligible studies were conducted as of 2010. Almost half of the studies were published in a journal (48%). The majority of the other studies are discussion/working papers (20%). We also include some reports of international organizations (16%) such as the National Center for Postsecondary research, some conference papers (8%), and dissertations (8%). Most authors came to the conclusion that their intervention had a positive effect on the main outcomes (64%). 24% of the studies reported a negative effect, and 12% found mixed effects. Note that some studies found only very small effects.

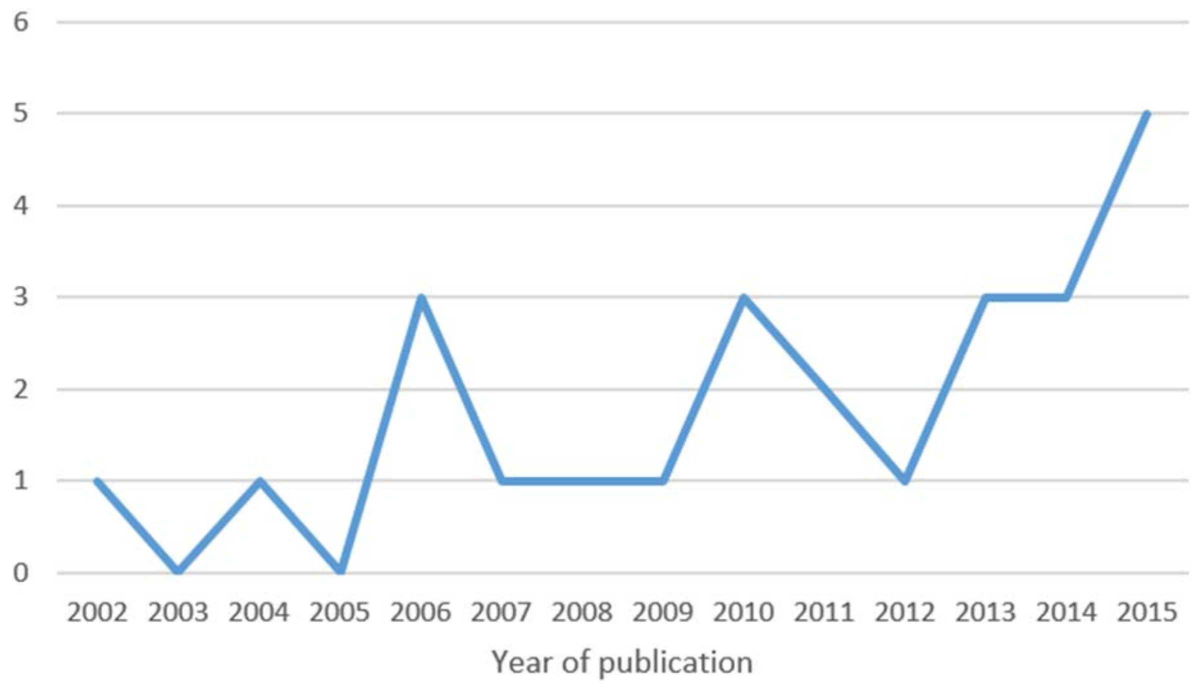

Figure 2: Number of included studies by year of publication

3.2.1.1 Academic Dismissal (AD) policy

We included eight studies concerning Academic Dismissal or Academic Probation. Half of the studies used a quasi-experimental design, mostly a regression-discontinuity design (RDD). The other four made use of cohort comparison. Despite the fact academic probation is a well-known intervention in many countries (e.g. the Netherlands, the US, and Canada), little empirical evidence was found. This might be due to two reasons. First, in countries such as the US and Canada, academic probation is common practice and is installed in most universities and colleges. As a result, most studies focusing on academic dismissal policies in these countries try to find the effect of interventions on dropout rates and graduation rates of students already on probation, or try to find which characteristics predict whether a student will be on probation. Second, in other countries academic dismissal policies are a novelty. As a result, there are no good experimental or quasi-experimental studies.

Lindo et al. (2010) found that being placed on academic probation after the first year discourages some students from returning to school, while the students who remain improve their subsequent performance. Similarly, Fletcher and Tokmouline (2010), Casey et al. (2015) and Chi and Dow (2014) have also shown that retention rates decrease. While Casey et al. (2015) found a negative effect on graduation rates, Chi and Down (2014) found a positive effect on completion rates. Fletcher and Tokmouline (2011), however, showed mixed evidence concerning the effect on graduation rates. Duijndam and Scheepers (2009) indicated that compared with programs where no AD policy is in place, the introduction of an AD policy in a business management program resulted in earlier withdrawal of students in the first year. This shows that fewer students linger in the program because of the introduction of a credit threshold. In addition, Arnold (2015) compared Bachelor programs in the academic years 2002-03 and 2007-08. During these years some programs implemented an AD policy (i.e. intervention group), while others did not (i.e. the control group). He found that the introduction of an AD policy increased the first-year dropout rate and the 4-year completion rate of university students.

3.1.2. Student-faculty mentoring

We also found seven studies focusing on student-faculty mentoring. Four of these studies made use of a quasi-experimental design (such as matching or a Difference-in-differences (DiD) estimator). The other three studies made use of a random experiment. The studies were conducted in Germany and the US. Although student mentoring is a practice that is often applied, we did not find many studies investigating the effect of student-faculty mentoring on performance indicators of higher education students. We did, however, find many studies focusing on the effect of peer mentoring on student performance indicators. Peer mentoring is a practice whereby the mentoring function is often fulfilled by a more experienced student (e.g. a graduate student or a senior student). Since we are only interested in the effect of faculty-student interaction, we did not take the latter group of studies into account in our review.

Given the costs of mentoring, one might expect a substantial number of papers on the effect of mentoring on student outcomes. However, the evaluation research in this field is limited. Bettinger and Baker (2014) used a randomized experiment to test the effectiveness of ‘Inside Track’. This intervention consists of an individualized coaching service which has proved to increase retention and graduation rates. Using matching, Kot (2014) showed that centralized advising increased first-year attritions. A study of Brock and Richburg-Hayes (2006) investigated the effect of an open-door program on student outcomes. Moreover, they found that enhanced counseling layered with financial incentives leads to fewer dropouts. This is in line with a study of Campbell and Campbell (2007) which showed that retention increases due to mentoring. However, graduation rates remain unchanged[6]. Hoffer (2010) studied the effect of an experimental mentoring program in US university, and found a positive effect on retention. A study of Swanson (2005) investigated the different effects between normal advising meetings, development advising meetings, and strength-based advising meetings. It was shown that the most intensive advising led to higher retention rates compared with normal advising. The study of Visher, Butcher and Cerna (2011) showed no effect on retention.

3.1.3. Need-based grants

In this meta-analysis, 10 studies on the effect of need-based grants were included. One study made use of a randomized design. The other studies implemented quasi-experimental designs (mostly RDD). Most studies were conducted in the US and in Denmark. We also found a German, Italian and French study. We observed a number of studies focusing on merit-based aid. Students are eligible for this aid when they reach certain performance thresholds. We did not include these studies in our review. Only studies whereby initial eligibility is based on socio-economic status or income are taken into account. Note that the criteria that students need to fulfil in order to retain the scholarship after the first-year are not important in this review.

Castleman & Long (2012) examined the impact of the Florida Student Access Grant (FSAG) on academic outcomes such as enrollment and college persistence. Using a regression-discontinuity strategy, the authors showed that aid positively affects enrollment and persistence. Moreover, the grant also increases the likelihood of obtaining a Bachelor's degree by 4.6 percentage points. The finding concerning enrollment and retention is in line with many other studies, such as Goldrick-Rab, Douglas, Kelchen and Benson (2012) and Fack and Grenet (2015). Bettinger (2015) assessed the Ohio College Opportunity Grant initiative, and, through a difference-in-differences identification strategy, highlighted its positive effect on first-year performance, such as a reduction in dropout rates. Several outcomes of academic performance have been considered by Arendt (2013) to assess the effect of the Danish reform of the student grant and loan system. The results suggest a positive impact on the dropout rates, but no overall effect on completion rates. Agasisti and Murtinu (2014) assessed the impact of receiving a grant for a cohort of students enrolled at the "Politecnico di Milano" in the academic year 2007/2008. The empirical analysis focuses on a wider range of academic results, i.e. dropout and time to graduation. The authors found that obtaining a grant positively affected retention and completion, especially for immigrants, students whose family reside in another region, and those attending engineering courses. Studies of Fack and Grenet (2015) and Alon (2011), which both exploit a quasi-experimental design, also found no effect on graduation.

[6] Note that in the study of Campbell & Campbell (2007) the target group of the intervention consists of newly enrolled students (also transfer students). Although they are not defined as first-year students, they are in their first-year at the institution, and face the same issues as first-year students. Consequently, we include this study in the review.

3.2. Meta-analysis

We estimate the overall mean effect size (*d*) using inverse variance random effects weights. When the interventions have a positive impact on the outcome variables (e.g. an increase in enrollment or retention), we scale the standardized mean differences (Cohen's d) as positive. When the intervention has a negative impact (e.g. graduation decreased) or no effect (e.g. the enrollment rate in both groups remained the same), we scale the effect size as negative and zero, respectively. An effect size estimate of .5 reflects a 5/10 standard deviation improvement for students in the treatment condition compared with students in the control conditions. We also present heterogeneity data, including the I^2, the tau^2 (between studies) and the Q-test[7] for each intervention analysis. All these metrics show how well the mean effects represent the sample of studies in the analysis. Given the expected variation in samples, countries, and design methods, the variability in effect size is also large.

