Trickle down education - ripple effects in college admissions

Mikkel Høst Gandil*

June 2023

Abstract

Admitting an applicant to an educational program will free up a slot to be filled at her next-best alternative. This paper investigates the impact of such ripple effects in an educational market and their significance for policy evaluation. By reconstructing the national centralized allocation mechanism for higher education in Denmark, I analyze the ripple effects resulting from local capacity changes. I use reported and estimated preferences to simulate thousands of local supply policies and determine the magnitude of ripple effects. I find that ripple effects are substantial as admitting one applicant affects almost three other applicants on average and vary substantially across educational programs. Ripple effects play a crucial role in policy evaluation. I estimate returns to education in terms of earnings, and show that ignoring ripple effects results in missing a significant portion of the variation in effects on earnings and in many cases can change policy recommendations. The findings highlight the importance of considering applicant sorting and rationing in policymaking related to education. The results demonstrate the relevance of ripple effects and market clearing for policy interventions in various markets and sectors.

JEL Codes: C63,I21,I22,I24, I26, J24, H52
Keywords: Field of study, College admission, Program evaluation, RDD

*Department of Economics, University of Oslo. Postboks 1095 Blindern 0317, Oslo, Norway. Telephone: +45 28736769. Mail: mikkel.gandil@econ.uio.no. The Norwegian Research Council funded this research under project no. 275906. I am grateful to Edwin Leuven, Bjørn Meyer, Anders Humlum, Sturla Løkken, Anders Munk-Nielsen, Monique de Haan, Manudeep Bhuller, Paolo Giovanni Piacquadio, Lasse Jensen, Morten Nielsen, Simon Damkjær as well as seminar participants for helpful comments and suggestions. An earlier version of this paper has been circulated under the title "Substitution effects in College Admissions".

1 Introduction

When a policymaker considers changing the size of an educational program, it is crucial to consider the associated returns on investment. However, in oversubscribed programs, admitting an applicant to one program creates an available spot in their next-best alternative, which can then be filled by an applicant from a different program. As a result, local supply changes can have ripple effects throughout the market, impacting other applicants and programs. Do these ripple effects increase or diminish the social returns on investment in education? And how significant are these effects in magnitude? Despite their potential importance for policy evaluation, these questions remain largely unexplored empirically.

This paper aims to identify and analyze ripple effects in an educational market where applicants can have an impact on each other's admissions. By reconstructing the centralized allocation mechanism for higher education in Denmark, I can track all applicants and examine how local supply changes propagate through different programs. To capture the effects of significant supply changes, I estimate random utility models to allow applicants to change behavior in counterfactual scenarios. Using reported and estimated preferences, I simulate thousands of local supply changes and determine the magnitude of ripple effects by comparing the number of affected applicants to those who move into or out of expanding or contracting programs. To my knowledge, this study represents the first attempt to quantify ripple effects within a national higher education market using full counterfactual market equilibria.

Within applied microeconomics and labor, program evaluations traditionally ignore ripple effects. To connect the presence of ripple effect with program evaluation, I demonstrate that the Local Average Treatment Effects (LATEs) for marginal applicants serve as policy-relevant treatment effects. However, to fully consider the ripple effect, treatment effect estimates are necessary for all programs that applicants may switch to after a local supply change. Additionally, one needs to know the flows of applicants between programs. Consequently, the informational demands on a social planner are significantly greater compared to a program manager and misalignment of incentives between the social planner and program manager may arise as a result.

To analyze ripple effects empirically, I combine simulated allocations of applicants under counterfactual supply with estimates of returns in earnings. I estimate the LATEs for program-specific marginal applicants using a regression discontinuity design (RDD) integrated into the allocation mechanism, utilizing three decades of Danish register data. To capture the ripple effects on earnings, I compare changes in earnings for the marginal applicants in programs where supply is altered to the aggregate changes for all applicants in the admission round. In cases of significant supply changes, non-parametric LATEs may not be suitable due to shifts in composition and limited empirical support. Therefore, I also utilize parametric selection models to predict potential outcomes. My approach extends beyond traditional program evaluation, allowing estimation not only of program-specific returns but also the externalities caused by ripple effects in the entire market for higher education.

