Optimal admission to higher education

Karsten Albæk*

June 24, 2016

Abstract

This paper analyses admission decisions when students from different high school tracks apply for admission to university programmes. I derive a criterion that is optimal in the sense that it maximizes the graduation rates of the university programmes. The paper contains an empirical analysis that documents the relevance of theory and illustrates how to apply optimal admission procedures. Indirect gains from optimal admission procedures include the potential for increasing entire cohorts of students' probability of graduating with a higher education degree, thereby increasing the skill level of the work force.

Keywords: Educational economics, college admission, human capital

JEL Classification: I21.

*SFI - The Danish National Centre for Social Research, Herluf Trolles Gade 11, DK-1052 Copenhagen, Email: kal@sfi.dk, tlf. +45 22 473 473.

1 Introduction

High graduation rates and the smooth transition of students through the higher educational system are important elements for attaining a high skill level of the work force. Admission policies to higher education constitute a crucial factor for the allocation of students to different programmes and institutions, for graduation rates, and for incentives in the secondary educational system.

University admission policies are under debate in many countries. For instance, the British debate includes the lack of applicants to university programmes that require a strong background in mathematics and science, because upper-secondary school 'students choose "easier" subjects to get higher grades' (Sunday Times, August 17, 2003, p. 6). One virtue of the admission system that this paper proposes is that it can remove the inconsistency between the rational choices of upper-secondary school students and the aim of the education system of providing students with a good background for obtaining a university degree.

This paper analyses the following issue. Students in upper-secondary school (also denoted 'high school') choose different types of high school curricula (hereafter 'tracks'). These tracks vary with respect to, for example, the amount of mathematics. After high school the students apply to a higher educational institution (hereafter 'college' or 'university') or a programme at a university. If the Grade Point Average (GPA) from high school is sufficiently high, the student is admitted, if not, the student is not admitted. Amongst the admitted students, some graduate from the university programme whilst others do not. The probability of graduating from the programme may depend on the student's choice of track in high school (e.g. students with a strong background in mathematics may more successfully complete certain university programmes than students with less mathematics).

This paper first analyses admission criteria to a programme to which students from different high school tracks apply. The analysis is carried out under the following assumptions. (1) The University bases the admission criteria on the GPA and the student's high school tracks. (2) The probability that a student graduates from the programme increases with GPA and varies between students from different high school tracks. (3) The number of admitted students to the programme is fixed. (4) The goal of the university is to maximize the number of students that graduate from the programme. These assumptions imply that the university will set admission GPA thresholds for each high school track (i.e. students with GPAs above the threshold are admitted, whereas those with GPAs below are not). I prove that the maximum graduation rate is obtained when admission thresholds are set such that the marginal graduation rates of students from the various high school tracks are equal. If the marginal graduation rates are not equal, the university can decrease the number of students from a high school track with a low graduation probability and increase the number of students from a track with a high graduation probability, which implies that the average graduation rate increases. The admission rule is optimal in the sense that it maximizes graduation rates.

The paper contains an empirical analysis that documents the relevance of theory and illustrates the application of the rules for optimal admission. I estimate graduation probabilities non-parametrically as a function of GPA for students from three high school tracks entering four social science programmes at the University of Copenhagen. The data contain information about high school GPA, high school track and the bachelor degree programme to which the student is admitted. The graduation probabilities form the basis for calculating the admission thresholds for each of the three tracks in the four programmes. Furthermore, I calculate the increase in the graduation rate from a change in the admission system from a common GPA threshold for the three high school tracks (the present system) to the optimal admission system, in which the thresholds differ for the three tracks.

After presenting the theory and the empirical analysis, I discuss students' reactions to a transition to an optimal admission system. Two types of potential reactions are relevant. First, the students may behave strategically during the application and admission process, that is, the change in admission system may change the students' choices of programmes and universities. I outline the circumstances under which strategic behaviour of students can be ruled out. Second, I analyse how a transition to optimal admission may change students' choices of subjects in high school. Optimal admission changes the intake of students at the margin towards students from high school tracks that have a documented positive impact on the students' graduation probabilities. As the new system makes admission easier for students from these high school tracks, more students may thus choose the tracks that make admission easier. Students who choose these high school tracks are, per construction, more likely to pass university exams. A change to optimal admission to universities thus has the potential to increase entire cohorts of students' likelihood of graduating with a higher education degree.

Two recent articles provide a framework for relating the topic and the contribution of the present paper to the previous economic literature on higher education. The survey by Altonji et al. (2012) formulates a theory for the choice of first, the high school field of study, then the choice of college major, and finally the choice of occupation in the labour market. This sequential choice model is supposed to be solved backwards, such that exogenous shocks to occupational specific wages and consumption are reflected in choices in the educational system. The elements of the theory include multiple levels of education and curriculum choices, and that the agents update their beliefs about ability and preferences. Altonji et al. (2012) also survey the empirical literature on the impact of choices in high school and college on labour market outcomes and note that most authors take simpler approaches as 'it is a complicated model; to include all the detail we have suggested may be almost impossible in practice' (Altonji et al. (2012) p. 14).

This paper focuses on the part of the model in Altonji et al. (2012) that relates to the transition from high school to tertiary education (university or college). Changes in admission criteria alter the admission chances for high school students from different tracks and may thus have an impact on their choice of high school track. The optimal admission in this paper is derived from maximization of graduation rates at college. Neither in their theoretical model nor in their empirical survey do Altonji et al. (2012) consider college attrition explicitly. Thus the inclusion of attrition in college in a dynamic choice model entails an extension of the (already complicated) model by Altonji et al. (2012).

Moreover, Fu (2014) analyses applications, admission and enrolment in the college market. She assumes that a private college maximizes a payoff function consisting of the weighted sum of the net tuition revenue of the students and the abilities of the admitted students. The number of admitted students is fixed in the model. The framework is an equilibrium game theoretical model that includes the choice of students, and she estimates this model using data for the US college market. Fu (2014) extends earlier contributions on tuition choice and admission criteria (e.g. Epple et al. (2006) and Gary-Bobo and Trannoy (2008)).

None of these contributions consider university attrition rates as a part of an element of relevance for admission criteria to universities. However, under additional assumptions, the criteria function in this paper becomes equivalent to the payoff function for a private college in Fu (2014). Assume that the state provides funding to a public university only for students who pass exams. This assumption implies that maximization of graduation rates for a public university becomes equivalent to maximization of tuition fees for a private college. In Denmark the state funds universities only for students, who pass exams, not for students who do not, thus making the funding element of the criteria function for a Danish public university equivalent to that for an American private college in Fu (2014). The other element in the criteria function in Fu (2014) is the quality of the admitted students. However, if the quality is measured by students' ability to graduate from college, then maximization of the quality of admitted students becomes identical to maximization of graduation probability.

Although neither Altonji et al. (2012) nor Fu (2014) consider college attrition, this problem is an issue in the empirical literature on American higher education. The American literature includes Bound and Turner (2011), who survey the economics literature on college enrolment and completion with a focus on the limited US growth in college enrolment and flat college completion rates over time. The empirical analysis includes issues such as student preparedness, student funding, parental background, college resources and the costs and gains of college education. Bowen et al. (2009) study completion in US public colleges. One of the results of their empirical analysis is that high school grades are much better for predicting college completion than the Scholastic Assessment Test (SAT) scores. Another strand of the American literature deals with the ability of the SAT test to predict college admission (e.g. Manski and Wise 1983) and freshman grades (e.g. Rothstein 2004). In combination with high school grades, US colleges and universities use the SAT score for determining admission. For example, the University of California constructs an admission index consisting of a weighted average of SAT scores and high school grades. If a student from a public secondary school in California has a score above a certain threshold on this index, this student is automatically admitted to the University of California system[1]. In addition to pre-college grades learning about academic performance during college plays a role according to Stinebrickner and Stinebrickner (2014) who assess that 45 per cent of dropout in the first two years of college can be attributed to what students learn about their academic performance. The

[1] See the 'admission index' at University of California, http://admission.universityofcalifornia.edu/.

problem of attrition in the US higher educational system figures prominently in the educational and sociological literature, where the seminal contribution is Tinto (1993).

The interplay between university preferences for students (as expressed in, e.g. admission rules) and student preferences for universities is analysed in the literature on the 'college admission problem' (Gale and Shapley 1962, and Roth and Sotomayor 1990). The contributions in this literature are useful for analysing potential strategic behaviour by students. A separate section of this paper analyses this issue and clarifies how the contribution in this paper relates to the literature on the college admission problem.

The social science programmes in the empirical analysis are three-year structured programmes at the University of Copenhagen leading to a bachelor degree in a specific field (or 'major' in American terminology). As in many other European countries the choice of programme is concurrent with initial enrolment: students are admitted to specific programmes within the university (Bound and Turner 2011, p. 586). The empirical results in this paper may thus be of direct relevance to policy makers in many countries.

The contributions in this paper are applicable at several levels of decision-making: for constructing admission rules for one programme at one university, for joint admission rules for several programmes in one university, and for joint admission criteria for programmes across several universities.

Increased graduation rate from higher education, better preparation in upper-secondary school for enrolment in higher education, and the consequent higher skill level of the work force may not be the only objectives of higher education admission policies. Other goals may include access to higher education for groups of students from a variety of backgrounds. Such goals might or might not be in conflict with the efficiency considerations in this paper.

To the extent that other goals are in conflict with efficiency, this paper establishes a framework for calculating the costs of pursuing these goals in terms of lower graduation rates. However, as Bound and Turner (2011), p. 603 state: 'That college completion is a central outcome of higher education and a critical input for labor-market success and economic growth is not in dispute'.

The paper is organised as follows. Section 2 derives the optimal admission rules that maximise graduation rates in the higher education system. The section also derives an analytical expression for the approximate increase in graduation rates. Section 3 contains a breif description of the Danish educational system. Section 4 presents the data. Section 5 demonstrates how to estimate the parameters of interest and use them for construction of optimal admission rules. Section 6 establishes the relation between optimal admission procedures and the issues considered in the literature on the college admission problem. Section 7 traces the impact of a change from non-optimal to optimal admission procedures on the choice of subjects by students in secondary school. Section 8 analyses how this behavioural change alters the observed graduation rates for groups of secondary school students. The impact on graduation rates depends on the extent to which differences in graduation rates are due to self-selection or to a heterogeneous impact of subjects across students. Section 9 concludes.

2 Maximisation of graduation rates

This section contains the formal analysis of admission to higher education viewed as an optimising problem. I first outline the framework and then the content of the section.