3.2.1. Effects of the Academic Dismissal (AD) policy

In this section, we report the overall effects of academic probation on retention and graduation. In Figure 3, the results are represented for 8 studies (and 11 effect sizes) that measure the effect on retention rates. We observe that the average treatment effect is negative (*d*=-.17; 95% CI [-.23; -.12]), and ranges from -.45 to -.04. All studies report a negative effect size. The difference between students on academic probation and regular students is a fifth of a standard deviation. Moreover, based on the Percentage Improvement in Treatment over Control (BESD), an academic dismissal policy leads to an 8.5% decrease in retention (see Table 5). Following Cohen (1988) and Hattie (2015) this is a small effect size. As expected, heterogeneity statistics indicate large variability across effect sizes ($Q=155.09$; $df= 10$; $p<.001$; $I^2=93.06$; $tau^2= .0077$).

Next, we consider the effect of AD policy on graduation within the nominal study time and within 6 years. Five studies or eight effect sizes are included in the analysis of graduation within the nominal time (see Figure 4). Collectively, the average effect size is zero (*d*=.00; 95% CI [-.04; .04]), and ranges from -.08 to -.13. This suggests that an AD policy does not have an effect on graduation within the nominal study time. This is not surprising given that three cases report a negative effect size and two reported no effect size. Only three cases show a positive effect. We also find a lot of heterogeneity ($Q=43.88$; $df= 7$; $p<.001$; $I^2=84.00$; $tau^2= .0024$).

We also do not find a statistical effect of academic probation on graduation within 6 years (see Figure 5). Seven effect sizes are used in the analysis. A small overall negative effect size is

[7] For any observed Q, a low P value provides evidence of the heterogeneity of intervention effects (i.e. that studies do not share a common effect size). The Tau^2 test presents an estimate of the magnitude of variation between studies. The I^2 statistic describes the proportion of variability in effect estimates due to heterogeneity rather than to chance (Higgins, 2008)

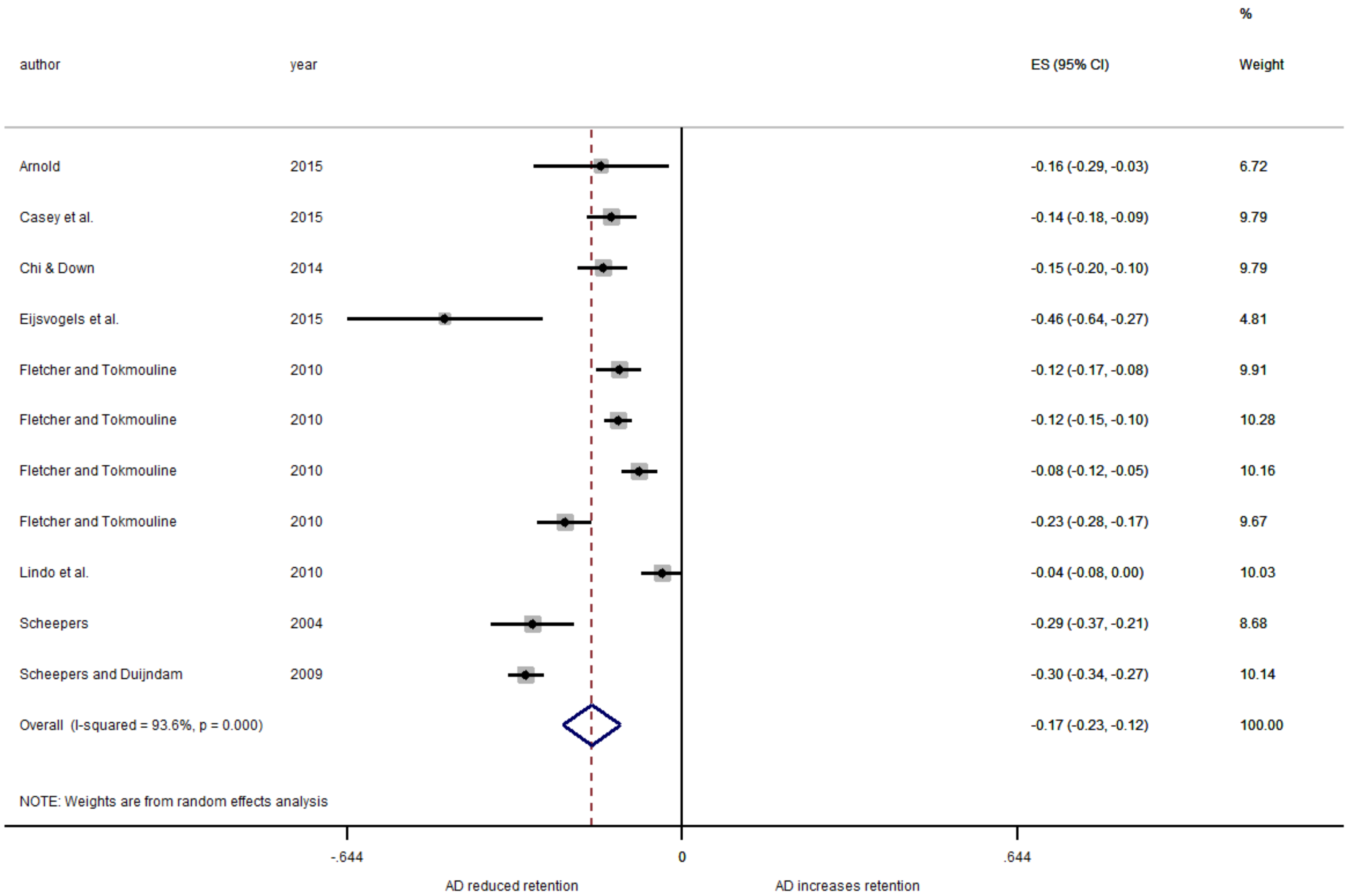

Figure 3: Main effects of AD policy on retention rates. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output.

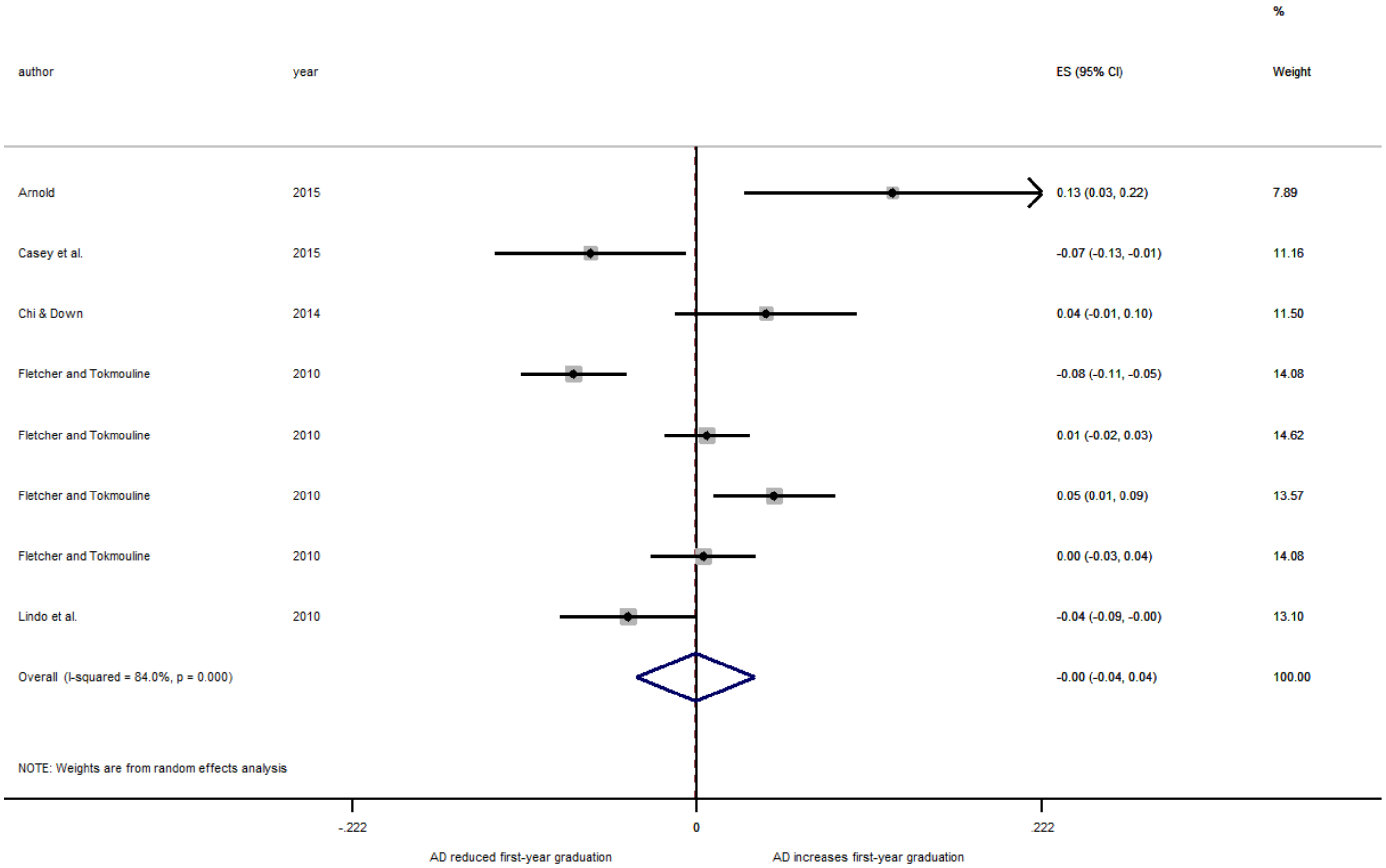

Figure 4: Main effects of AD policy on graduation within the nominal time. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output