I find that ripple effects are large and heterogeneous across educational programs. Using reported preferences, I find that when 10 marginal applicants are let into a program, on average an additional 5.7 applicants shift into other programs, an externality of almost 60 percent. When I base simulations on applications constructed from my utility models, the externality grows to 260 percent. The size of the ripple effects varies across programs, with some programs affecting more than ten applicants every time they let one marginal applicant enter.

Ripple effects matter for policy evaluation. I find that between 15 and 21 percent of the variation in earnings effects is missed by ignoring ripple effects. These shares masks considerable heterogeneity among programs. Despite having similar educational content and estimated gains for marginal applicants, different programs exhibit varying total gains when accounting for ripple effects. I find a positive externality for 60 percent of programs, indicating that returns from expanding supply are generally underestimated. Some programs have positive marginal effects which are offset by ripple effects and vice versa.

A central question is to which extend researchers can use less detailed data to predict ripple effects. To address this, I construct a program network, where programs are linked based on their appearance on the same application list. Within this network, I calculate program centrality. I also record entrance thresholds in terms of GPA. Both measures of popularity serve as predictors for the number of additional applicants affected by a supply change. In addition, a higher degree of centrality and higher cutoffs correlate positively with higher-order returns resulting from ripple effects. This suggests that popular programs may often be under-supplied if programs set capacity themselves. GPA cutoffs, which are publicly available, can directly inform policymaking. However, it is important to note that the ripple effects on earnings exhibit considerable heterogeneity.

This paper highlights the significance of applicant sorting and rationing in the context of targeted supply-side policies. It emphasizes the need for policymakers to adopt a comprehensive approach when making decisions regarding the allocation of educational resources. Furthermore, the findings suggest that prestigious institutions may not fully consider the societal benefits when determining their capacity.

These results are derived from an analysis of the centralized college admission system in Denmark. However, the underlying qualitative insight regarding the importance of ripple effects extends beyond centralized markets and higher education. It serves as a specific example of how chains of substitution can either amplify or counteract the effects of targeted policies. Similar effects can be anticipated in various markets, both within and across countries, spanning different educational sectors as well as markets for housing or labor.

This paper is related to several different branches of economic research. Firstly, there is a large body of research focusing on the returns to tertiary education. Due to the nature of the educational systems, studies on American data typically focus on returns to institutions (Dale and Krueger, 2002; Hastings, Neilson, and Zimmerman, 2013; Zimmerman, 2014; Arcidiacono, Aucejo, and Hotz, 2016; Andrews, Li, and Lovenheim, 2016; Mountjoy and Hickman, 2020), Several European and South-American papers investigate returns to specific fields, see Altonji, Arcidiacono, and Maurel (2016) for a review. Among these, a number of studies employ regression discontinuity designs (RDD) to estimate returns to fields of study, either in terms of admission (Hastings, Neilson, and Zimmerman, 2013) or completion Kirkeboen, Leuven, and

Mogstad (2016) of the marginal applicant of a given program. Several papers have used such approaches to estimate returns to admittance or completion on Danish admission data (Heinesen and Hvid, 2019; Humlum and Meyer, 2020; Daly, Jensen, and Le Maire, 2020; Andersen, Hørlück, and Sørensen, 2020; Heinesen, Hvid, Kirkebøen, Leuven, and Mogstad, 2022).

A second related body of research finds its origin in the work on school choice by Gale and Shapley (1962) and Abdulkadiroğlu and Sönmez (2003) among others. Azevedo and Leshno (2016) develop large market asymptotics of stable matching mechanisms and show that they can be represented in a framework where cutoff structure plays the role of prices in clearing demand and supply. Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) and Abdulkadiroglu, Angrist, Narita, and Pathak (2022) use the same large market framework to construct a research designs for the estimation of causal effects exploiting that random tie-breaking moves applicants between programs. These articles utilize applicant switching in centralized mechanisms to generate reliable estimates of returns to programs. In my approach, I delve into a similar context but emphasize the significance of ripple effects in their own right for policy evaluation.