I assume that the authority that decides admission criteria to programmes at a university has a goal that can be stated as an objective function. I further assume that the goal of the authority is to maximise the graduation rates of those programmes.

As mentioned in the introduction, this objective function is equivalent to the objective function in Fu (2014) if the university is funded for students who pass their exams at the university, but not for students who do not pass. Furthermore, I assume that the number of students admitted to each programme is fixed, which makes the assumption equivalent to the 'fixed capacity' assumption in Fu (2014).

First, this section proves the admission rule that is the result of the maximization problem, the 'optimal admission rule'. Second, the section calculates how much optimal admission changes admission thresholds compared to a system with a common threshold (i.e. students with GPAs above these thresholds are admitted and students with GPAs below are not). Third, the section develops an expression for the increase in graduation rates that is the result of introducing optimal admission rules.

As grades in Denmark are comparable across high schools, GPAs are used as a measure of general ability to study at universities. The admission authority has information about the distribution of the applicants to each programme with respect to their high school tracks and high school GPAs.

Students can choose several tracks or types of school in the secondary school system before applying to higher education. Secondary school tracks are denoted i, $i = 1, 2, ..., n$. The threshold GPA for track i is denoted g_i, and the total number of admitted students from track i becomes

$$\Lambda_i (g_i) = A_i \int_{g_i}^{g^{\max}} f_i (k) \, dk, \tag{1}$$

where $g^{\max}$ is the maximum grade, $f_i (k)$ is the density function for the number of applicants with a GPA of k, and A_i is the total number of applicants with a secondary school certificate from track i.[2]

This section takes the choice of students as given, implying that the function, $f_i (k)$, is taken as given. Sections 6, 7 and 8 in the paper analyse the choice of students given the change to optimal admission.

Graduation is denoted by y, which takes the value 1 for graduating and 0 for not graduating. The probability of graduating for a student from track i with GPA k is denoted by $p_i (k)$, that is

$$p_i (k) = P (y = 1 | k, \text{path } i) = E (y | k, \text{path } i) .$$

This section assumes that $p_i (k)$ increases in k, $\partial p_i (k) / \partial k > 0$. The case when higher grades in secondary school do not increase graduation probability is simple and is covered by an example in the empirical section.

The expected number of students graduating from the programme becomes

$$K_i (g_i) = A_i \int_{g_i}^{g^{\max}} p_i (k) \; f_i (k) \, dk.$$

The number of students finishing the programme decreases in the admission threshold, whilst $K_i (g_i) / \Lambda_i (g_i)$, the share of students graduating from the programme, increases in the admission threshold.

[2] The interpretation of the expression is straightforward if students apply to only one university programme. However, students often apply to more than one programme. In this case the function is the density of students after the process has taken place that allocates students to different programmes. See the next section for a description of this allocation process in Denmark.

The goal is to maximise the number of students graduating from the programme, $\sum_{i=1}^{n} K_i(g_i)$, given a fixed number admissions

$$\Lambda = \sum_{i=1}^{n} \Lambda_i(g_i).$$

The number of admitted students, Λ, is exogenously given.

The solution must entail that the admission thresholds do not exceed the maximum grade

$$g_i \leq g^{\max}, \; i = 1, 2, ..., n.$$

The Kuhn-Tucker stationarity conditions for maximisation become

$$\begin{aligned} \frac{\partial}{\partial g_i}\left[\sum_{i=1}^{n} K_i(g_i)\right] + \lambda \frac{\partial}{\partial g_i}\left[\sum_{i=1}^{n} \Lambda_i(g_i)\right] - \mu_i &= \\ -A_i p_i(g_i) \; f_i(g_i) + \lambda A_i f_i(g_i) - \mu_i &= 0, \; i = 1, 2, ..., n. \end{aligned} \tag{2}$$

The complementary slackness conditions are

$$\mu_i (g_i - g^{\max}) = 0, \; i = 1, 2, ..., n.$$

If $g_i = g^{\max}$, students from group i are not admitted to the programme. Interior solutions, $g_i < g^{\max}$, imply $\mu_i = 0$ for the admitted groups. Hence, for two groups admitted to the programme, group i and group j, a reformulation of (2) gives

$$\lambda = p_i(g_i) \; = p_j(g_j). \tag{3}$$

This is the rule for optimal admission to higher education. Threshold GPAs for groups of students should be set to equalise marginal graduation probabilities. The expected number of graduating students is maximised when the last admitted student from each of the groups has the same probability of graduating from the programme. The threshold values neither depend on the number of applicants, A_i, nor on the distribution of students according to grades in secondary school, $f_i(g_i)$.

Figure 1 contains a stylised illustration of the optimal admission rule in the case of two groups of students, group i and group j. The conditional graduation probability is assumed to be linear, with the same slope for both groups.

Figure 1 around here

The admission system with a common threshold GPA value, $\overline{g}$, is illustrated by the dotted vertical line, denoted 'Not optimal'. The difference in graduation probability between the two groups corresponds to the distance $p_i(\overline{g}) - p_j(\overline{g})$. Transition to the optimal admission system involves a move from $p_i(\overline{g})$ to $p_i(g_i)$ for group i students and a move from $p_j(\overline{g})$ to $p_j(g_j)$ for group j students. On the horizontal dotted line, the marginal graduation probability is the same for the two groups, $p_i(g_i) = p_j(g_j)$. This admission rule is optimal in the sense that no further improvement in the aggregate graduation rate is possible for a given number of admitted students. Optimal admission to higher education is horizontal, not vertical.

The second step in the analysis is to obtain the differences in admission thresholds in terms of the parameters of the problem, the slopes of the conditional graduation functions, and the difference in initial graduation probabilities. A Taylor expansion of the conditional graduation probability function around $\overline{g}$ gives

$$p_i(g_i) = p_i(\overline{g}) + p_i'(\overline{g}_i)(g_i - \overline{g}), \quad i = 1, 2, ..., n, \tag{4}$$

where $p_i'(\overline{g}_i)$ is the slope of the conditional graduation probability function for students from track i evaluated at GPA level $\overline{g}_i$, which lies between g_i and $\overline{g}$. Furthermore, assume a constant slope in the relevant range of grades, $p_i'(\overline{g}_i) = p_i'$. For another group of students, j, assume a constant slope in graduation probability, $p_j'(\overline{g}_j) = p_j'$, which differs from the slope of group i by a constant, $p_j' = p_i' + \pi_{ji}$.

These assumptions about the slopes for group i and j, in combination with the optimality condition (3), yield the difference in admission thresholds

$$g_j - g_i = \left[(p_i(\overline{g}) - p_j(\overline{g})) - (g_j - \overline{g})\pi_{ji}\right]/p_i'. \tag{5}$$

The first part of the right-hand side, $(p_i(\overline{g}) - p_j(\overline{g}))/p_i'$, is the difference in admission thresholds if group i and group j have the common slope p_i'. If the graduation probability for group i is higher than that for group j, evaluated at the common threshold $\overline{g}$, $p_i(\overline{g}) > p_j(\overline{g})$, the threshold GPA for group j has to be higher than that for group i, $g_j > g_i$. The larger the difference in graduation between the two groups at the common threshold, the larger the difference in the threshold GPAs will be. When the denominator p_i' is large, that is, when a large difference exists in graduation rates between students with high and low GPAs, a small difference between the threshold GPAs amongst the two groups will equalise the marginal graduation rates.

The second part of the right-hand side of (5), $(g_j - \overline{g})\, \pi_{ji}/p_i'$, adjusts the difference in threshold GPAs for the impact of different slopes in conditional graduation probability. If group j has a larger slope than group i, $\pi_{ji} > 0$, the difference in threshold GPAs, $g_j - g_i$, is reduced, whilst $\pi_{ji} < 0$ implies a larger difference in threshold GPAs relative to the case with a common slope.

In the case of a common slope, I have obtained the difference in admission grades between groups of students but not the change from the common threshold $g_i - \overline{g}$. As Figure 1 illustrates, I have determined the magnitude of the horizontal line from $p_i\,(g_i)$ to $p_j\,(g_j)$ but not the height or location of this line segment. I now consider how to determine this location that depends on the relative number of students in the groups.

When an admission threshold is lowered, the number of new students depends on the distribution of applicants, which is an unknown. The following calculations apply the simplifying assumption that the densities are constant in the relevant ranges of GPAs, $f_i\,(k) = f_i$ (amounting to a zero order approximation of the true, unknown distribution by a local uniform distribution function)[3]. From (1) I obtain the number of admitted students before and after the change in the admission threshold, and the change in the number of admitted students becomes

$$\Delta\Lambda_i\,(g_i) = -A_i f_i \Delta g_i, \quad \Delta g_i = g_i - \overline{g}.$$

As $\sum_{i=1}^{n} \Delta\Lambda_i\,(g_i) = 0$, the change in the admission threshold for group j becomes

$$\Delta g_j = g_j - \overline{g} = \sum_{i=1}^{n} s_i\,(g_j - g_i)\,, \tag{6}$$

where

$$s_i = A_i f_i / \sum_{i=1}^{n} A_i f_i, \quad \sum_{i=1}^{n} s_i = 1.$$

[3]It is straightforward to adjust the exposition in the following by replacing the assumption of a local uniform distribution with alternative distributions. However, such an amendment would make the exposition substantially more involved, with limited gain in insight.

As f_i is the density of applicants and A_i the number of applicants, s_i is thus the share of group i students amongst the applicants.

Expression (6) shows that the admission threshold for each group of students can be calculated from the share of students of the different groups and the differences in admission thresholds from expression (5). In the case of common slopes, the change in the admission threshold appears directly on the left-hand side of (6). When the slopes differ for groups of students, expression (5) inserted into (6) yields an equation with the solution

$$\Delta g_j = g_j - \overline{g} = \sum_{i=1}^{n} s_i \left\{ \left[p_i\left(\overline{g}\right) - p_j\left(\overline{g}\right) \right] / p_i' \right\} / \left(1 + \sum_{i=1}^{n} s_i \pi_{ji} / p_i' \right). \tag{7}$$

In Figure 1 a large share of group i students will result in a large increase in the threshold GPA for group j students, and, correspondingly, a small decrease in the threshold GPA for group i students. The horizontal distance from g_i to g_j in Figure 1 is divided into two line segments corresponding to the share of the two groups in the pool of applicants. This division determines the height of the dotted horizontal line denoted 'Optimal'.

The final step is to assess the magnitude of the gain in graduation rates by applying the optimal admission rule. I deduct an analytical expression for the approximate gain under the previously applied simplifying assumption of constant densities, $f_i\left(k\right) = f_i$, as in (6) and constant and identical slopes of the graduation probability functions, $p_i' = p_j' = p'$, (when the slopes differ, a sensitivity analysis is performed by inserting alternative values). Furthermore, assume without loss of generality that the grade distribution is centred at 0, that is, $\overline{g} = 0$.