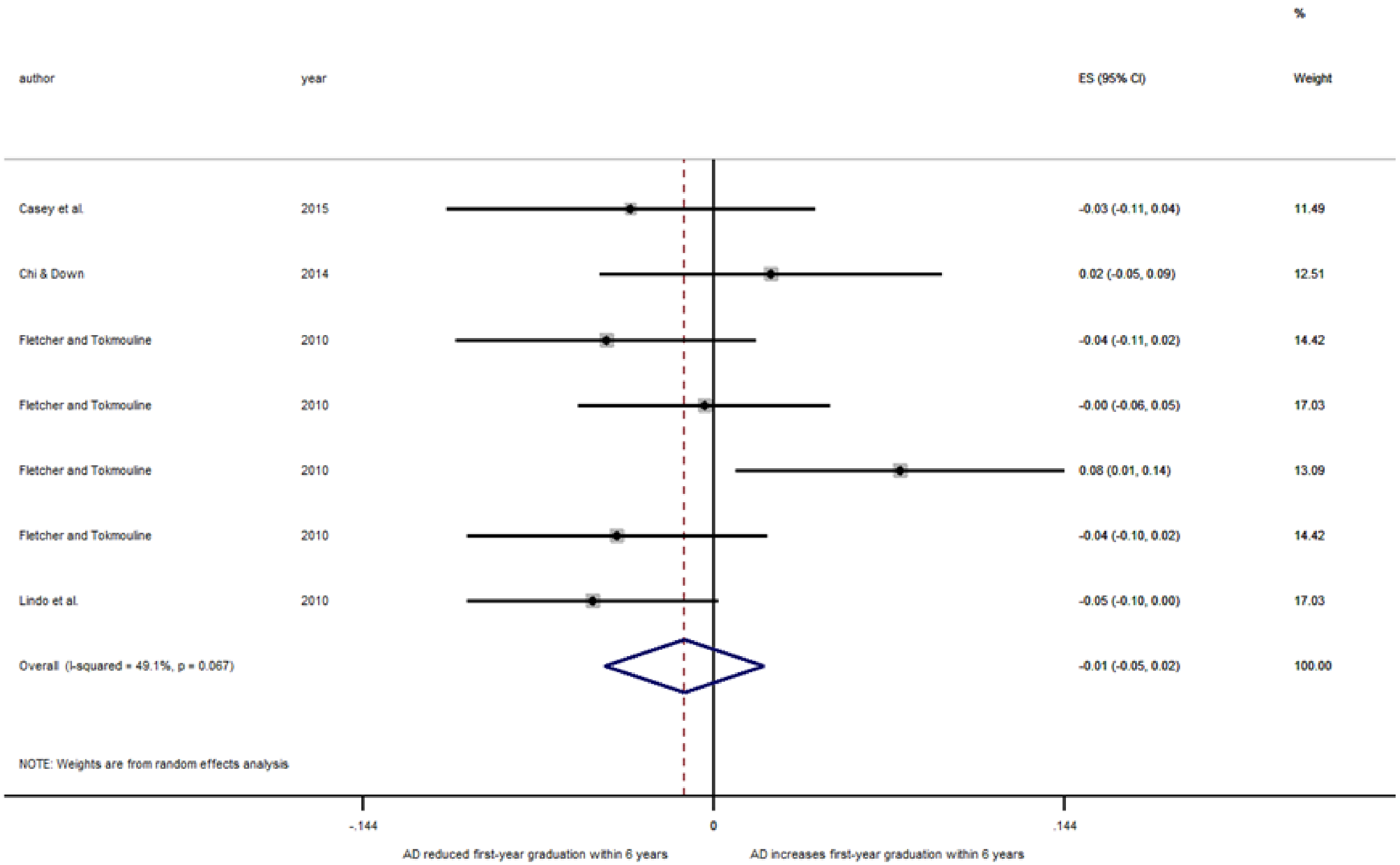

Figure 5: Main effects of AD policy on graduation within 6 years. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output**.**

found (*d*=-.01; 95% CI [-.05; .02]), and the effects range from -.05 to -.08. This suggests that an AD policy does not have an effect on graduation within 6 years. Moreover, five cases report a negative effect size and only two cases show a positive effect. We also find a moderate amount of heterogeneity (Q=11.78; df= 6; p<.010; I^2=49.10; tau^2= .0010).

3.2.2. Effect of student-faculty mentoring

This section reports the effect of student-faculty mentoring on two outcome variables (i.e. retention and graduation). Seven studies are included in estimating the effect of mentoring on student retention (see Figure 6). We observe an average positive effect size
(*d*=.15; 95% CI [.06; .23]), and the effects range from -.02 to .41. The difference in retention rate between the mentored and non-mentored students is .15 of a standard deviation. Although student-faculty mentoring seems to reduce dropout significantly, this effect size is perceived as small (Cohen, 1988; Hattie, 2015). Moreover, as a result of mentoring, dropout decreases by .7.5% (see Table 5). Only one study reports a negative effect, the other studies all show a positive effect. Again, we find a substantial amount of heterogeneity (Q=19.90; df= 6; p<.05; I^2=69.9%; tau^2= .0066).
Looking at the graduation rates[8], we find a positive average effect of mentoring (*d*=.10; 95% CI [.01; .19]; range from .09 to .14) (see Figure 7). Between the students who received mentoring and those who did not there is a .10 difference of a standard deviation. This indicates that mentoring leads to a 5% increase in graduation rates. Cohen (1988) and Hattie (2015) identify a Cohen's d of .10 as a small effect. However, the number of studies included in these analyses is limited (i.e. two studies), so the results have to be interpreted with caution. We do not find evidence of heterogeneity, but this is probably due to the small number of studies (Q=.27; df= 1; p=.602; I^2=0%; tau^2= .0000).

[8] Note that the authors speak about the completion of a degree, but do not mention in how many years after the start of the program.

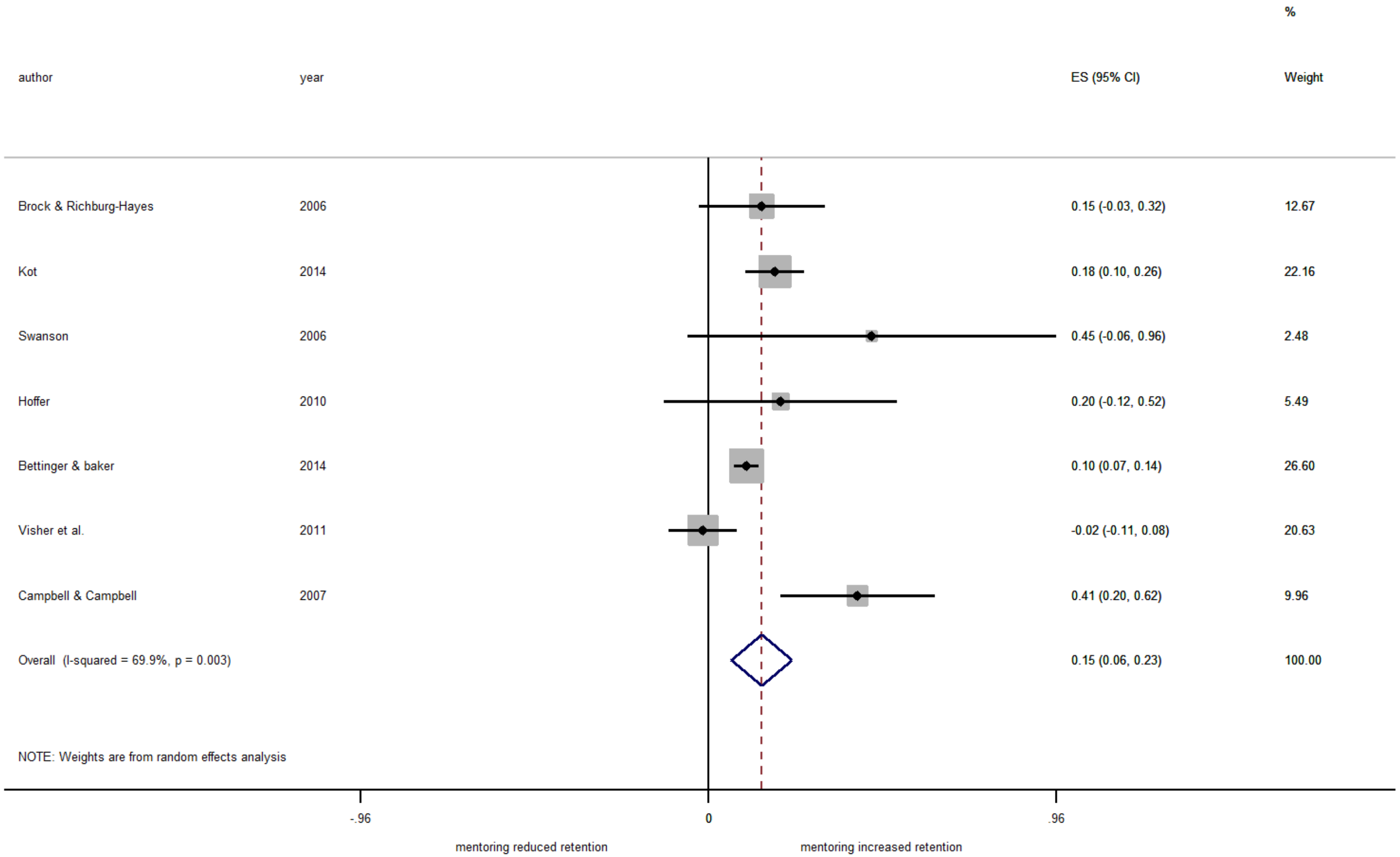

Figure 6: Main effects of mentoring on retention. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output.