Several papers exploit the large market properties to estimate welfare in mechanisms where truth-telling is not a dominant strategy (Abdulkadiroğlu, Agarwal, and Pathak, 2017; Agarwal and Somaini, 2018; Kapor, Neilson, and Zimmerman, 2020). Contrary to these papers, my analysis of outcomes will be in terms of earnings, which is uncontroversial in the social welfare literature (Sandmo, 2015). Fack, Grenet, and He (2019) provide the approach that I use to estimate utility models in contexts where a truthful full ordering of programs is either not optimal or infeasible for the applicants. Although my primary focus is not explicitly on applicant welfare, the Danish mechanism enables me to gauge changes in welfare. As the mechanism encourages applicants to provide an honest (partial) ordering of programs, expanding capacity consistently leads to an increase in applicant welfare (and conversely, a decrease when capacity is reduced).

A few papers investigate applicant flows explicitly. Agarwal (2015, 2017) investigates the effects of changing capacities on allocations in the centralized medical residency matching market in the US and Bucarey (2018) investigates the crowding out of applicants in Chile when introducing financial aid. My paper differs from theirs in two significant ways. Firstly, the Danish system is fully centralized and I have access to all inputs as well as the algorithm. This means that simulations reflect realistic counterfactual scenarios of an entire educational market. Secondly, I link ripple effects directly to the traditional treatment effect literature and the returns to fields and compute the externalities which are not commonly considered in these literatures.[1]

The importance of ripple effects for policy interventions has been investigated by Manning and Petrongolo (2017) in the context of local labor markets where markets overlap geographically. The authors document that geographic ripple effects dilute the effect of local stimulus policies. In Danish higher education, I show that treatment effects are on average not attenuated but rather increased by the presence of ripple effects. In their working paper Kirkeboen, Leuven, and Mogstad (2016) briefly investigate the indirect effects of increasing capacity in sci-

[1]Additionally, Tanaka, Narita, and Moriguchi (2020) provide reduced-form evidence that crowding out occurred when introducing centralized admission to education in Japan.

ence programs using first-stage estimates of complier compositions from their fuzzy regression discontinuity framework and IV estimates of returns conditional on the next-best alternative. While the Scandinavian centralized setting is similar to this paper, my paper adds to their analysis in significant ways. Firstly, they consider one round of indirect effects at the field level and ignore longer chains of applicants. In contrast, I follow each applicant potentially through their entire rank-ordered list of programs including the possibility of non-assignment. Additionally, I allow for heterogeneity within fields and across applicants by parametrizing outcomes. I show that the program level focus is important as even within fields there are large differences in ripple effects that are missed by aggregation. Lastly, I estimate preferences, and therefore do not underestimate ripple effects due to truncated application lists.

Kline and Walters (2016) explicitly investigate the role of substitution for cost-benefit analysis in the context of the Head Start program in the US and show that accounting for the cost of the non-admitted compliers in other public programs can significantly alter the cost-benefit ratio of an expansion.[2] They briefly investigate the importance of rationed substitutes and conjecture that not accounting for rationing provides a lower bound on the rate of return to program expansion. In the context of Danish higher education, I show that this does not always hold. In estimating and predicting counterfactual outcomes, I draw inspiration from Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) where the authors investigate the role of value-added for high school choice in New York City.

The paper is structured as follows: Section 2 introduces the conceptualization of ripple effects within a theoretical framework. Section 3 provides an overview of the institutional context and the simulation model used. Section 4 presents the data used in the analysis and provides descriptive statistics. Section 5 focuses on the estimation of preferences. Section 6 presents the simulation results and quantifies the ripple effects by examining the number of affected applicants. Section 7 characterizes the ripple effects in terms of predicted changes in earnings and explores their relationship to program evaluation. Finally, Section 8 concludes the paper.