With these assumptions the change in the number of group i students graduating from the programme becomes

$$\Delta K_i\left(g_i\right) = A_i f_i \int_{g_i}^{0} p_i\left(k\right)\ dk,$$

which in normalised form becomes

$$\Delta\kappa_i\left(g_i\right) = \Delta K_i\left(g_i\right) / \sum_{i=1}^{n} A_i f_i = s_i \int_{g_i}^{0} p_i\left(k\right)\ dk. \tag{8}$$

As the number of admitted students does not change, the introduction of optimal admission rules implies an increase in the graduation rate K/Λ by

$$\Delta K/\Lambda = \left(\sum_{i=1}^{n} A_i f_i/\Lambda \right) \sum_{i=1}^{n} \Delta\kappa_i \left(g_i \right). \tag{9}$$

The increase in the graduation rate is thus proportional to the sum of the normalised changes in graduation, $\sum_{i=1}^{n} \Delta\kappa_i \left(g_i \right)$, where the scaling factor is $\sum_{i=1}^{n} A_i f_i/\Lambda$.

Inserting (4) in (8) gives

$$\Delta\kappa_i \left(g_i \right) = -s_i p_i \left(0 \right) g_i - \frac{1}{2} s_i p' g_i^2. \tag{10}$$

A decrease in the threshold, $g_i < 0$, implies that more students from the group enter and graduate from the programme. The first term on the right-hand side is positive and corresponds to a graduation probability for all new admitted students of $p_i \left(0 \right)$. The second term is negative, taking into account that the graduation probability for the marginal student decreases when the threshold is lowered. In Figure 1 the first term corresponds to a rectangle with base $g_i \overline{g}$ and height $p_i \left(\overline{g} \right)$, whilst the second term corresponds to a triangle with base $g_i \overline{g}$ and height $p' g_i \overline{g}$ (situated at the location of the downward sloping arrow). When fewer students are admitted, $g_i > 0$, the interpretation of (10) is analogous.

Adding (10) over the different groups of applicants and applying (6) yields the following expression for the change in graduation for the programme

$$\sum_{i=1}^{n} \Delta\kappa_i \left(g_i \right) = -\sum_{i=1}^{n} s_i p_i \left(0 \right) \sum_{j=1}^{n} s_j \left(g_i - g_j \right) - \frac{1}{2} p' \sum_{i=1}^{n} s_i g_i^2. \tag{11}$$

The first term on the right-hand side of (11) is the sum of the effects when students affected by the change in admission thresholds are assumed to have graduation probabilities equal to the marginal students admitted at the common threshold. The second term is the sum of the terms that correct for the fact that not all students in the groups that experience either increases or decreases in size have graduation probabilities equal to the marginal students admitted at the common threshold.

Inserting (5) into (11) yields the expression for the increase in the graduation rate for the programme in terms of the parameters of the problem.

The case of $n = 3$ will suffice for both the purpose of analytical insight and the use of this insight in applications ($n = 3$ corresponds to the number of groups in the empirical section that follows).

The result for the gain in the graduation rate for the programme is

$$\sum_{i=1}^{3} \Delta\kappa_i (g_i) = \left\{ \left[s_1 s_2 (s_1 + s_2) d_{21}^2 + s_2 s_3 (s_2 + s_3) d_{32}^2 + s_1 s_3 (s_1 + s_3) d_{31}^2 \right] / 2p' \right\}$$
$$+ s_1 s_2 s_3 (d_{21} d_{23} + d_{32} d_{31} + d_{31} d_{21}) / p', \tag{12}$$

where $d_{21} = p_2(0) - p_1(0)$, $d_{32} = p_3(0) - p_2(0)$ and $d_{31} = p_3(0) - p_1(0)$.

The gain in the graduation rate is high when the initial differences in graduation rates between the groups are large (large numerical values of $p_2(0) - p_1(0)$, $p_3(0) - p_1(0)$ and $p_3(0) - p_2(0)$). A large increase in the graduation rate as GPAs increase (high p') implies a small gain in the aggregate graduation rate, whilst a small value of this slope implies that many students with low graduation rates are replaced by students with high graduation rates. Reshuffling between two groups gives the highest gain in the graduation rate when the two groups are of equal size (e.g. $s_1 = s_2$), as equal group size implies that many students with low graduation rates are replaced by many students with high graduation rates.

Expression (12) for the gain in the aggregate graduation rate applies to one programme. Programmes will differ in increases in graduation rates according to the variation in (1) differences in initial graduation rates between various groups of students, (2) the derivative of the graduation probability as a function of GPA, and (3) the composition of the intake of students from different groups.

3 Institutional setting

The following brief description of the Danish education system provides the background for understanding the empirical results in the paper. The description includes the upper secondary school and the bachelor level higher educational systems.

The Danish upper secondary school system is essentially a two-tiered system in which about 40% enter apprenticeship training,[4] whilst the majority of

[4]Similar to that in Germany, see Albæk (2009) for an overview of the Danish appren-

the remaining 60% enter upper secondary schools ('gymnasium'), which are a prerequisite for admission to higher education. Upon entrance to the upper secondary school, students choose between different tracks. The data for the analysis in this paper stems from a period in which students in the major upper secondary schools (the 'general' high school) had to choose between the 'mathematics track' (with emphasis on mathematics and science) and the 'language track' (with emphasis on languages). In addition to these two tracks separate upper secondary schools offer courses with more specialised curricula, which also qualify students for admission to higher education.

The Danish Ministry of Education attempts to make grades comparable across secondary schools. Written exams set by the Ministry are mandatory in all secondary schools, and grading is monitored by the Ministry. Oral exams are conducted by both the teacher and an external moderator. Thus the expectation is that the GPAs from Danish high schools are better predictors of graduation than those from the more fragmented US high school system (even though US GPAs nevertheless are good predictors for college completion (see Bowen et al. (2009, ch. 6)).

Upon successful completion of upper secondary school, students may be eligible to enter university, the first step of which is typically a bachelor degree. A bachelor student is expected to graduate after three years of study. The success criteria in this paper is graduation within four years. A *central* national admission office receives data for every programme at every college and university in the country. These data include the number of students to be admitted and the prerequisites that the students have to fulfil for admission to each particular programme (e.g. successful completion of certain high school subjects). Each high school student submits a ranked preference list of desired programmes to the central admission office (students cannot apply directly to Danish colleges or universities).

The central admission office ranks all applicants to the programmes according to their high school GPA. The student with the highest GPA is admitted to the top programme on her preference list, the student with the next highest GPA is also allocated to the top programme on her preference list and so forth. At some stage the allocation process gets to a point, where there are no more places available in the programme of the next students' top priority, and she is admitted to her second priority (if no places are available on this programme, the algorithm continues to her third priority). Then the

ticeship system.

next student is allocated to a programme and so forth until the allocation process stops, either from lack of students or from lack of education places in the programmes. The central admission office then notifies each student as to which programme and university they have been admitted and notifies the universities which students have been admitted to which programmes.

Students not admitted to the programmes for which they have applied are referred to a list of programmes with vacant places. Such programmes are characterised by good employment prospects for graduates in the present Danish labour market, such as engineering and primary school teaching. This indicates that the Danish government cares about the numbers of students in higher education and that a goal of the government is to increase enrolment in programmes from which graduates have good labour market prospects.

Other government goals with respect to higher education can partly be inferred from the funding of higher education. All funding is provided from general revenue, as students do not pay tuition fees. On the contrary, university students receive a stipend while studying (the amount was 9,500 Euros in 2015). In addition, students are entitled to receive government loans, to be repaid after graduation (the amount was 4,800 Euros in 2015). These figures indicate that the Danish government also cares about equality of access to higher education.

However, the stipend payments cease if the student does not pass exams within a specified time (each passed exam yields some points and each programme has a number of points that must be obtained each semester). This indicates that the Danish government cares about efficiency in higher education.

As noted in the introduction, the government only pays higher educational institutions for students who pass their exams, whilst no funding is provided for students who do not pass. This funding scheme also indicates that the Danish government cares about efficiency in higher education.

4 Data

This section describes the data that I use for illustrating the optimal admission rules that were derived in section two of the paper. The application of the data follows in the next section.

I use data for students admitted to the four largest bachelor programmes in social sciences at the University of Copenhagen. The data contain all 10,418 students admitted to these four programmes between summer 1984 and summer 2001, and contain information about high school GPAs, high school tracks and the bachelor programme to which the student is admitted.[5]

As Table 1 shows, the largest share (46%) of university students in the four social sciences programmes come from the mathematical track of the upper secondary school, whilst 25% come from the language track. The remaining 30% come from various types of specialised uppper secondary schools. For the present purpose this last group of students is categorised as 'other'.

Table 1 around here

Table 1 shows that the average graduation rate for all students in the social sciences was 55%. This rate is higher than the figures for the US in Bound, Lovenheim and Turner (2010). Students who do not graduate consist of both dropouts from higher education and those who transfer to another programme.

Table 1 ranks the programmes according to total number enrolled. The law programme with 47% of the students is the largest, whilst the psychology programme with 12% is the smallest.

On average, students from the general secondary schools have higher graduation rates than students from the 'other' group: the difference is a significant 21 percentage points. In total no significant difference in graduation rates appears between the mathematics and the language groups. However, when significant differences exist between the mathematics and the language groups within the four individual programmes, the mathematics students have higher graduation rates. The low aggregate graduation rate for mathematics students is partly a consequence of the low graduation rate in the economics programme, which attracts a substantial share of all mathematics students.

Figure 2 shows the high school GPA distribution for the students in the four university programmes: the top panel displays the distribution of the three secondary school tracks for each of the programmes.[6] The main in-

[5] The data stem from the admission office at the University of Copenhagen.

[6] The secondary school grades of the students in the data are standardised by subtracting the mean of the grades and dividing by the standard deviation for the population

ference drawn is that only minor differences appear in grade distribution amongst the three groups in any of the programmes.

Figure 2 around here

In the bottom panel of Figure 2, the three groups are combined, showing that economics students have lower grades than those in the law programme; the law students in turn have lower grades than those in political science and psychology. Most students are admitted to the various programmes solely on the basis of their secondary school GPAs. The threshold GPAs varied between 0 for the economics programme to nearly 2 for psychology. As the grade distribution for the population of Danish secondary school students is approximately normal, only a tiny fraction of secondary school students is eligible for social sciences programmes other than economics and law. However, a minority of varying magnitude (10-30%) are admitted on the basis of supplementary criteria (e.g. work experience, volunteer work). These admissions imply that data exists not only for students with GPAs above the acceptance threshold but also for those with GPAs below it.