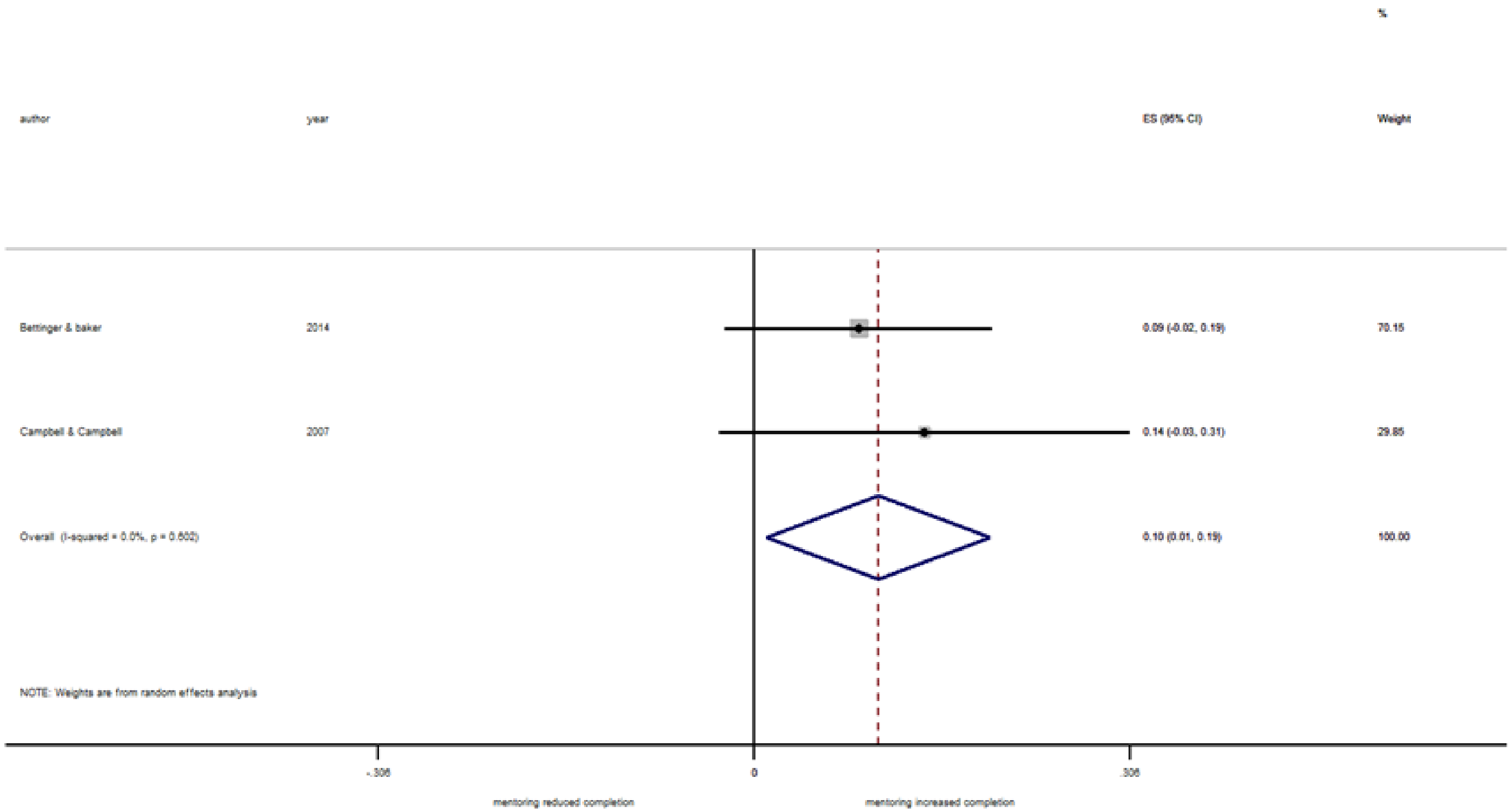

Figure 7: Main effects of mentoring on completion. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output.

3.2.3. Effects of need-based grants

We consider the effect of need-based grants on three outcome variables[9]. We start by looking at the effect of need-based grants on enrollment (see Figure 8). For these analyses we include six studies. We observe an average positive effect of need-based grants (d=.05; 95% CI [.03; .07]; range from .03 to .13). Moreover, there is a .05 difference of a standard deviation between the control and intervention group. Due to need-based grants, enrollment has increased by 2.5% (see Table 5). This effect is highly significant but small (Cohen, 1988). Thus, all studies show positive effects. We find a moderate amount of heterogeneity (Q=9.45; df= 5; p=.092; I^2=47.1%; tau^2= .0001).

Next, we present the results of the effect of need-based grants on retention (see Figure.9). Six studies are included, and we observe an average positive, but small, significant effect (d=.05; 95% CI [.04; .06]; range from .01 to .12). Hence, financed and non-financed students show a .05 difference of a standard deviation. This indicates that retention has risen by 2.5% (See Table 5). Again, all studies show a positive effect. We observe no heterogeneity (Q=4.85; df= 5; p=.434; I^2=0%; tau^2= .0000).

Finally, we show the results of the effect of need-based grants on graduation (see Figure 10). It is important to note that some authors mention graduation within the nominal study time, others use graduation within 5 or 6 years and one author looks at graduation independently of the necessary time. We used five studies and seven effect sizes (i.e. some authors use multiple outcome variables). We observe an average positive, small significant effect of grants on graduation (d=.05; 95% CI [.01; .08]). All the observed effect sizes range from .01 to .18. Graduation rates have increased by 2.5% due to need-based grants (see Table 5). Again, there is evidence of heterogeneity (Q=13.06; df= 6; p=.042; I^2=54.0%; tau^2=

[9] We transformed the effect sizes so they present the effect of a €1000 increase in grant. Castleman & Long (2012) indicate that this is common practice.

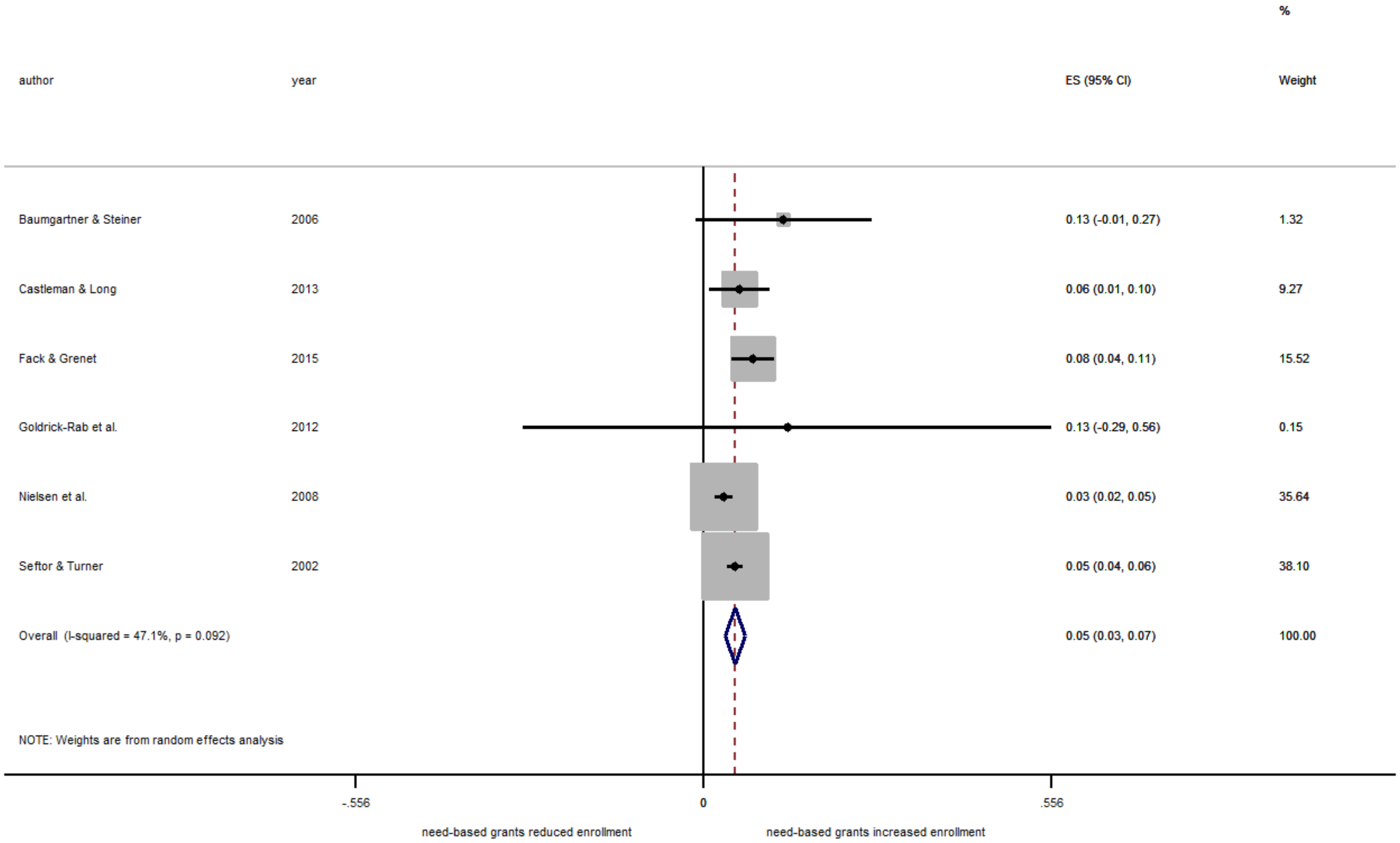

Figure 8: Main effects of need-based grants on enrollment. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output.