2 Ripple effects and policy evaluation in centralized mechanisms

I now construct a simple framework for policy evaluation in a market with ripple effects and show what the policy-relevant treatment effects are. I employ a stylized model of a centralized matching market. Similarly to Azevedo and Leshno (2016) I assume a continuum economy where a finite set of programs, P, indexed by p, is matched to a continuum of applicants. The capacity of a program is given as S_p. Applicants are defined by their type, $\theta = (\succ^\theta, e^\theta, Y_\theta, X_\theta)$. where $\succ^\theta$ is a strict preference ordering of programs and $e^\theta \in (0,1)^P$ is a vector of eligibility score, where programs prefer applicants with higher eligibility scores. $Y_\theta = \{Y_p^\theta\}$ is a vector of length P of potential outcomes. Finally, X_θ is a set of observable characteristics. Let η be the

[2] Feller, Grindal, Miratrix, Page, et al. (2016) in the same context find differential effects depending on the counterfactual allocation.

probability measure over the set of all types.

Azevedo and Leshno (2016) show that a stable matching in a continuum economy is unique and can be characterized by a vector of program cutoffs, $C \in (0,1)^P$ which equate demand and supply. The demand of applicant type θ, $D^\theta(C)$ is defined as her favorite program among those where her program-specific eligibility score, e_p^θ, exceeds the cutoff. The cutoffs can be thought of as market-clearing prices. The market is cleared using Student Proposing Deferred Acceptance (DA). In this mechanism it is the dominant strategy for applicants to report true preferences regardless of other applicants or supply and no applicant has justified envy towards other applicants (Abdulkadiroğlu and Sönmez, 2003).

To quantify the effects of supply changes on outcomes, assume that program p performs a marginal expansion of capacity, dS_p. Given that p is oversubscribed, the change in the number of admitted applicants to p equals the change in capacity. The vector C' contains the cutoffs in the stable matching using the new vector of supplies. Equating supply and demand it follows that:

$$\begin{aligned} dS_p = dD(C_p) =& \eta(\{\theta : D^\theta(C') = p, D^\theta(C) \neq p\}) \\ =& \int_\theta \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) \neq p\} d\eta(\theta) \\ =& \sum_{p' \neq p} \int_\theta \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) = p'\} d\eta(\theta). \end{aligned} \tag{1}$$

This formalization shows that marginal applicants come from other programs and that these applicants are heterogeneous in type. By moving into program p, the applicants will realize their potential outcome associated with p, Y_p^θ. The aggregate change in outcomes for this set of applicants, therefore, is given by:

$$\Delta_p = \sum_{p' \neq p} \int_\theta \left[Y_p^\theta - Y_{p'}^\theta \right] \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) = p'\} d\eta(\theta) \tag{2}$$

Dividing Equation (2) with (1) gives the average gain in outcomes of the marginal applicants associated with a marginal increase in program p, which is equivalent to the Local Average Treatment Effect (LATE, Imbens and Angrist (1994)):

$$LATE_p = \frac{\sum_{p' \neq p} \int_\theta \left[Y_p^\theta - Y_{p'}^\theta \right] \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) = p'\} d\eta(\theta)}{\sum_{p' \neq p} \int_\theta \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) = p'\} d\eta(\theta)} \tag{3}$$

Equation (3) shows that an IV estimator with a binary admission treatment will readily identify the treatment effect on compliers without observing types of applicants.[3] The irrelevance of knowledge of the margin of choice is a specific formulation for the centralized DA mechanism of the general identification results presented by Heckman, Urzua, and Vytlacil (2008).

[3]In other words, even without full rank-ordered lists and knowledge of alternatives, the IV estimate identifies the effect of admission on the margin. To see this, note that the complier share, as defined in the LATE framework in this context, can be obtained by summing over alternatives and integrating over types: $Pr\left(D_i(C') = p, D_i(C) \neq p\right) = \sum_{p' \neq p} Pr(D_i(C') = p, D_i(C) = p') = \sum_{p' \neq p} \int_\theta \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C) = p'\} d\eta(\theta)$.