5 Empirical analysis of optimal admission

This section illustrates the application of the rules for optimal admission developed in section two. I illustrate the application using the data for the four social science bachelor programmes at the University of Copenhagen, described in the previous section. The data are used for deriving admission rules that maximize the graduation probability for students admitted to these four programmes. Furthermore, I calculate the resulting increase in graduation probability for the programmes.

According to the optimal admission rule in section two, expression (5), two pieces of information are sufficient for calculating the difference in GPAs between different groups in the case of a common slope: (a) the difference

of Danish secondary school students. The grading scale in the sampling period was the '13-scale', which ranges from 0 to 13. The average GPA for all students in the common high school was approximately 8, and the standard deviation was about 1.

in graduation probability at the common threshold and (b) the slope of the graduation probability with respect to grades. Furthermore, to identify the location of the differentiated thresholds according to expression (6), we need information about (c), the local share of the groups of students. According to expression (12), this information is also sufficient for calculating the approximate gain in graduation rates.

The empirical analysis thus demands not only estimation of the graduation probability function conditional on GPA, which is likely to exhibit non-linearities, but also the slope of the conditional graduation probability function. In contrast to many standard non-parametric regression methods, local polynomial regressions enable slope estimation and I thus choose this method to identify the position and the slope of the conditional graduation probability functions.

A brief exposition of the procedure is as follows. The graduation probability for student s is denoted y_s, which takes the value 1 for graduating and 0 for not graduating. The independent variable, the student GPA, is denoted x_s. Following Wand and Jones (1995, p. 118), the conditional expectation $E\left[y_s|x_s = x_0\right]$ from polynomial regression can be computed from the solution to the minimization problem

$$\min_{a,b_1,b_2} \sum_{s=1}^{N} \left(y_s - a - b_1\left(x_s - x_o\right) - b_2\left(x_s - x_o\right)^2\right) K\left(\frac{x_s - x_o}{h_N}\right)$$

where $K\left(\right)$ is a kernel function and h_N is a bandwidth that converges to zero as $N \to \infty$. The degree of the polynomial must be at least one in order to take the derivative and estimate the slope of the conditional expectation function. The conditional expectation functions estimated from a polynomial of degree two exhibit slightly more curvature than the functions estimated from a local linear estimator, and I set the degree of the polynomial to two. The kernels are Gaussian. The bandwidth is set at 0.62 for the largest group of students, mathematics students in the law programme. The bandwidths for the remaining groups are increased according to the asymptotics of the plug-in bandwidth, which minimise the mean square error of the regression[7].

[7]For a programme with sample size n the bandwith becomes $h_0\left(n_0/n\right)^{2/9}$, where h_0 is the bandwith for the law programme and n_0 is the sample size for the law programme. This expression is obtained from the formulas in Wand and Jones (1995, p. 139).

All data points enter the estimations, including the tails of the grade distribution, where data are slim and inference imprecise. I show the results for GPAs from -1 to 2.5, the relevant range for admission policy.

Figure 3 presents the estimates of the expected probability of graduating, conditional on secondary school GPA. The figure shows the expected graduation probability for the three secondary school tracks for each of the four programmes. The expected graduation probabilities are calculated from the local polynomial regression functions. The conditional expectation functions are highly nonlinear, demonstrating that non-parametric estimation is adequate.

Figure 3 around here

In the three largest programmes – law, economics and political science – the graduation probability increases with secondary school GPA for most tracks of students (the exceptions are mathematics students in economics with very high GPAs and other students in political science with GPAs above 1). In psychology, graduation probabilities are constant for mathematics students with GPAs above 0 and for language students with GPAs above 1.[8]

Optimal admission rules appear in Figure 3 as horizontal lines. For the law programme, the marginal graduation probability for the groups is set at 0.60. The graduation probability for other students is below the horizontal line for all levels of GPA. Consequently, this group should not be admitted to the law programme. The height of the horizontal line is found as follows: when the group of other students does not enter the programme, the intake of mathematics and language students increases to ensure an unaltered number of students in the programme. This increase takes place through lowering the common threshold for mathematics and language students to 0.5.

[8] If no students below the common admission thresholds were admitted, the conditional graduation functions would be truncated at the common threshold. In such a case extrapolation of the conditional graduation functions below the threshold is neccessary for constructing optimal admission thresholds. Graduation probabilities below the common thresholds in Figure 3 are not estimates of population parameters but are contingent upon the policy determining admission of students with GPAs below the common threshold. To the extent that institutions are able to identify students with a high probability of graduating, contingent on their GPA, the graduation probabilities below the common threshold in Figure 3 are higher than the average amongst the applicants.

The values of the graduation probabilities in Figure 3 appear in Table 1: 0.649 for mathematics students and 0.586 for language students. The slopes of the curves in Figure 3 appear in Figure 4, and for the law programme the slopes at GPA level 0.5 also appear in Table 1: 0.090 for the mathematics students and 0.040 for the language students. In expression (7) I insert the information about the slopes, the shares of the two groups and the difference in graduation probability, thereby obtaining the threshold changes that appear in Table 1: a decrease in the threshold for mathematics students of 0.5 and an increase for language students of 0.5.[9] The graduation probability is 0.6 for both mathematics students with a GPA level of 0 and for language students with a GPA level of 1 according to Figure 3. The optimal admission rule for the law programme is thus shown as a horizontal line with the height of 0.6.

Figure 4 around here

The gain in the graduation rate for the law programme is assessed in three steps: an increase when the university admits only the previous number of students from the general secondary school (5.9 percentage points), a decrease when the admission threshold is lowered and more general secondary school students are admitted (-0.3 percentage points), and an increase in the graduation rate when a differentiated threshold equalises the marginal graduation rate between mathematics and language students (0.7 percentage points).[10] The result is the increase in graduation rate by 6.2 percentage points shown in Table 1, corresponding to an increase of 10.4% in the number of graduating law students (0.065 divided by the previous average graduation rate of 0.598).

For the economics programme, all three groups of students are admitted after the change from a common threshold to the optimal admission rule. The common threshold is 0.0, and for each of the three groups, Table 1 lists the

[9]The top panel of Figure 2 shows that the density functions are approximately identical for the three groups of students. I assume that the (unobserved) distributions of the applicants are also identical and that the shares of students in Table 1 apply in the calculations.

[10]The gain from implementing a differentiated threshold is obtained from (9) and (12). On the basis of simulations, I assess the correction factor in (9) to be 0.5 for the programmes analysed in this paper.

estimated graduation probabilities in Figure 3 and the corresponding slopes of the conditional graduation functions in Figure 4. Inserting these values into (7) gives the change in admission threshold listed in Table 1: -0.2 for mathematics students, 0.1 for language students and 0.6 for other students.

The optimal difference in admission thresholds between mathematics and language students is thus 0.3 for the economics programme, which is smaller than the threshold difference of 1.0 between mathematics and language students for the law programme. It should thus be easier for language students to be admitted to the economics programme than to the law programme. This finding is valid despite language students in the relevant ranges of GPAs having more difficulty graduating in economics than in law (as the vertical distance between the conditional graduation functions in Figure 3 makes clear).[11] The reason is that the slopes of the conditional graduation functions are substantially higher in the economics programme than in the law programme in the relevant GPA ranges, as both Figure 4 and the values of the slopes at the common thresholds in Table 1 make clear.

For the economics programme the gain in the aggregate graduation rate from adopting the optimal admission rule is modest. Application of the approximation (12) yields an increase in the aggregate graduation rate of 0.7 percentage points when the slope for the other group enters the calculation (a sensitivity check yields 0.5 percentage points when inserting the larger slope for mathematics group, whilst the smaller slope for the language group gives 1.0 percentage point). The main reason for the small gain is the high association between grades and graduation probability in the economics programme, implying a small reallocation between the groups before the graduation rates for the marginal students are equalised. As the initial graduation rate is 0.401 for the economics programme, the expected increase in the number of graduating students is 1.8%.

The political science programme has a common threshold for mathematics and language students, whilst no students from the 'other' group are admitted. The increase in the graduation rate is estimated at 6.8 percentage points.

For the psychology programme, the horizontal line corresponds to a graduation probability of 0.67. Irrespective of GPA, mathematics students have

[11] At a GPA level of 1.0 (the admission threshold for language students in the law programme), language students have a 9% lower graduation probability than mathematics students in the law programme but an 18% lower graduation probability in the economics programme.

graduation probabilities above the line, whilst language students and students from the 'other' group have graduation probabilities on or below the line. As mathematics students have higher graduation rates than the two other groups, irrespective of GPA, admitting only mathematics students maximises the graduation rate for the psychology programme. This finding proves valid despite the comparatively low advantage of being a mathematics students in the psychology programme: the difference in graduation probability is about 5 percentage points in the GPA range 1-2.5.

The highest increase in graduation comes from changes in admission to the psychology programme. In this programme higher grades do not lead to a higher graduation probability, so expression (12) for the gain in the graduation rate is not valid (as $p' = 0$). However, a simple calculation shows that the gain in terms of increase in aggregate graduation rate is a substantial 13.6 percentage points.[12] As the initial graduation rate was 55.9%, introducing the optimal admission rule is expected to increase the graduation rate of psychology students by 24.3%.

The difference in graduation rates between mathematics students and language students is statistically significant for the law, economics and psychology programmes (see the standard errors in table 1). Figure 3 shows small differences in graduation rates in the lower part of the grade distributions for both the law and the economics programmes; thus the larger differences in graduation rates in the upper part are significant.

For the four social science programmes in total, the gain in graduation rate from a transition to an optimal admission system is substantial. The estimated increase in the number of graduating students is a sizeable 10.3%.

This section has illustrated how to maximise graduation rates in higher education programmes by admitting more students with a high graduation probability and fewer students with a low graduation probability. The results show that mathematics students, conditional on GPA, should be admitted more frequently to the law, economics and psychology programmes than students from other high school tracks.

The change in admission structure can have derived consequences both for students' choice of university programmes and for students' choices of subjects in secondary school. I consider these issues in the following sections.

[12] The statistics for graduation rates and shares in Table 1 enter the calcuation as 0.287x(0.695-0.621)+0.472(0.695-0.452)=0.136.

6 Optimal admission and the college admission problem

In this section I clarify the relation between optimal admission procedures, as defined in this paper, and the literature on the 'college admission problem', which also addresses the question of allocation of students to colleges. This section addresses the question of whether the change in admission policy suggested in this paper leads to strategic behaviour amongst either students or colleges, which is a topic in literature on the college admission problem.