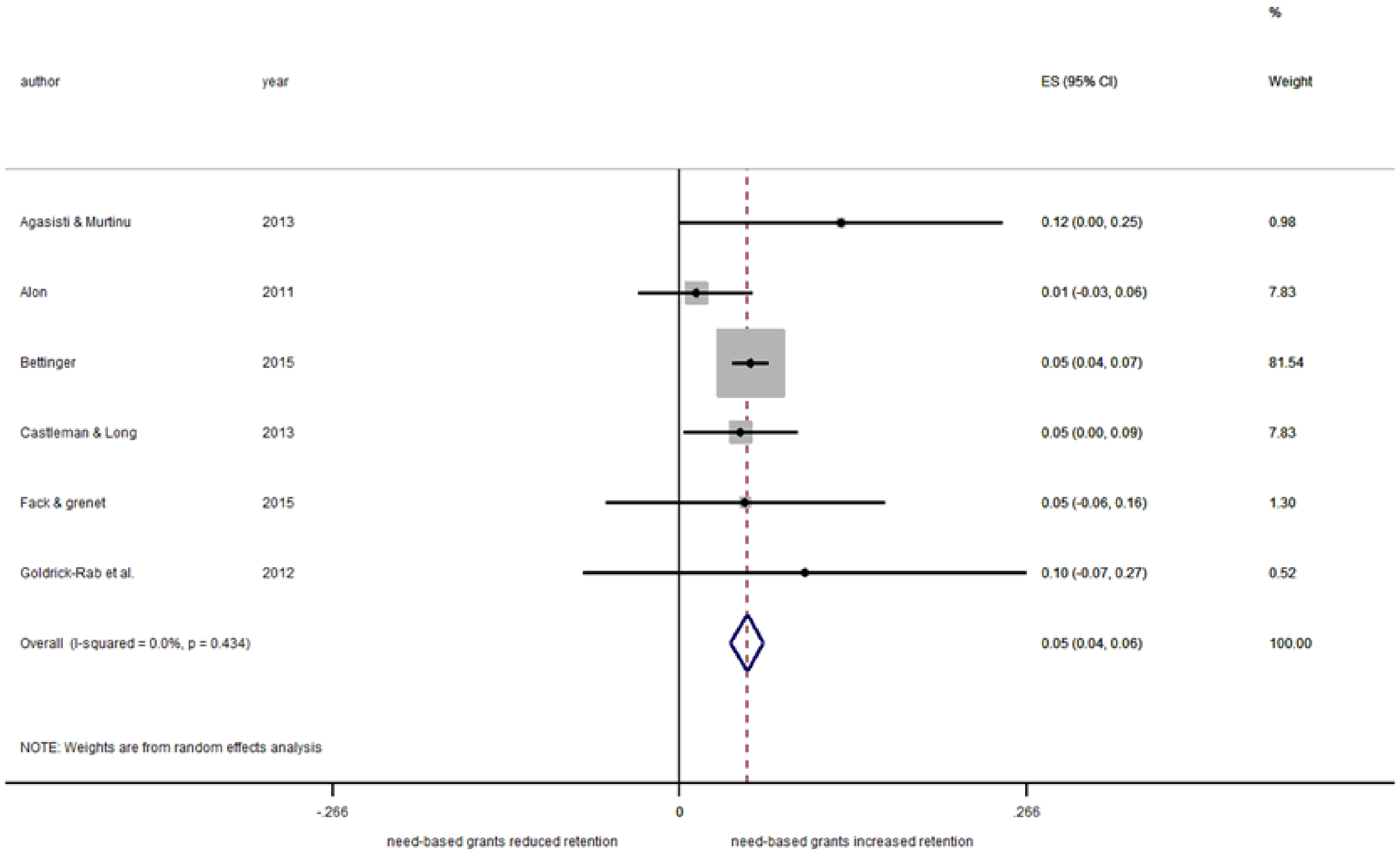

Figure 9: Main effects of need-based grants on retention. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output

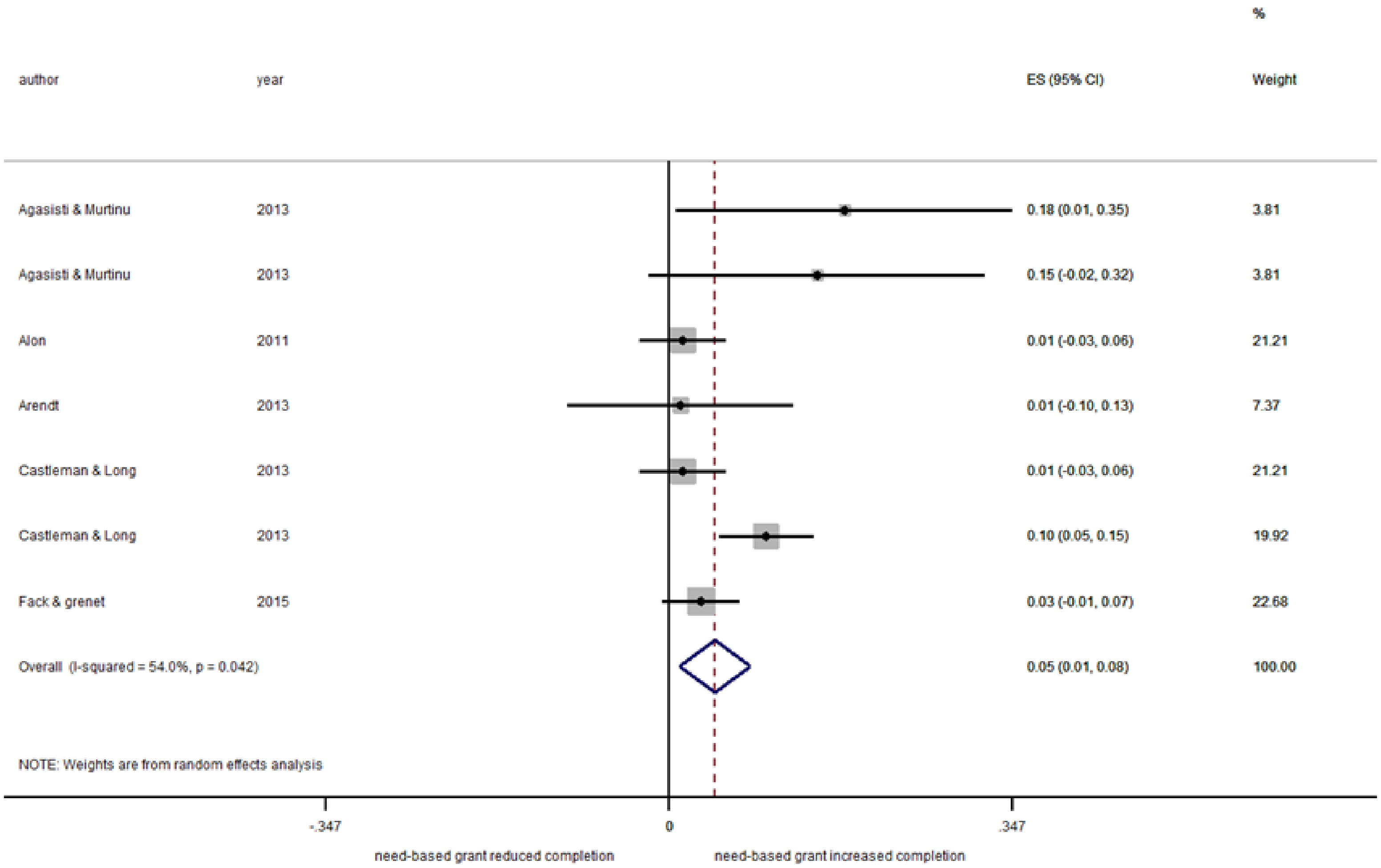

Figure 10: Main effects of need-based grants on graduation. In the forest plot, the area of each box represents the contribution of the corresponding study to the meta-analysis. The center of the box represents the size of the treatment effect. The confidence interval for the treatment effect is shown by the black line. The diamond shows the summary treatment effect with the left and right extremes representing the corresponding confidence interval. Stata 14 is used to produce this output.