This also formalizes the reasoning of Kirkeboen, Leuven, and Mogstad (2016) that the LATE of admission to program p is identified in a fuzzy regression discontinuity design (RDD) using threshold crossing as an instrument and controlling for the position in the waiting list in a DA-mechanism.[4]

Using Equations (1) and (3) the aggregate effect on the newly admitted to program p can then be formulated in terms of a LATE and a change in supply.

$$\Delta_p = LATE_p \times dS_p = LATE_p \times dD_p(C), \tag{4}$$

which shows that the LATE is the policy-relevant treatment effect (Heckman and Vytlacil, 2001). However, using only (4) to evaluate a supply change policy overlooks that not only does the demand of program p change, but the demands of other programs change as well. The link between changed supply in one program and changes in the entire vector of cutoffs reflects the substitution of applicants and is equivalent to a pecuniary externality in a standard supply and demand framework. To evaluate a policy of program expansion, an earnings-maximizing social planner computes the total gain, Δ^T, which is the sum of program-specific gains from a change in program p:

$$\Delta^T = \underbrace{LATE_p \times dD_p(C)}_{\text{First-order}} + \underbrace{\sum_{p' \neq p} LATE_{p'} \times dD_{p'}(C)}_{\text{Higher-order}}, \tag{5}$$

where $dD_{p'}(C)$ is the size of the inflow of applicants, i.e. from at the margin of non-admission to the program. In case of a program expansion, the DA mechanism ensures that all applicant weakly move to a more preferred program. For contractions all applicants move weakly to a less preferred program. For a program contraction, the LATEs should, therefore, be estimated for compliers on the margin of getting into a higher priority.[5]

Equation (5) shows that the informational requirement for policy evaluation using the LATE framework is substantial. The social planner needs knowledge of all changes in cutoffs and estimated LATEs for all programs where $dC_{p'}/dS_p \neq 0$. The large informational requirement is not a specific feature for centralized mechanism, as substitutes are often rationed (see Kline and Walters (2016) for an example using Head Start in the US).

As mentioned above, the LATE can be estimated with a fuzzy RDD. This design is valid under three assumptions; Independence of the instrument, an exclusion restriction, and a monotonicity condition. These assumptions are plausible within an admission system based on the Deferred Acceptance mechanism. Infinitely close to the program cutoff, an offer is as good as random, ensuring that independence of potential outcomes is satisfied. The exclusion restriction is plausible as the instrument reflects a decision rule in the algorithm – a program offer

[4]Kirkeboen, Leuven, and Mogstad (2016) argue that one needs to condition on the next-best program to obtain interpretable LATEs. This is only the case if one wants to uncover heterogeneity in returns. A standard fuzzy RDD is sufficient to uncover the LATE of interest in the context of marginal supply changes, as it implicitly weights compliers on the margin.

[5]Note that this property is informative of welfare. An expansion (contraction) of capacity will always weakly increase (decrease) welfare.

is the only plausible way that it might affect outcomes. Following Imbens and Angrist (1994), applicants can be separated into four sets; never-takers, always-takers, defiers, and compliers. Never-takers are those who would not be admitted regardless of crossing the cutoff or not and consists of two subsets: Those who do not apply to the program and those admitted to higher prioritized programs. The set of always-takers is empty.[6] By the same reasoning, the set of defiers is empty; crossing the cutoff can only make applicants (weakly) more likely to accept be accepted. The standard monotonicity condition for a just-identified IV is therefore plausible.

2.1 Alternative to LATE for policy evaluation when compliers are known

The usefulness of LATEs around cutoffs of admission for policy evaluation rests on supply changes being sufficiently small. However, for non-marginal changes in supply, the set of compliers might change. Contrary to regular IV setups, in this setting compliers can be identified in the data through simulation. It is, therefore, useful to reframe Equation (5) in terms of potential outcomes and add structure.