The seminal contribution to the college admission problem is Gale and Shapley (1962), who began by analysing a simpler problem, the 'stability of marriages'. This problem consists of two sets of agents, where each element of one of the sets has ordered preferences for the elements in the other set. The elements of the two sets form one-to-one matchings. The college admission problem also consists of two sets of agents, but with the two sets instead forming many-to-one matchings.

A core element in the matching literature is to what extent agents can obtain a better matching by misrepresenting their true preference, that is, by engaging in strategic behaviour. According to Roth and Sotomayor (1990), strategic behaviour appears to be an optimal strategy in many matching problems.

However, recent contributions (Abdulkadiroglu and Sönmez (1998) and Svensson (1999)) have considered a particular matching procedure whereby strategic behaviour is not an optimal strategy, namely serial dictatorship. A good example of serial dictatorship in practice is the admission process to higher education in Denmark, described in the section 'institutional setting'. The central admission office ranks all applicants according to their high school GPA. The student with the highest GPA chooses first (the first dictator), the student with the next highest GPA chooses second (the second serial dictator) and so forth until the allocation process stops.

The particular algorithm applied to matching students and college programmes in Denmark does not give students any incentives to misstate their true preferences either before or after a change to optimal admission procedures (moreover, in the terminology of the matching literature, the matching is both 'stable' and 'Pareto efficient', see Abdulkadiroglu and Sönmez (1998) and Svensson (1999)). Although strategic behaviour as a result of a change in admission procedures might be a problem in other systems for matching students and colleges, this possibility is outside the scope of the present paper and is left for future research.

The literature on the college admission problem analyses admission to college given students' preferences for colleges and colleges' preferences for students. This paper considers how colleges can construct admission rules that increase graduation rates. In terms of the college admission problem, this paper thus deals with how colleges or programmes can form preferences for students (i.e. how they can rank students). The literature on the college admission problem takes preferences as data in the analysis. Neither this paper nor the literature on the college admission problem considers how students form preferences.

7 The choice of subjects in secondary school

The change in admission rules may have an impact on the students' choice of subjects in secondary school. This section establishes a framework for analysing the choice of subjects by secondary school students. Within this framework I trace the impact of a change in admission rules on the choice of subjects.

Students' choice of subjects is relevant for the topic of this paper because the curriculum in high school may have an impact on the probability of graduating from university. I analyse this topic in the next section, which builds on the analysis in this section.

The analytical framework in both this and the following section is the standard model of individual choice in econometrics. The model is applied in, for example, Heckman (1979) and reviewed in Maddala (1983).

Section two of the paper considered the choice of the authority that determines admission criteria, given the choice of students. This section and the following section consider the choice of students, given the choice of the authority that determines admission criteria.

Assume that secondary school students choose amongst two subjects: mathematics leading to an A-level in mathematics (the mathematics track) and an alternative subject leading to a B-level in mathematics (the non-mathematics track). The indirect utility for student s from choosing mathematics at A-level is specified as

$$U_A^s = \omega \gamma_A^s + w_s \delta_A + \epsilon_A^s, \tag{13}$$

where U_A^s is the utility from choosing mathematics at A-level, γ_A^s is the expected grade in mathematics, ω is the weight of mathematics in calculating the GPA used for admitting students to university programmes, w_s is a vector of variables determining the preferences for mathematics with the associated parameter vector δ_A, and ϵ_A^s is a term that includes preferences for mathematics beyond the expected grade and explanatory variables.

Correspondingly, the utility from choosing the alternative non-mathematics subject leading to a B-level in mathematics, is specified as

$$U_B^s = (1-\omega)\,\gamma_B^s + w_s \delta_B + \epsilon_B^s, \tag{14}$$

where U_B^s is the utility from choosing the subject, γ_B^s is the expected grade in the non-mathematics subject, $(1-\omega)$ is the weight of the non-mathematics subject, δ_B is a parameter vector associated with the explanatory variables and ϵ_B^s is the error term that includes preferences for the non-mathematics subject.

Mathematics is chosen by student s if

$$\omega \gamma_A^s - (1-\omega)\,\gamma_B^s + w_s\,(\delta_A - \delta_B) > \epsilon_B^s - \epsilon_A^s.$$

The variance of the error term is

$$\sigma^2 = \sigma_{\epsilon_B}^2 + \sigma_{\epsilon_A}^2 - 2\sigma_{\epsilon_B \epsilon_A}.$$

The following notation is adopted when the standard deviation is applied to normalise the error term and the covariates

$$\begin{aligned} \epsilon_s &= \frac{\epsilon_B^s - \epsilon_A^s}{\sigma} \\ z_s &= \frac{\omega \gamma_A^s - (1-\omega)\,\gamma_B^s}{\sigma} + \frac{w_s\,(\delta_A - \delta_B)}{\sigma}. \end{aligned} \tag{15}$$

The probability that student s chooses mathematics thus becomes

$$P\,(\epsilon_s < z_s) = G\,(z_s)\,,$$

where G is the distribution function of the random variable ϵ_s.

Differentiating yields

$$\frac{\partial G}{\partial \omega} = g\left(z_s\right)\frac{\partial z_s}{\partial \omega} > 0, \quad \frac{\partial z_s}{\partial \omega} = \frac{\gamma_A^s + \gamma_B^s}{\sigma} > 0,$$

where g is the density function. If the weight of mathematics in the admission criterion is increased, more students will choose this subject.

Three main assumptions drive the proof. First, students choose subjects by comparing the gain or utility from choosing mathematics with the utility from choosing some alternative subject. Second, the expected grade is included in the gain or utility from the choices. Third, a spread exists in student preferences for choosing mathematics or the alternative. If these assumptions are fulfilled, an increased weight to mathematics in the calculation of the GPA will move students on the borderline of choosing this subject, whilst students with low relative preferences for mathematics will continue to study the alternative subject.

8 Behavioural change and changes in graduation rates

The previous section traced the effect of the policy change on students' choice of subjects in high school and a change in high school subjects may have an impact on the graduation probability of future applicants to university programmes. This section investigates the impact of the policy change on changes in graduation rates at universities. As in the previous section, this one considers the choice of students given the choice of the authority that determines admission criteria.

The analysis in this section is important for two reasons. First, better-prepared applicants to higher education imply higher graduation rates and potentially a higher skill level of the workforce. Second, a change in graduation probability amongst future students from different high school tracks implies that the authority that determines admission criteria may have to alter admission criteria.

However, in practice admission rules are adjusted frequently for other reasons, such as changes over time in the educational system and other factors that influence students' choices, which results in changes in the number and the composition of applicants to higher education.[13] An adjustment of admission criteria as a consequence of a change of students' choice of subjects in high school can thus be considered to be one extra element in the ongoing admission criteria adjustment process.[14]

This section extends the model in the previous section for students' choice of subjects in high school with a model for graduation from university. The graduation probability for student s in a particular university programme is denoted y_s. The graduation probability is approximated by the linear probability model

$$\begin{aligned} y_s &= X_s\beta + \alpha_s m_s + e_s \\ \alpha_s &= a_0 + a_s, \end{aligned} \tag{16}$$

where X_s is a vector of graduation determinants, including grades, β is the associated coefficient vector, m_s is an indicator variable taking the value 1 if mathematics is chosen and 0 otherwise, α_s measures the change in graduation probability that student s obtains as a consequence of studying mathematics (that is, α_s is the 'causal effect' of the mathematics subject on the graduation probability for student s), and e_s is the error term.

The term a_0 is the average change in graduation probability amongst the population of students, and a_s is the deviation for student s from the population average. The heterogeneous return to mathematics is specified as a random coefficient model, which is applied in the recent literature on the impact of schooling and training on labour market outcomes (see Card (1999), and Heckman et al. (1999)). To obtain closed form solutions, I assume that errors are multivariate normal as in the seminal contribution on the random coefficient model by Björklund and Moffitt (1987).

The expected graduation probability for a student who has studied mathematics is

[13] E.g. the University of California frequently adjusts the 'admission index' for California residents.

[14] Estimation of a structural model including identification of the distribution of unobserved preferences and abilities is beyond the scope of this paper. Identification requires instruments for selection into the different high school paths (see e.g. the discussion in Altonji et al. (2012)). The present data do not contain variables that are suitable as instruments in such an analysis.

$$E\left(y_s|X_s, m_s=1\right) = X_s\beta + a_0 + E\left(a_s + e_s|\epsilon_s < z_s\right), \tag{17}$$

where ϵ_s and z_s are defined in (15).

Applying the expression for a standard normal variable truncated from above, the expected value of the conditional error term becomes

$$E\left(\alpha_s + e_s|\epsilon_s < z_s\right) = -\sigma_{a+e,\epsilon}\frac{g\left(z_s\right)}{G\left(z_s\right)}, \tag{18}$$

where G and g are the distribution and density functions from the previous section of the paper, now assumed to follow the standard normal distribution.

The covariance $\sigma_{a+e,\epsilon}$ between the preference term ϵ_s in (15) and the sum of the error term e_s and the random coefficient term a_s in (16) becomes

$$\begin{aligned} \sigma_{a+e,\epsilon} &= \sigma_{a\epsilon} + \sigma_{e\epsilon} \\ \sigma_{a\epsilon} &= \frac{\sigma_{a\epsilon_B} - \sigma_{a\epsilon_A}}{\sigma} \\ \sigma_{e\epsilon} &= \frac{\sigma_{e\epsilon_B} - \sigma_{e\epsilon_A}}{\sigma}. \end{aligned} \tag{19}$$

The expected graduation probability for a student who has not studied mathematics is

$$E\left(y_s|X_s, m_s=0\right) = X_s\beta + E\left(e_s|\epsilon_s > z_s\right). \tag{20}$$

Applying the expression for a standard normal variable truncated from below, the expected value of the conditional error term becomes

$$E\left(e_s|\epsilon_s > z_s\right) = \sigma_{e,\epsilon}\frac{g\left(z_s\right)}{1-G\left(z_s\right)}. \tag{21}$$

The difference in expected graduation probabilities between the mathematics and the non-mathematics students thus becomes

$$\begin{aligned} & E\left(y_s|X_s, m_s=1\right) - E\left(y_s|X_s, m_s=0\right) \\ = \;& a_0 - \sigma_{a\epsilon}\frac{g\left(z_s\right)}{G\left(z_s\right)} - \sigma_{e,\epsilon}\left[\frac{g\left(z_s\right)}{G\left(z_s\right)} + \frac{g\left(z_s\right)}{1-G\left(z_s\right)}\right]. \end{aligned} \tag{22}$$

For the law, economics and psychology programmes at the University of Copenhagen, mathematics students have a higher graduation probability than non-mathematics students conditional on GPA (see fig. 3) and the difference (22) is thus positive.