.0010). Notice that, when we rerun the analysis only for graduation within the nominal time, the overall effect is somewhat smaller and insignificant (d=.030; 95% CI [-.01; .06]). When we look at the effect of grants on graduation after the nominal study time, we find a bigger effect size (d=.06; 95% CI [-.005; .125]). It seems that the effect of need-based grants increases over time.

3.3. Publication bias

Publication bias is an issue that can lead to biased results in a meta-analysis. Publication bias occurs when not all research concerning mentoring, academic probation and need-based grants has been published or reported. Studies may not be reported or published because of non-significant results, or because the results are not valued properly by journal editors or by other researchers. When these studies report different results from those in the analysis, bias arises.

The majority of the studies used in this review are journal articles (46.15%). Roughly half of the studies are located outside journals, and the possibility exists that we do not have to cope with publication bias. However, we also want to draw attention to two caveats. First of all, most articles in this review sample are in the field of education economics. A tradition in this field is that unpublished articles are often made available online because of the long time that elapses before publication. As a result, some of our "unpublished" papers may be published in the future. Secondly, it is not really clear what unpublished means in the digital area. Indeed, most documents can be easily obtained via internet searches independently of being controlled by commercial publishers.

To assess the possibility of publication bias we visually examine a funnel plot and examine an egger regression test for funnel plot asymmetry (Egger, Smith, Schneider & Minder, 1997). Looking at the funnel plot (with our total review sample of 25 studies and 55 outcomes), we observe that it is fairly symmetric (see Figure 11). This is probably because that few studies have large standard errors. The egger test for funnel plot symmetry also indicates a positive insignificant association between the effect size and standard error (b=.024, p= .19, 95% CI [-.01; .06]). For completeness, we also executed the egger test for each intervention. We only conducted a trim and fill analysis (Duval & Tweedie, 2000) if the egger test indicated publication bias. Note that a general effect size per intervention is not interpretable due to the different outcome variables. First of all, we do not observe a significant association between effect size and the standard error based on the egger test (b=-.073, p= .19, 95% CI [-.19; .04]) for the effect of the AD policy. On the other hand, for faculty-student mentoring, the egger test indicates an asymmetric funnel plot (b=.079, p= .049, 95% CI [.000; .1549]). However, after trimming and filling 12 hypothetical effect sizes, the random effects size is still positive and significant (see Table 4). Finally, we observe a significant positive association between the effect size and the standard error (b=.041, p < .01, 95% CI [.028; .054]) for the effect of need-based grants. Again, the trim and fill analysis makes it clear that after trimming and filling 23 hypothetical effect sizes, the random effect size remains positive and statistically significant (see Table 4). We can conclude that although the egger test of two of the interventions points towards publication bias, this bias is unlikely to have an appreciable effect on the findings.

Table 4: Trim and fill analysis.

Intervention	**Status**	**Pooled Est.**	**Conf. Interval**	**p**	**No. of studies**
Mentoring	regular	.131	[.07;.20]	.00	9
	filled	.103	[.03;.17]	.00	12
Grants	regular	.05	[.04; .05]	.00	19
	filled	.05	[.04;.05]	.00	23

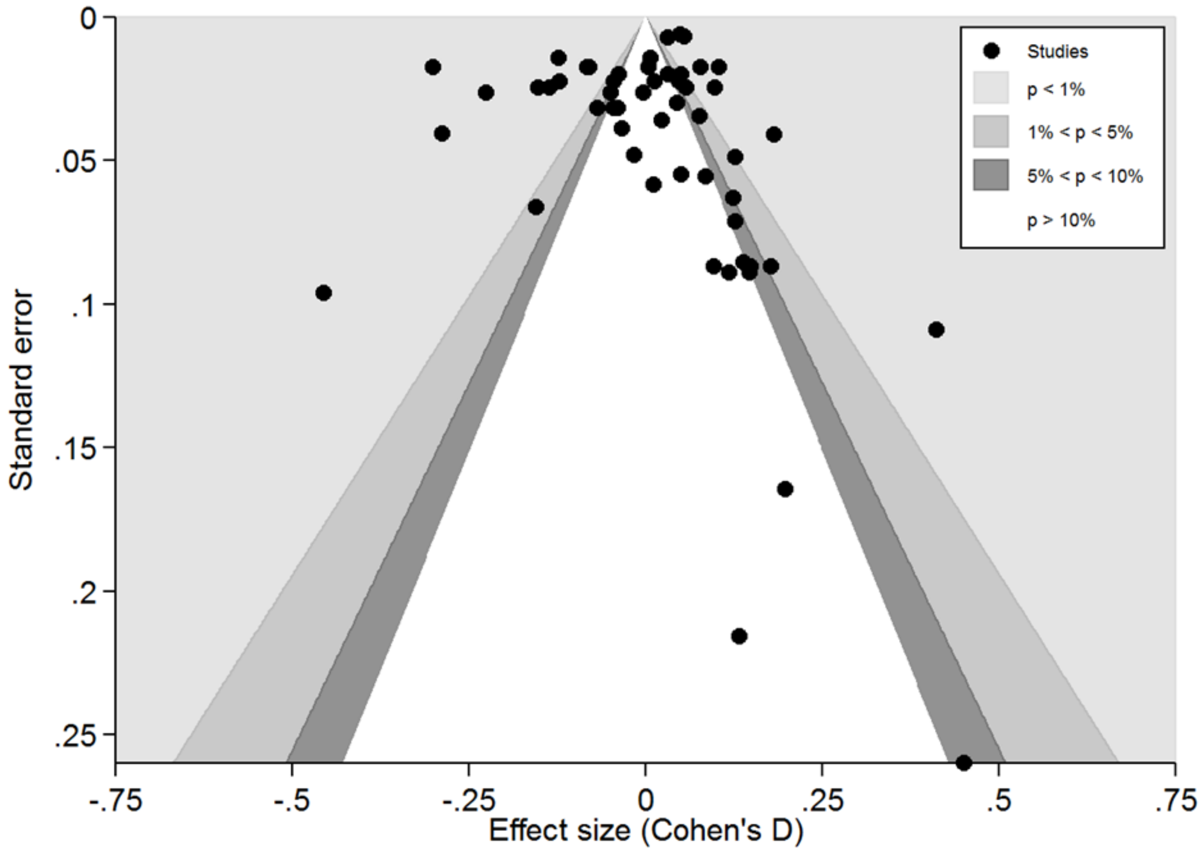

Figure 11: Funnel plot of Standard error by effect size. The center of the funnel plot indicates the fixed-effects summary estimates. The sloping lines indicate the expected 95% confidence interval for a given standard error when assuming that there is no heterogeneity between studies. Stata 14 is used to produce this output.

4. Conclusion

In this study, we have identified 25 studies and 54 effect sizes of experimental, quasi-experimental, and cohort studies. All of these studies investigate the impact of an academic dismissal policy, faculty-student mentoring, or need-based grants on student outcomes. In particular, we studied the impact of these interventions on enrollment, retention, and graduation. Table 5 summarizes the results of our analyses.

In sum, we observe for two of the three interventions positive and significant effects on the outcome variables. The academic dismissal policy has a negative significant effect on retention (d=-.17), no effect on graduation within the nominal study time (d=.00), and a small negative effect on graduation within 6 years (d=- .01). Thus, the AD policy has led to an increase of dropout of 8.4% of the intervention group in comparison to the control group (see Table 5). This indicates that because of an academic dismissal policy, eight extra students (out of a 100 students) drop out compared to the situation where there was no academic dismissal policy. Student-faculty mentoring, on the other hand, seems to have a positive and significant effect on retention (d=.15) and graduation (d=.10). From Table 5, we observe that the average intervention effect of mentoring on retention is equivalent to a 7.5% improvement in retention and a 5% increase in graduation for the intervention groups. This implies that if 100 students are being mentored, 7 students are retained and 5 graduate compared with if they had received the control treatment. This is a more positive finding compared with the results of

Table 5: Summary of average effect sizes for overall intervention effects

Intervention	Outcome	Standardized mean effect (d) N of effect sizes in parentheses	BESD (Percentage Improvement in Treatment over Control)
AD policy	Retention	-.17 (11)	-8.4%
	Graduation within nominal time	.00 (8)	0%
	Graduation within 6 year	-.01 (7)	-.5%
Mentoring	Retention	.15 (7)	7.5%
	Graduation	.10 (2)	5%
Need-based grants	Enrollment	.05 (6)	2.5%
	Retention	.05(6)	2.5%
	Graduation	.05 (7)	2.5%

Hattie (2015) concerning the effect of mentoring (d=.09) on student achievement. Need-based grants also have positive significant effects on enrollment (d=.05), retention (d=.05), and graduation (d=.05). Need-based grants lead to an increase in enrolment, retention, and graduation of 2.5% for the intervention group. This implies that if 100 students receive need-based grants, 2-3 extra students enroll, are retained, and eventually graduate compared with if these students did not receive treatment. These findings are somewhat lower than the results of Hattie (2015), who found an effect size of .23 concerning the effect of finances on student outcomes. Despite the statistical significance of most findings, the average effect of all interventions are perceived as small (d<.40; Cohen, 1988; Hattie, 2015).