I assume that potential outcomes can take only positive values and assume the following parametric structure:

$$Y_p^\theta = \exp\left(X_\theta \beta_p\right), \tag{6}$$

where X_θ is a vector of observable characteristics of type θ and β_p are program-specific parameters that map the characteristics to outcomes. The aggregated effects of a supply change from equation (2) can then be written as

$$\Delta^T = \sum_p \sum_{p' \neq p} \int_\theta \left[\exp\left(X_\theta \beta_p\right) - \exp\left(X_\theta \beta_{p'}\right)\right] \mathbf{1}\{\theta : D^\theta(C') = p, D^\theta(C') = p'\} d\eta(\theta) \tag{7}$$

Relative to the formulation with LATEs outlined in Equation (5), this formulation trades off one kind of structure for another. The parametric structure is restrictive in that it selects the features that are relevant for predicting outcomes. Further, the program-specific parameters must be estimated. On the other hand, the complier compositions are computed directly by simulating the assignment mechanism. Given estimates of β_p, I can calculate all the empirical equivalents of the elements of Equation (7). In other words, I can implement non-marginal supply changes and calculate expected changes in outcomes as well as outcome levels across the whole distribution of applicants.[7]

My argument for computing counterfactual equilibria is based on several assumptions worthy of discussion. Firstly, I have so far assumed that I observe a complete ranking of programs for every type of applicant. However, this is rarely the case in practice and certainly not in the Danish context (see Section 3 for institutional context). Applicants often list only a subset of

[6] To see this, recall the definition of the cutoff as the minimum eligibility score of those admitted. Due to the lack of justified envy, a necessary condition for admission is crossing this cutoff. This rules out the existence of always takers.

[7] Note, however, that I assume β_p to be constant under counterfactual supply. I thereby assume away general equilibrium effects that affect payoffs. I discuss this assumption in the conclusion.

programs due to administrative restrictions of submitted rank-ordered lists or application costs. To address this, I estimate preferences by assuming a stable allocation and then construct a full ranking of programs in Section 5.[8] Secondly, I assume that the mapping from applicant observables to program-specific potential outcomes is fixed. This implies stable skill prices and rules out general equilibrium effects of changed supply in the labor market. While this assumption may hold for programs that offer transferable skills within the EU and internationally, it may be stronger for licensed occupations like law and medicine. Thirdly, I keep the outside option constant, which includes the option of delaying application. Consequently, I disregard dynamic effects such as applying earlier if the supply increases in a particular year. By ignoring the temporal dimension, I likely underestimate the ripple effect in terms of applicant flows. These assumptions limit the nature of the counterfactual equilibria that I can compute and I do not claim to capture all general equilibrium effects. Rather, the ripple effects analyzed in this paper are observed in the cross-sectional dimension and can be understood as one-time and possibly anticipated changes in supply.

3 Context and simulation of the assignment mechanism

In Denmark, tertiary programs are generally divided between short-cycle 2-year programs, medium-cycle professional bachelors (such as teaching), and long-cycle academic bachelors at universities (where the majority of graduates will proceed to complete a master's degree). Education is free and students receive generous grants and loan opportunities to cover living expenses. In line with most continental European systems, applicants apply to specific program-institution combinations and enrollments mainly occur in the fall semester. In general, programs set their own capacities though some programs with high unemployment rates of graduates are prohibited from expanding capacity. Programs receive funds from the central government based on completed coursework and graduation rates.

Allocation to tertiary education is administered by the Ministry of Higher Education and Science in a centralized allocation system (in Danish: *Den Koordinerede Tilmelding*). Each year the allocation round matches between 70 and 90 thousand applicants to around 800 programs. When applicants outnumber slots in a program, applicants are generally admitted in two quotas. Quota 1 (abbreviated Q1) admits applicants according to the grade-point average (GPA) from secondary school which is a combination of nationally standardized exams and continuous assessment during the term. An applicant will enter with the same GPA in the rankings in multiple programs and all applicants with a GPA or foreign equivalent can be admitted in Q1. The GPA is calculated with one decimal and a lottery number is used as a tie-breaker. An alternative is Quota 2 (abbreviated Q2), in which the ranking criteria are chosen by the educational institution under constraints set by the ministry. The most popular approaches in Q2 are combinations of specific course grades and CV requirements, though there is a lot of variation in these criteria. The ranking process is performed by the program admission offices and the

[8]If applicants are forced to have only a limited set of options, this corresponds to them correctly anticipating cutoffs knowing the changes in capacities and only report the relevant section of their lists.

ministry only observes the final ranking in Q2, see Gandil and Leuven (2022) for details.