Each of the three terms on the right-hand side of (22) can contribute to a positive difference in the graduation probability between mathematics and non-mathematics students. The first term contributes to a positive difference if the average student increases the graduation probability by taking mathematics ($a_0 > 0$). The second term contributes to a positive difference if and only if (see (19))

$$\sigma_{a\epsilon} < 0 \Leftrightarrow \sigma_{a\epsilon_B} < \sigma_{a\epsilon_A}.$$

This inequality arises when students with high preferences for mathematics experience higher increases in graduating from the programme (the a_s term in (16)) by studying mathematics in secondary school than do students with low preferences for mathematics. The third term on the right-hand side (22) contributes to a positive difference if and only if (see (19))

$$\sigma_{e,\epsilon} < 0 \Leftrightarrow \sigma_{e\epsilon_B} < \sigma_{e\epsilon_A}.$$

This case applies when students with high preferences for mathematics have a higher unobserved probability of graduating from the programme (the error term e_s in (16)) than students with low preferences for mathematics (relative to the non-mathematics subject).

The sum of the three terms on the right-hand side of (22) is positive but none of the three terms are necessarily positive. If, for example two of the terms are zero, then the third must constitute the whole of the positive left-hand side of (22).

I now consider the consequences of a policy change that gives mathematics a higher weight in the admission criterion on the graduation probability both for students who choose mathematics and for students who do not. The change in the graduation probability for students who choose mathematics follows from differentiating (17) taking (18) into account. As $\partial g\left(z_s\right)/\partial z_s = -z_s g\left(z_s\right)$, the result becomes

$$\frac{\partial}{\partial \omega}\left[-\sigma_{a+e,\epsilon}\frac{g\left(z_s\right)}{G\left(z_s\right)}\right] = \left(\sigma_{a\epsilon} + \sigma_{e\epsilon}\right)\left(z_s + \frac{g\left(z_s\right)}{G\left(z_s\right)}\right)\frac{g\left(z_s\right)}{G\left(z_s\right)}\frac{\partial z_s}{\partial \omega}. \qquad (23)$$

For $z_s > -g\left(z_s\right)/G\left(z_s\right)$ the term in the brackets on the right-hand side is positive, and calculations show that this expression is positive for all values of z_s.

The effect of the policy change on graduation for students who do not choose mathematics follows from differentiating (20) taking (21) into account

$$\frac{\partial}{\partial \omega}\left[\sigma_{e,\epsilon}\frac{g\left(z_s\right)}{1-G\left(z_s\right)}\right]=\sigma_{e,\epsilon}\left(-z_s+\frac{g\left(z_s\right)}{1-G\left(z_s\right)}\right)\frac{g\left(z_s\right)}{1-G\left(z_s\right)}\frac{\partial z_s}{\partial \omega}. \tag{24}$$

For $z_s < g\left(z_s\right)/\left(1-G\left(z_s\right)\right)$ the term in the brackets on the right-hand side is positive, and calculations show that this expression is positive for all values of z_s.

The effect of the policy change on the graduation probability depends on the reason for mathematics students having higher graduation probabilities than non-mathematic students. The previous discussion shows that mathematics students can have higher graduation probabilities than non-mathematic students for three reasons, and I now consider them in turn.

First, the higher graduation probabilities associated with the mathematics subject is the result of a higher average graduation probability ($a_0 > 0$), whilst no heterogeneity exists in the returns from choosing mathematics ($\sigma_{a\epsilon} = 0$), or in graduation probabilities ($\sigma_{e,\epsilon} = 0$). Thus the right-hand side of both (23) and (24) is zero, and the policy change has thus no impact on the graduation probabilities of either mathematics or non-mathematics students. The new students who choose the mathematics subject as a consequence of the policy change experience the same increase in graduation probability as the previous students. In terms of Figure 1, the curves for mathematics and non-mathematics students do not change.

Second, the higher graduation probabilities associated with the mathematics subject is the result of heterogeneity in the returns from choosing mathematics ($\sigma_{a\epsilon} < 0$), whilst there is neither an increase in graduation probability for the average student following the subject ($a_0 = 0$), nor heterogeneity in graduation probabilities ($\sigma_{e,\epsilon} = 0$). In this case the right-hand side of (23) is negative, the graduation probability for students following the mathematics subject decreases, whilst the right-hand side of (24) is zero, there is no change in graduation probability for students, who do not choose the mathematics subject. Whilst the policy change induces more students to take mathematics, these students, on average, derive less from mathematics with respect to increases in graduation probability than those who took this subject before the policy change. In Figure 1, the upper curve moves downwards, whilst the lower curve remains unchanged.

Third, the higher graduation probabilities associated with the mathematics subject is the result of heterogeneity in graduation probabilities ($\sigma_{e,\epsilon} < 0$), whilst there is neither an increase in graduation probability for the average student following the subject ($a_0 = 0$), nor heterogeneity in the returns from choosing mathematics ($\sigma_{a\epsilon} = 0$). The right-hand side of (23) is negative, the graduation probability decreases for students who choose the mathematics subject. Whilst the policy change induces more students to choose mathematics, these students, on average, have lower preferences for mathematics and do not have the same graduation probability in the programme as those who took the subject before the policy change. The upper curve in Figure 1 thus moves downwards. The right-hand side of (24) is also negative, the graduation probability decreases for students who do not choose the mathematics subject. The policy change induces more students to choose mathematics. As a result, the remaining students – those who do not choose mathematics – have lower preferences for mathematics and thus a smaller graduation probability than those who shift to mathematics. The lower curve in Figure 1 also moves downwards.

Under the assumption that giving higher weight to mathematics in the admission criteria induces more students to take this subject, the analysis has identified three cases of outcomes. When the impact of the mathematics subject on graduation is the same for all students, the policy change has no impact on the conditional expected graduation functions, and the curves in Figure 1 remain unchanged. To the extent that (some of) the difference in graduation rates between mathematics and non-mathematics students is due to heterogeneous returns from mathematics, the expected graduation rate for mathematics students will decrease whilst the expected graduation rate for non-mathematics students remains the same as before the policy change. The difference between the two groups in terms of graduating from the programme will thus diminish. In contrast, to the extent that (some of) the difference in graduation between mathematics and non-mathematics students is due to a higher graduation probability for students with a high relative preference for mathematics, the expected graduation rate for mathematics students will decrease, and the same holds for the expected graduation rate for non-mathematics students.

The analysis in the section is based on equation (22), which shows how the difference in the graduation probability between mathematics and non-mathematics students can be decomposed into different components, given the assumptions of the model. The magnitude of the components depends on the size of the causal impact of mathematics on graduating from university (the term $\alpha_s = a_0 + a_s$ in equation (16)) and preferences for choosing mathematics relative to the non-mathematics subject (the term ϵ_s in equation (15)).

However, this decomposition has no bearing on the analysis in section 2 of graduation rates and the gains in graduation rates obtained by introducing optimal admission rules. The analyses in sections 2 and 3 hinge solely on the empirical observation that mathematics students have a higher graduation rate than non-mathematics students, that is, that the left-hand side of equation (22) is positive. These direct gains from optimal admission accrue irrespective of whether the difference in graduation probabilities between mathematics and non-mathematics students is due to a causal effect of mathematics on the graduation probability (the two first terms on the right-hand side of (22)) or to self-selection (the third term on the right-hand side of (22)). However, indirect gains may exist if optimal admission induces more students to choose mathematics and if these students obtain a higher graduation probability, that is, if there is a causal impact of the mathematics subject on graduation probabilities.[15]

This section shows that the changes in the incentive structure of the education system may have consequences for the observed outcome of choosing mathematics subjects. The changes in graduation probabilities for students from different high school tracks depend on the extent to which mathematics has a causal impact on graduating from university. Changes in the observed outcome may, in turn, imply that admission criteria to higher education need to be adjusted.

[15] Joensen and Nielsen (2009) provide evidence for derived gains of optimal admission as they find substantial effects of advanced high school mathematics on subsequent earnings. The authors exploit an educational reform in Denmark where a change in the content of the high school tracks implied that more students chose advanced mathematics. The main part of the effect on earnings is indirect and goes through choice of higher education.

9 Conclusion

This paper analyses admission policies of the higher education system and derives a policy rule that maximises graduation rates. This 'optimal admission' rule implies that the marginal graduation rates of students from various high school tracks are equalised. Introduction of optimal admission results in increases in graduation rates, and the paper thus contributes to the literature on college enrolment and completion, an important policy issue (see, e.g. Bound and Turner (2011)).

Optimal admission will result in a large increase in the graduation rate for a higher education programme if there are (1) large differences in initial graduation rates between various groups of students admitted to the programme, (2) a small derivative of the graduation probability as a function of GPA, and (3) an equal composition of the intake of students from different groups.

The application of the optimal admission rule to the social science programmes at the University of Copenhagen demonstrates that mathematics students, conditional on GPA, should be admitted more frequently to the law, economics and psychology programmes than students from other high school tracks. For the four social science programmes in total, the gain in graduation from a transition to an optimal admission policy is substantial.

In combination with the theory in this paper, analogous empirical analysis can form the basis of optimal admission rules in other educational systems. Optimal admission could be fine-tuned to single programmes at one university, or common rules could be applied to several programmes within or across universities.

Introduction of optimal admission rules is expected to alter the choice of subjects by upper-secondary school students in such a way that they try to better prepare themselves to graduate from programmes in the higher education system. This altered choice is an indirect gain from a changed admission system. Even when the entire difference in graduation rates between different groups of students is due to self-selection, optimal admission of students is a valid policy rule. In this case, whilst a direct gain of increased graduation rates exists, no indirect gain appears from applying a policy rule maximising graduation rates.

However, most empirical results show that only a minor part of the estimated economic gain of one more year of education is due to self-selection (see, e.g. the survey by Card (1999)). To the extent that self-selection also plays a minor role in the impact of upper-secondary school subjects for understanding the content of programmes in the higher education system, there are derived gains from admission policies that maximise graduation rates. These derived gains have the potential to result in entire cohorts of upper-secondary school students being more likely to graduate with a higher education degree.

The theory developed in this paper and the application of the theory in the empirical analysis demonstrates how to increase graduation rates and further the smooth transition of students through the higher educational system. The application of an optimal admission policy for higher education is thus a means of increasing the skill level of the work force, which is crucial for labour market outcomes and economic growth.