This meta-analysis provides some recommendations for policy makers. Even though the effect sizes of the discussed interventions are perceived as small, this does not indicate that they are not valuable. As Glass, McGaw and Smith (1981, p. 104) state "… the issue of whether an effect should be considered 'large' depends on a number of factors. These might include the costs of implementing the intervention, its practicable feasibility, the benefits associated with the difference produced and the value attached to those benefits, as well as the size of other effects produced by comparable interventions in the same context and with the same outcome". In the higher education context, this means that policy makers and practitioners should consider the costs associated with implementing different types of interventions relative to the potential gains they can expect in educational outcomes.

The implementation of an AD policy does not come with many costs. Due to an AD policy, dropout in the first year increases by 8.4%. Based on the assumption that students who dropped out because of the AD policy would also have dropped out in later years, an AD policy can lead to substantial cost savings. The average EU-28 country has a cost per student per year of €9,500 (Eurostat, 2016b). Each year 2.8 million Bachelor students enter higher education in the EU (Eurostat, 2016a). If we can accelerate the dropout decision of first-year students (8.4%) because of the AD policy, this leads to a cost saving of €2,200 million. However, the dismissed students need to be properly guided to a new higher education study and another higher education degree. As Larsen (2000, p3) states: "people without a degree in higher education, despite the fact that they use less time in the educational system, on average spend eight years less on the labor market, because they more often struggle with unemployment and more frequently end up on early retirement or welfare benefits". A financial grant of €1,000 leads to a 2.5% increase in enrolment, retention and graduation. This implies that in the EU-28 countries 70,000 extra students enter and earn a degree. EU governments have to pay €10,500 per student per year (i.e. average expenditure and grant). If the average student graduates within 5 years, this implies a cost of € 3,700 million. However, compared with high school graduates, Bachelor's degree holders have net present earnings of €158,000[10] or €11,000 million for 70,000 students (OECD, 2014) and the real benefits are much higher ,given, for example, the increased tax revenues and higher employability rate. For society, the net present value of the extra graduates comes down to €5,851 million[11] (OECD, 2014). Mentoring leads to an increase of 210,000 first-year Bachelor students, which means an increase of €33,000 million private net present earnings and 17,500 million public net present earnings. Although mentoring is perceived as a costly intervention, the benefits seem high enough to consider its implementation. Given the above cost-benefit analysis, the discussed interventions seem to be beneficial for students, higher educational institutions, and the society as a whole.

Other important considerations would be the scope of the problems of the higher education institutions or academic programs. Some policy makers may want to select the most promising students after the first year, while other practitioners may want to support students. For the former group, academic probation may be a good option, for the latter group need-based grants and mentoring seem to be the best intervention. Finally, practitioners need to take the effect of these interventions on other educational and personal outcomes into account (e.g. student satisfaction).

There are some directions for further research. The studies represented in this report cover literature from 1995 until 2015. Given the fact that we found more eligible studies in the last years (see Figure 4), this review needs to be updated in the near future. Further, academic probation needs to be further exploited. Indeed, at the moment we only found quasi-experimental studies from the US and Canada. Studies from the Netherlands only investigated the effect of academic dismissal policy using cohort research. However, academic probation can still have different effects depending on the notation. Since academic probation is becoming very popular in the Netherlands, a study using a good quasi-experimental design in the Netherlands in necessary. Next, we only exploited mentoring executed by faculty staff or professionals. Peer mentoring can also have an effect on educational outcomes. Future research could investigate this further. Finally, other grants than need-based grants could be researched.

[10] The private net present earnings is the average of the private net present earnings for men and for woman (OECD, 2014).

[11] The public net present earnings is the average of the public net present earning for men and for woman (OECD, 2014).

References

Agasisti, T., & Murtinu, S. (2014). Grants in Italian university: a look at the heterogeneity of their impact on students' performances. *Studies in Higher Education*, 1-27.

Alexander, F. K. (2000). The changing face of accountability: Monitoring and assessing institutional performance in higher education. *Journal of Higher Education*, 411-431.

Alon, S. (2011). Who Benefits Most from Financial Aid? The Heterogeneous Effect of Need-Based Grants on Students' College Persistence*. *Social Science Quarterly, 92*(3), 807-829.

Arendt, J. N. (2013). The effect of public financial aid on dropout from and completion of university education: evidence from a student grant reform. *Empirical Economics, 44*(3), 1545-1562.

Avdic, D., & Gartell, M. (2015). Working while studying? Student aid design and socioeconomic achievement disparities in higher education. *Labour Economics, 33*, 26-40.

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.

Bettinger, E. (2004). *How financial aid affects persistence*. In College choices: The economics of where to go, when to go, and how to pay for it (pp. 207-238). University of Chicago Press.

Bettinger, E., & Baker, R. (2014). The effects of student coaching: An evaluation of randomized experiment in student advising. *Education Evaluation and Policy Analysis, 36*(1), 3--19.

Bettinger, E. (2015). Need-based aid and college persistence the effects of the ohio college opportunity grant. *Educational Evaluation and Policy Analysis, 37*(1 suppl), 102S-119S.

Borenstein, M., Hedges, L. V., & Rothstein, H. (2007). *Introduction to meta-analysis*. Chichester, England: Wiley.

Boruch, R. F. (1997). *Randomized experiments for planning and evaluation: A practical guide*. Thousand Oaks, CA: Corwin

Brock, T., & Richburg-Hayes, L. (2006). Paying for Persistence. Early Results of a Louisiana Scholarship Program for Low-Income Parents Attending Community College. *MDRC,43.*

Campbell, T. A., & Campbell, D. E. (1997). Faculty/student mentor program: Effects on academic performance and retention. *Research in Higher Education, 38*(6), 727-742.

Campbell, T. A., & Campbell, D. E. (2007). Outcomes of mentoring at-risk college students: gender and ethnic matching effects. *Mentoring & Tutoring, 15*(2), 135-148.

Casey, M., Cline, J., Ost, B., & Qureshi, J. (2015). *Academic Probation, Student Performance and Strategic Behavior (working paper).* Retrieved at 07-04-2016, from: https://aefpweb.org/sites/default/files/webform/aefp40/Academic_Probation.pdf

Castleman, B.J., Long, B.T., (2012). *Looking beyond enrolment: the causal effect of need-based grants on college access, persistence, and graduation*. paper presented at the Annual Conference of the Association for Education Finance and Policy: Boston.

Chi, O. L., & Dow, A.W. (2014). *Improving Balance in Regression Discontinuity Design by Matching: Estimating the Effect of Academic Probation After the First Year of College*. Paper presented at SREE 2014 Improving Education Science and Practice:The Role of Replication, Cambridge, Harvard University

Deil-Amen, R., & Rosenbaum, J. E. (2003). The social prerequisites of success: Can college structure reduce the need for social know-how?. *The Annals of the American Academy of Political and Social Science, 586*(1), 120-143.

De Koning, B. B., Loyens, S. M. M., Rikers, R. M. J. P., Smeets, G., & van der Molen, H. T. (2013). 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(ahead-of-print),* 1-13.

Duijndam, F. J., & Scheepers, A. W. A. (2009). Bindend studieadvies [Academic dismissal policy]. *Hoger Onderwijs Management*, 5, 17-19.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot--based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455-463.

Dynarski, S. M. (2003). Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion. *American Economic Review 93*(1), 279-88.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629-634.

European Commission (2014). *Priorities.* Retrieved at 01-04-2016, from: http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/priorities/index_en.htm

European Commission (2015). *Targets.* Retrieved at 01-04-2016, from: http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/targets/index_en.htm

Eurostat (2016a). *Tertiary educational attainment*. Retrieved at 01-04-2016, from: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=t2020_41&plugin=1

Eurostat (2016b). *Total public expenditure on education*. Retrieved at 01-04-2016, from: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tps00158&plugin=1

Fack, G., & Grenet, J. (2013). *Improving college access and success for low-income students: Evidence from a large need-based grant program (working paper).* retrieved at 15-03-2016, from: http://pubs.aeaweb.org/doi/pdfplus/10.1257/app.20130423

Fletcher, J. M., & Tokmouline, M. (2010). *The Effects of Academic Probation on College Success: Lending Students a Hand or Kicking Them While They Are Down*? (Working paper). Retrieved at 04-07-2016, from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.600.115&rep=rep1&type=pdf

Gijbels, D., Van der Rijt, J., & Van de Watering, G. (2004). Het bindend studieadvies in het hoger wetenschappelijk onderwijs: worden de juiste studenten geselecteerd? [The binding study advice in higher scientific education: Are the right students selected?]. *Tijdschrift voor Hoger Onderwijs,* 22, 62-72.

Goldrick-Rab, S. (2010). Challenges and opportunities for improving community college student success. *Review of Educational Research, 80*(3), 437-469.

Goldrick-Rab, S., Douglas N.H., Kelchen, R., and Benson, J. (2012). *Need-Based Financial Aid and College Persistence: Experimental Evidence from Wisconsin.* Madison, WI: Institute for Research on Poverty.

Hattie, J. (2015). The applicability of Visible Learning to higher education.*Scholarship of Teaching and Learning in Psychology, 1*(1), 79. doi: 10.1037/stl0000021

Hoffer, S. B. (2010). *A Formal Mentoring Program and its Relationship to Academic Success and Retention Rates of First-Time Freshman College Students at Walla Walla University (Unpublished dissertation).* Andrews University, Michigan.