Each applicant can provided a rank-ordered lists of up to eight programs. Under each program, the applicant can signal whether they want to be evaluated in Q2. If so, the applicant will have to supply further information to the program in order to be ranked. Additionally, the applicant can signal that they want to be evaluated for a "standby" slot both in Q1 and Q2. This system with quotas and standby means that while the applicant can only rank eight programs, the system can observe an applicant with more priorities.[9] In the mechanism, along with this modified rank-ordered list, an eligibility score is observed for each applicant in each relevant quota.

Program admission offices observe the program-specific application and not the remaining programs on the rank-ordered list of the applicant. This makes it very difficult for the admission offices to act strategically. Prior to allocation, the institutions report a capacity for each quota. The ministry runs a preliminary allocation, after which programs are allowed to increase, but not decrease their capacity, The mechanism is then cleared again using updated capacities. My data is from the final run of the mechanism.[10]

The allocation mechanism is a modified Student Proposing Deferred Acceptance (abbreviation: DA, Gale and Shapley (1962); Abdulkadiroğlu and Sönmez (2003)). In principle the DA is strategy-proof. However, in practice, specific features create strategic incentives. For instance, the limit on the number of programs on the rank-ordered list may force applicants to truncate their lists. However, only 3 percent of the applicants submit a full list, implying a low level of truncation at the bottom of the list. As noted by Calsamiglia, Haeringer, and Klijn (2010); Fack, Grenet, and He (2019), applicants might leave out unrealistic programs at the top of the rank-ordered list. Further, applicants might leave out programs below a safe option. I return to this issue in Section 5.

Each year the summary statistics of the final allocation are made available on the Ministry website. This information includes the number of students allocated to each quota and remaining slots. Admission cutoffs in Q1 in terms of GPA are published and treated as front-page news in the media. Appendix Figure A.1 show that cutoffs in Quota 1 tend to be stable from year to year, which allows applicants to infer admission probabilities from previous admission rounds. Cutoffs in Q2 are not informative as programs vary in their ranking function which is not generally known – nor in most cases formalized.

3.1 Simulating the assignment mechanism

With knowledge of the algorithm and production data, I re-engineer the entire mechanism for the allocation of tertiary education in Denmark. I manage to allocate over 98 percent of

[9]For example, a single application from the perspective of the applicant can entail 4 applications from the perspective of the mechanism: Q1, Q2, Q1-standby, and Q2-standby. If admitted to a standby-slot, and an admitted applicant rejects an offer, the offer is given to an applicant in the standby-quota according to the ranking. If not admitted, a standby slot guarantees an offer in the next academic year. Some programs also provide the option of enrolling in the Spring semester, which will sometimes be represented as an additional program in the mechanism.

[10]In order to not waste slots, the algorithm is nested within a loop where non-filled slots are transferred between quotas between each iteration of the algorithm. This means that the algorithm does not necessarily terminate, though the resulting matching is stable due to the properties of DA.