Acknowledgements

Thanks for comments from the editor, two referees, Atila Abdulkadiroglu, John Kennes, Dean Lillard, Rune Vejlin and seminar participants at the Copenhagen Business School, the conference of the Danish Economic Society, the Barcelona workshop of the European Network for Research on Matching Practices in Education and Early Labour Markets, the Congress of the European Economic Association in Gothenburg, and the International Workshop on Applied Economics of Education in Catanzaro.

Literature

Abdulkadiroglu, A. and T. Sönmez, 1998. Random serial dictatorship and the core from random endowments in house allocation problems. Econometrica, vol. 66 (3), 689-701.

Albæk, K., 2009. The Danish apprenticeship system, 1931-2002: the role of subsidies and institutions. Applied Economics Quarterly, vol. 55(1), 39-60.

Altonji, J.G., E. Blom and C. Meghir, 2012. Heterogeneity in human capital investments, high school curriculum, college major, and careers. NBER Working Paper no. 17985.

Björklund, A. and R. Moffitt, 1987. The estimation of wage gains and welfare gains in self-selection models. Review of Economics and Statistics, vol. 69(2), 42-49.

Bound, J., M.F. Lovenheim and S. Turner, 2010. Why have college completion rates declined? An Analysis of Changing Student Preparation and Collegiate Resources. American Economic Journal: Applied Economics, vol. 2(3), 129-157.

Bound, J., and S. Turner, 2011. Dropouts and diplomas: the divergence in Collegiate outcomes. In: Hanushek, Machin and Woessmann (Eds.), Handbook of the Economics of Education, vol. 4. Elsevier, Amsterdam, pp. 573-613.

Bowen, W.G., M.M. Chingos and M.S. McPherson, 2009. Crossing the finish line. Completing college at America's public universities. Princeton University Press, Princeton.

Card, D., 1999. The causal effect of education on earnings. In: Ashenfelter and Card (Eds.), Handbook of Labour Economics, vol. 3A. Elsevier, Amsterdam, pp. 1801-63.

Epple D., R. Romano, and H. Sieg. 2006. Admission, tuition, and financial aid policies in the market for higher education. Econometrica, vol. 74 (4): 885–928.

Fu, C. 2014. Equilibrium tuition, applications, admissions, and enrollment in the college market. Journal of Political Economy, vol. 122 (2), 225-281.

Gale, D. and L.S. Shapley. 1962. College Admissions and the Stability of Marriage. The American Mathematical Monthly, vol. 69 (1), 9-15.

Gary-Bobo, R. J., and A. Trannoy. 2008. Efficient Tuition Fees and Examinations. Journal of the European Economic Association vol. 6 (6), 1211–1243.

Heckman, J.J. 1979. Sample Selection Bias as a Specification Error, Econometrica, vol. 47(1), 153-161.

Heckman, J.J., R.J. LaLonde and J.A. Smith, 1999. The economics and econometrics of active labor market programs. In: O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, vol. 3A. Elsevier, Amsterdam, pp. 1865-2097.

Joensen, J.S. and H.S. Nielsen. 2009. Is There a Causal Effect of High School Math on Labor Market Outcomes? Journal of Human Resources. vol. 44(1), 171-198.

Maddala, G.S. 1983. Limited-Dependent and Qualitative Variables in Econometrics. Econometric Society Monographs. Cambridge: Cambridge University Press.

Manski, C.F. and D.A. Wise, 1983. College Choice in America. Harvard University Press, Cambridge.

Roth, A.E. and M.A.O. Sotomayor, 1990. Two-sided matching. A study in game-theoretic modeling and analysis. Econometric Society Monographs no. 18. Cambridge University Press, Cambridge.

Rothstein, J.M., 2004. College performance predictions and the SAT. Journal of Econometrics, vol. 121(1-2), 297-317.

Stinebrickner, R. and T. Stinebrickner, 2014. Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. Journal of Labor Economics, vol. 32(3), 601-644.

Svensson, L.-G., 1999. Strategy-proof allocation of indivisible goods. Social Choice and Welfare, vol. 16, 557-567.

Tinto, V., 1993. Leaving College. 2nd ed. University of Chicago Press, Chicago.

Wand, M.P. and M.C. Jones, 1995. Kernel Smoothing. Chapman and Hall, London.

Fig. 1. *Transition from admission system with common threshold to optimal admission system*