Kane, T. J. (2003). *A quasi-experimental estimate of the impact of financial aid on college-going (No. w9703)*. National Bureau of Economic Research.

Kot, F. C. (2014). The impact of centralized advising on first-year academic performance and second-year enrollment behavior. *Research in higher education, 55*(6), 527-563.

Lindo, J. M., Sanders, N. J., & Oreopoulos, P. (2010). Ability, Gender, and Performance Standards: Evidence from Academic Probation. *American economic journal. Applied economics, 2*(2), 95-117.

OECD (2014). Education at a Glance 2014. Retrieved from https://www.oecd.org/edu/Education-at-a-Glance-2014.pdf

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students*. San Francisco Jossey-Bass.

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

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education, 30*(2), 175-194.

Salinitri, G. (2005). The effects of formal mentoring on the retention rates for first-year, low achieving students. *Canadian Journal of Education/Revue canadienne de l'education*, 853-873.

Sandner, M. (2015). The effects of high-quality student mentoring. *Economics Letters, 136*, 227-232.

Schudde, L. & Scott Clayton, J. (2014). *Pell grants as performance-based aid? An examination of satisfactory academic progress requirements in the nation's largest need-based aid program (Working Paper).* Columbia University Academic Commons. Retrieved at 05-05-2016, from: http://dx.doi.org/10.7916/D89K491C.

Swanson, J.E. (2006). *Success in the First Year: Impact of Alternative advising on students of a liberal arts college (Unpublished Dissertation).* Westerm Michigan University, Michigan.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research, 45*(1), 89-125.

Tinto, V. (1987). *Leaving college: rethinking the cause and cures of student attrition*. Chicago: University of Chicago Press.

Van Heerden, A. (2013). *Dropout and study progress in Higher Education*. Hogeschool van Amsterdam -- TIER, mimeo.

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A Regression--Discontinuity Approach. *International Economic Review, 43*(4), 1249-1287.

Van Klaveren, C. P. B. J., de Wolf, I. F., & de Witte, K. (2015). Systematic Reviews in Education Research; when do effect studies provide evidence?.*Contemporary economic perspectives in Education*, 11-34. Retrieved from https://www.researchgate.net/profile/Chris_Klaveren/publication/280717607_Contemporary_Economic_Perspectives_in_Education/links/55d2dd7f08aec1b0429efb9a.pdf

Visher, M., Butcher, K. F., & Cerna, O. S. (2011). *Guiding Math Students to Campus Services: An Impact Evaluation of the Beacon Program at South Texas College*. Society for Research on Educational Effectiveness.

Webber, D. A. (2012). Expenditures and postsecondary graduation: An investigation using individual-level data from the state of Ohio. *Economics of Education Review, 31*(5), 615-618.

Woelders, L., Visser, W., & Rijksbaron, S. (2013). Bindend studieadvies in relatie tot studieresultaat. [Binding study advice in relation to study results]. *OnderwijsInnovatie*, 26-27.

Yorke, M., & Longden, B. (2004). *Retention & Student Success in Higher Education*. Maidenhead: Open University Press’’

Appendix

Author	Year	Country	Level of education	Method	Outcome	ES method used	Treatment analysis N	Control analysis N	Notes
Agasisti & Martinu	2013	Italy	University	Matching	Retention	Mean difference/ Pooled standard deviation	354	869	The effect size is recalculated in order to show the effect of a grant of €1000
					Graduation within the nominal study time	Mean difference/ Pooled standard deviation	166	670	See above.
					Graduation within 6 years	Mean difference/ Pooled standard deviation	166	670	See above.
Alon	2011	VS	Retention	IV	Retention	Unstandardized regression Coefficient	4,037	3,711	The unstandardized regression coefficient is recalculated in order to show the effect of a

									grant of €1000
					Completion	Unstandardized regression Coefficient	3,711	4,029	See above.
Arendt	2013	DK	University	Natural experiment	Graduation	Logit coefficient	618	557	Assume .526/.474 split. Based on summary statistics. We also assumed an average increase of €2170 to transform the logit coefficient so it represented the effect of an €1000 increase in grant.
Arnold	2015	NL	University	Cohort analysis	Retention	Unstandardized regression Coefficient	528	1,923	
					Graduation within the nominal	Unstandardized regression coefficient	525	1,911	

					study time				
Baumgartner & Steiner	2006	DE	College and universities	DiD	Enrollment	Unstandardized regression coefficient	347	451	
Bettinger	2015	VS	University	DiD	Retention	Unstandardized regression coefficient	41,062	42,197	Averaged the coefficient of the subsample of winners and losers. The unstandardized regression coefficient is recalculated in order to show the effect of a grant of €1000
Bettinger & Baker	2014	VS	University	RCT	Retention	Unstandardized regression coefficient	8,049	5,506	Assume .59/.41 split. Based on summary statistics.
					Graduation	Unstandardized regression coefficient	799	546	See above.
Brock & Richburg-	2006	VS	College	RCT	Enrollment	Frequencies [PROBIT]	197/67	191/82	

Hayes									
Campbell & Campbell	2007	VS	University	Matching	Retention	Frequencies [PROBIT}	290/49	250/89	
					Graduation	Frequencies [PROBIT]	198/141	177/162	
Casey et al.	2015	VS	University	RDD	Retention	Unstandardized regression coefficient	1,928	12,906	Assume .13/.87 split. Based on summary statistics.
					Graduation within the nominal study time	Unstandardized regression coefficient	1,125	7,526	See above.
					Graduation within 6 years	Unstandardized regression coefficient	744	4,979	See above.
Castleman & Long	2013	VS	University and college	RDD	Enrollment	Unstandardized regression coefficient	3,628	3,289	Assume .53/.61 split. Based on summary statistics. The unstandardized regression coefficient is recalculated in order to show

									the effect of a grant of €1000
					Retention	Unstandardized regression coefficient	3,591	3,962	See above.
					Graduation within the nominal study time	Unstandardized regression coefficient	4,280	3,881	See above.
					Graduation within 6 years	Unstandardized regression coefficient	3,628	3,289	See above.
Chi & Down	2014	Canada	University	RDD with matched data	Retention	Unstandardized regression coefficient	2,538	3,969	Assume .39/.61 split. Based on summary statistics.
					Graduation within the nominal study time	Unstandardized regression coefficient	1,860	2,909	See above.
					Graduation within 6 years	Unstandardized regression coefficient	1,288	2,015	See above.

Eijsvogels et al.	2015	NL	University	Cohort analysis	Retention	Frequencies [PROBIT]	835/151	607/48	
Fack & Grenet	2015	FR	University	RDD	Enrollment	Unstandardized regression coefficient	11,527	4,940	Assume .70/.30 split. Based on summary statistics. The unstandardized regression coefficient is recalculated in order to show the effect of a grant of €1000.
					Retention	Unstandardized regression coefficient	10,487	4,495	See above.
					Graduation	Unstandardized regression coefficient	7,666	3,285	See above.
Fletcher & Tokmouline	2010	VS	University	RDD	Retention	Unstandardized regression coefficient	University A: 5,182 University B: 6,317	University A: 11,533 University B: 15,025 University C:	Split of University A: .31/.69 University B: .29/.71

							University C: 2,089 University D: 3,621	5,648 University D: 4,780	University C: .27/.73 University D: .43/.57
					Graduation within the nominal study time	Unstandardized regression coefficient	University A: 4,315 University B: 6,135 University C: 1,470 University D: 2,377	University A: 9,604 University B: 15,021 University C: 3,676 University D: 3,150	See above.
					Graduation within 6 years	Unstandardized regression coefficient	University A: 3,610 University B: 5,126 University C: 1,128 University	University A: 8,034 University B: 12,533 University C: 3,049 University D: 2,442	See above.

							D: 1,842		
Goldrick-Rab et al.	2012	VS	University	RCT	Enrollment	Frequencies [PROBIT]	590/10	880/20	The effect size is recalculated in order to show the effect of a grant of €1000.
					Retention	Frequencies [PROBIT]	526/74	778/122	See above.
Kot	2014	VS	University	Matching	retention	Odds Ratio	1,238	1,143	
Hoffer	2010	VS	University	RCT	Retention	T-test	76	73	
					Graduation within the nominal study time	Unstandardized regression coefficient	3,087	5,734	See above.
					Graduation within 6 years	Unstandardized regression coefficient	2,102	3,903	See above.
Nielsen et al.	2008	DE	College and University	DiD	Enrollment	Unstandardized regression coefficient	40,173	41,408	The unstandardized regression coefficient is recalculated in order to show

									the effect of a grant of €1000
Scheepers	2004	NL	University	Cohort Analysis	Retention	Frequencies [PROBIT]	1,542/758	1,261/368	
Scheepers & Duijndam	2009	NL	University	Cohort Analysis	Retention	Frequencies [PROBIT]	6,324/3,258	7,854/2,346	
Seftor & Turner	2002	VS	College	DiD	Enrollment	Unstandardized regression coefficient	57,911	51,901	Average effect of two subsamples. The unstandardized regression coefficient is recalculated in order to show the effect of a grant of €1000
Swanson	2006	VS	University	RCT	Retention	Chi-Square 3x2 design	46/5	47/17	
Visher et al.	2011	VS	University	RCT	Retention	Proportions	1,067	1,098	

TIER WORKING PAPER SERIES
TIER WP 17/04

ISBN 978-94-003-0119-1