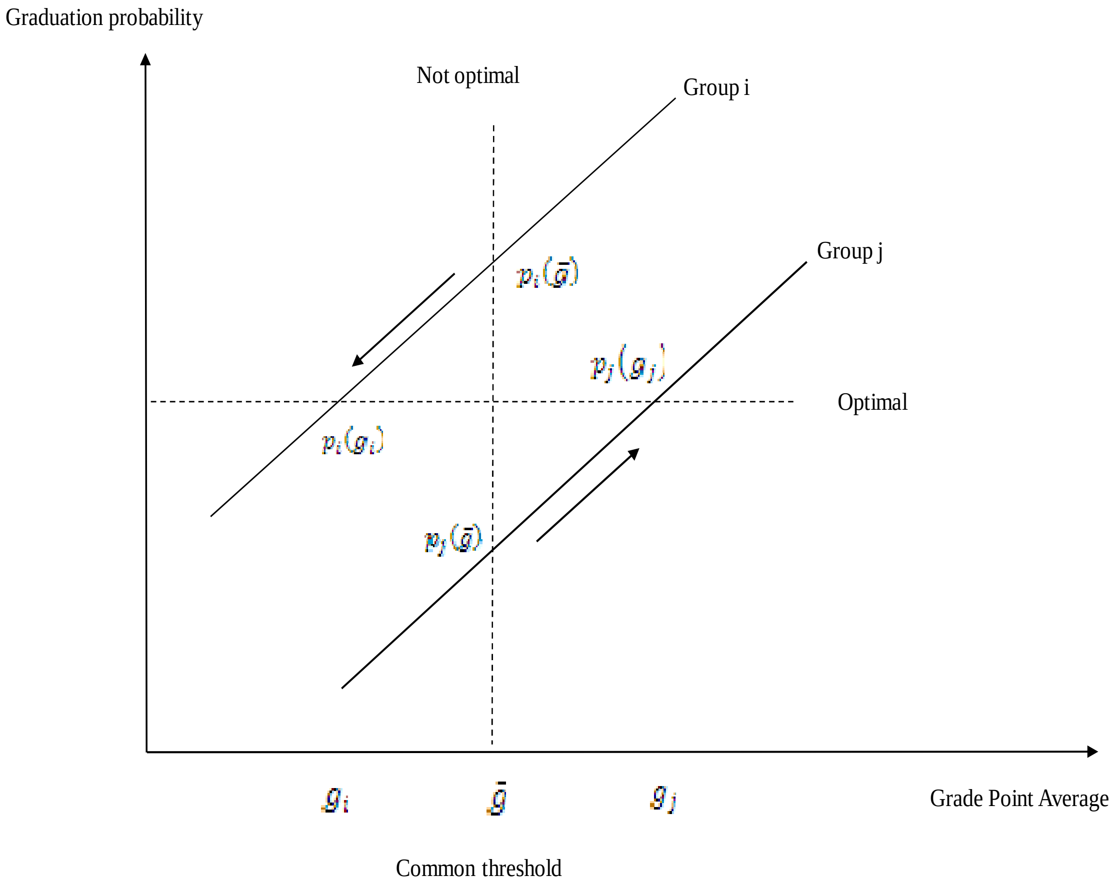


Fig.2. GPA for admitted students in social science

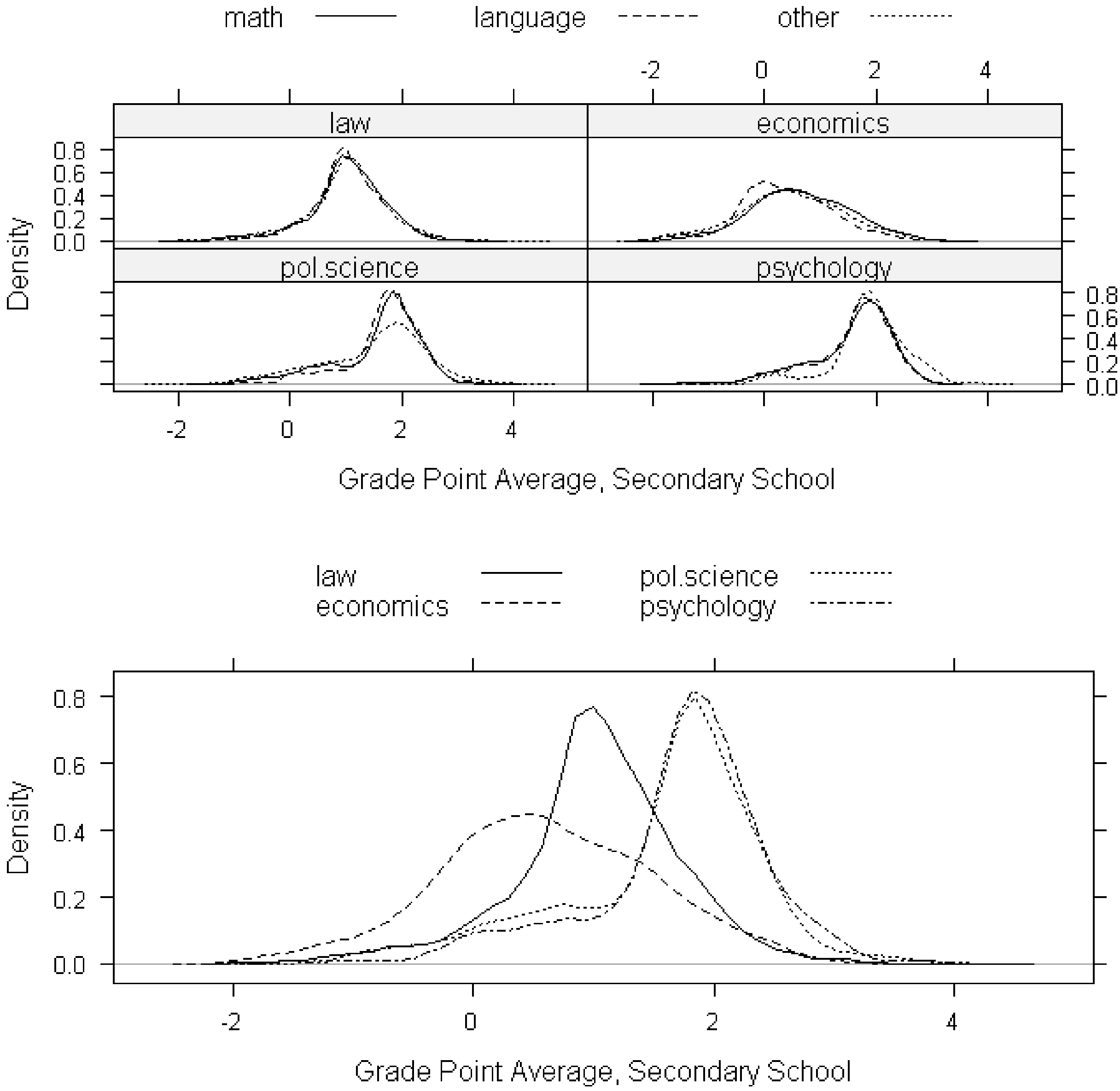
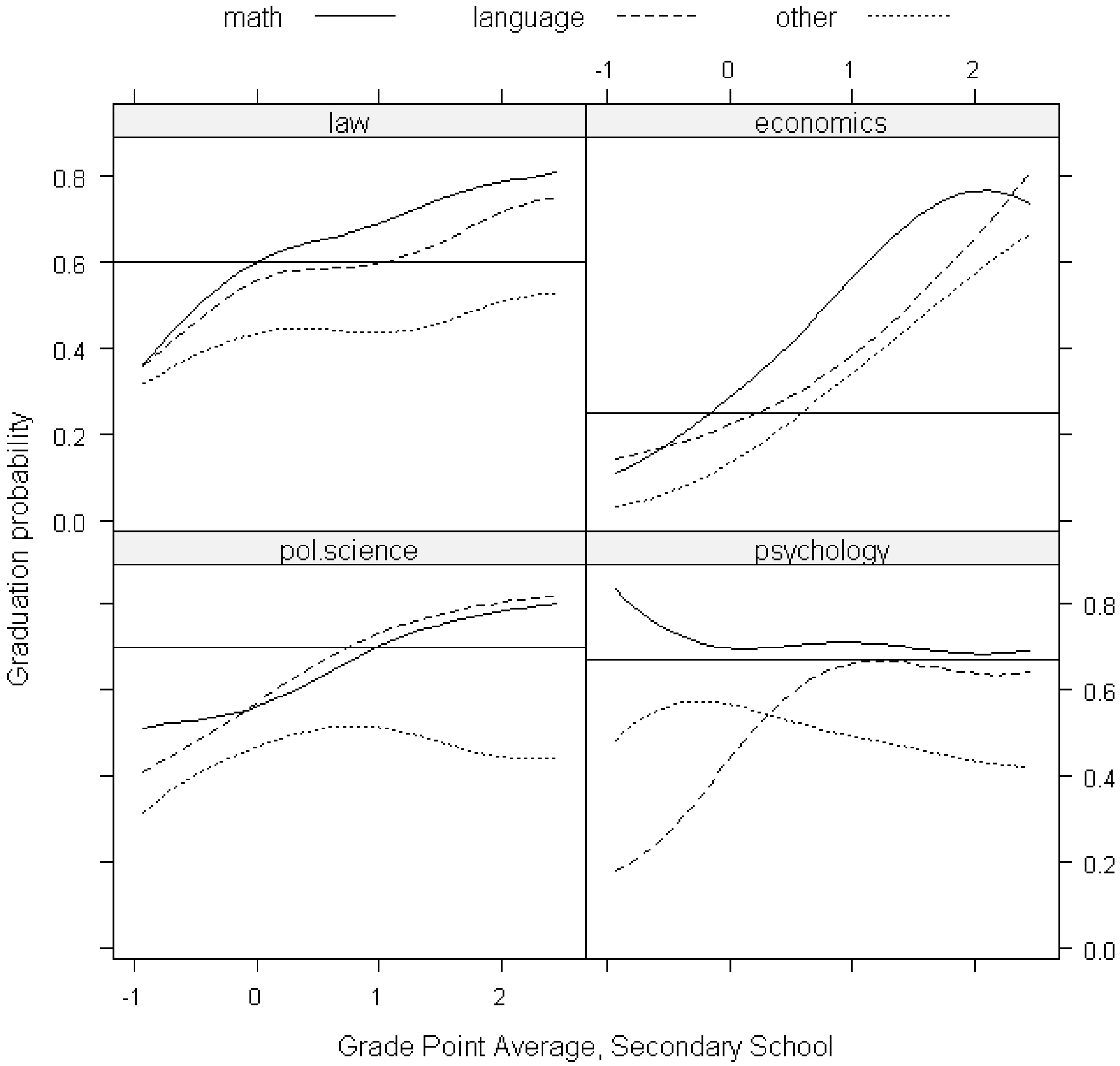


Fig. 3. Graduation probability and GPA in Social Science



Note: Optimal admission rules shown as horisontal lines

Fig. 4. Slope of Conditional Expected Graduation Functions

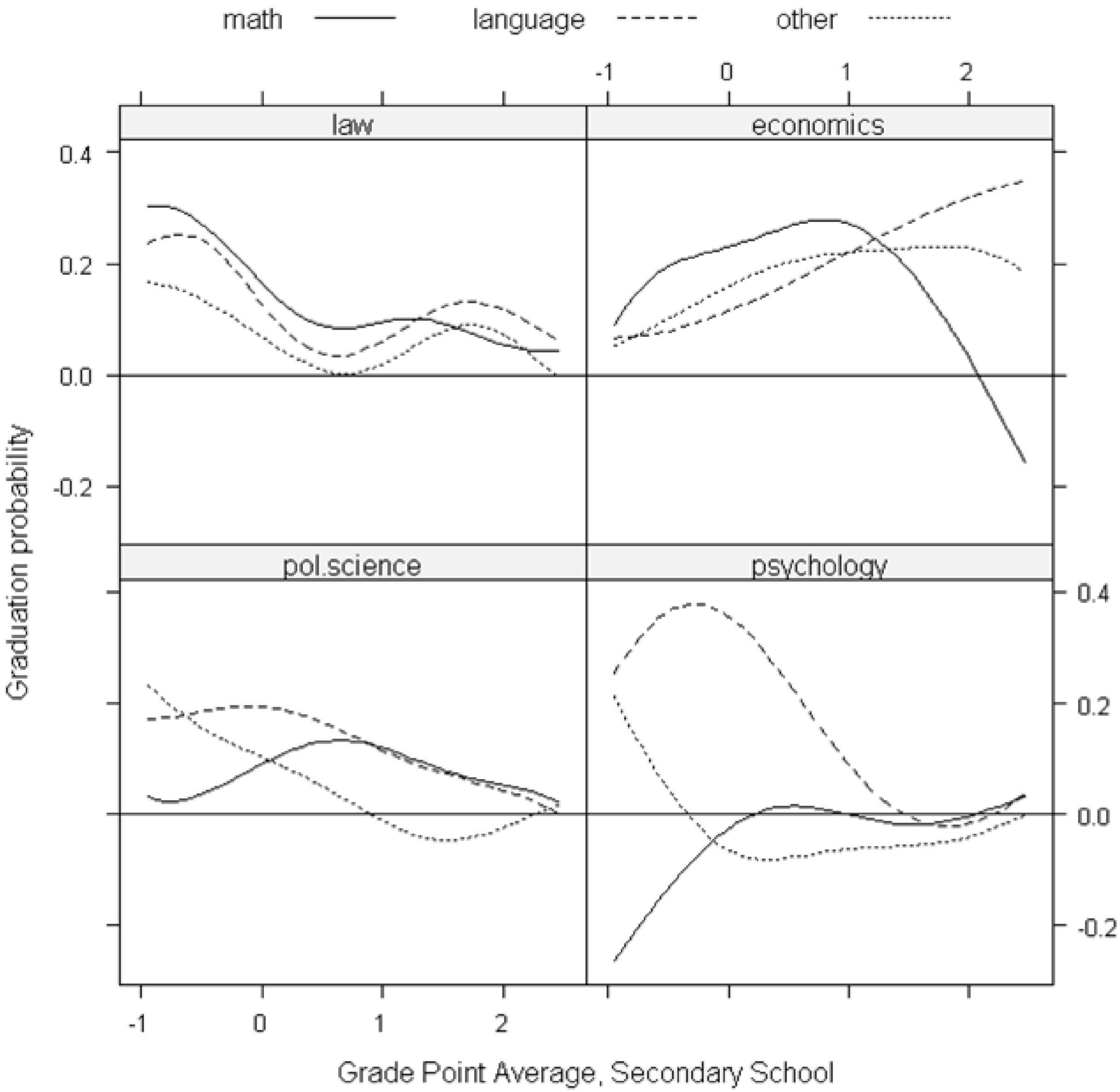


Table 1. *Sample statistics for three branches of secondary school and statistics for optimal admission.*

	Sample statistics			Common threshold[1]		Threshold	Graduation change[3]	
	Share	Graduation rate		Grad. rate	Slope	change[2]	Points	Relative
Programme:								
Law								
Math	0.386	0.693	(0.011)	0.649	0.090	-0.460		
Language	0.329	0.616	(0.012)	0.586	0.040	0.539		
Other	0.285	0.450	(0.013)					
Total	1.000	0.598	(0.007)				0.062	0.104
Economics								
Math	0.685	0.466	(0.012)	0.284	0.229	-0.221		
Language	0.042	0.308	(0.042)	0.221	0.113	0.055		
Other	0.273	0.252	(0.016)	0.131	0.157	0.613		
Total	1.000	0.401	(0.009)				0.007	0.018
Pol. Science								
Math	0.455	0.735	(0.017)					
Language	0.299	0.766	(0.020)					
Other	0.246	0.460	(0.026)					
Total	1.000	0.677	(0.012)				0.068	0.100
Psychology								
Math	0.241	0.695	(0.027)					
Language	0.287	0.621	(0.026)					
Other	0.472	0.452	(0.021)					
Total	1.000	0.559	(0.014)				0.136	0.243
Total								
Math	0.457	0.610	(0.007)					
Language	0.245	0.629	(0.010)					
Other	0.298	0.404	(0.009)					
Total	1.000	0.553	(0.005)				0.057	0.103

Note: Standard errors in parenthesis. Admission to the University of Copenhagen Summer 1984 - Summer 2001. Total number of admitted students was 10418 distributed on 47.2% in law, 26.3% in economics, 14.7% in political science and 11.8% in psychology. Graduation with a bachelor degree within four years of study. [1] Graduation probabilities and slopes are measured at the common threshold. [2] Changes in admission thresholds as a consequence of optimal admission. [3] Percentage points increase in graduation rate as a consequence of optimal admission and increase relative to previous level of graduation.