(a) Decrease of 10 percent (b) Increase of 10 percent

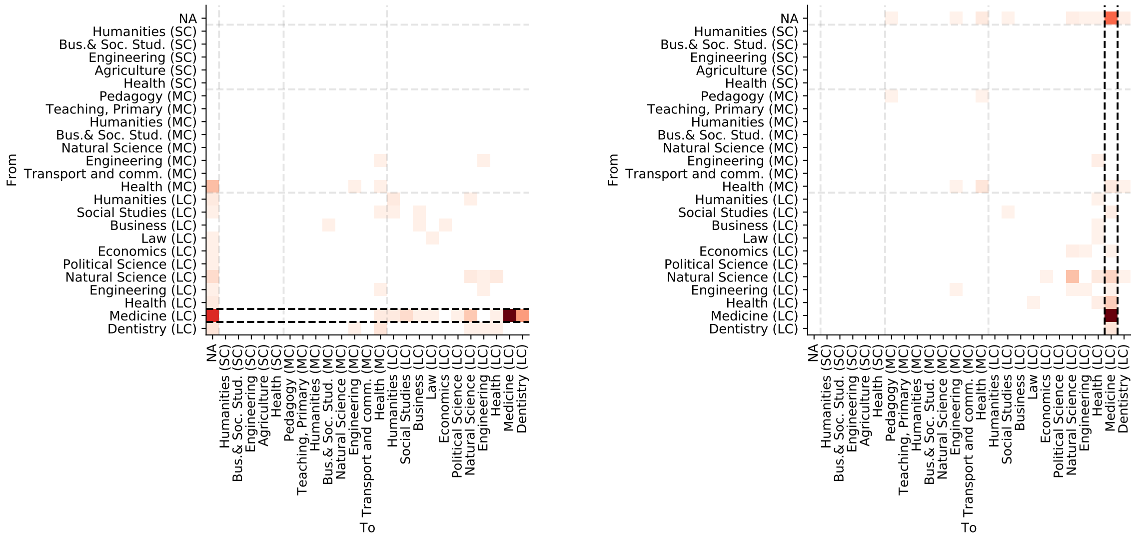


Figure 1: Applicant flows following a change in capacity for Medicine at KU

The simulations are based on a re-engineered version of the Danish assignment mechanism using 2016-data. In Figure 1a the number of slots in Quota 1 for Medicine at the University of Copenhagen is decreased by 10 percent. In Figure 1a the capacity is increased by 10 percent. Programs are grouped into field-length combinations, where the length can be short-cycle (SC, 1 to 3 years), medium-cycle (MC,3 to 4 years), and long cycle academic programs (LC, 3 years or more, academic bachelor programs). The dashed lines indicate the first order margin by field×length.

applicants correctly in numerous years.

To illustrate the simulation model, Figure 1 presents the result of changing the capacity of the Medicine program at the University of Copenhagen (KU) in the 2016 admission round. This program is heavily oversubscribed and has one of the highest GPA cutoffs in Quota 1. Figure 1a shows the flows resulting from a decrease in the number of slots in Quota 1 of 10 percent. The y-axis shows which combination of field and program length applicants come from, while the x-axis shows where they end up. The dashed line indicates the outflows from the long-cycle Medicine group, which includes the program at University of Copenhagen. The darker the square, the larger the flow.

The largest flow in Figure 1a is *between* Medicine programs. In other words, pushing applicants out of Medicine at KU pushes them into other Medicine programs. The second-largest flow is on the extensive margin, pushing applicants out of admission entirely. This partly reflects that some applicants only apply for a single program (and may apply again the following year.). Figure 1a also shows a large flow into Dentistry, indicating that these fields are substitutes. In turn, applicants are pushed out of Dentistry as evidenced by the flows in the lowest row. The same pattern can be seen for the medium-cycle health programs, which mostly consist of Nursing programs.

Figure 1b shows the result of an *increase* of capacity at the University of Copenhagen. While many flows occur along the same margins, this is not universally true. In other words, substitution patterns are complex and complier shares from a contraction do not equal the complier shares of an expansion.[11] Fortunately, my simulation model keeps track of these flows which accurately reflect what would have occurred had the supply been changed while holding demand constant.[12] Again, holding demand constant is a restrictive assumption, which I loosen by estimating preferences.

4 Data and descriptive statistics

Population data on income, grade-point averages (GPA), initiated educational spells and socio-economic variables are obtained from Statistics Denmark. Raw production data from the admission system is provided by the Danish Ministry of Higher Education. This data contains waiting lists for each program-quota combination, priorities of applicants and capacity, and the admittance outcome. Data is available from 1993, but due to lack of data simulation with a satisfactory level of precision is possible from 2016.[13] I construct two separate samples which I describe below. The first sample consists of a single admission round and is used for preference estimation and simulation, while the second sample is used to estimate potential outcomes and treatment effects.

Simulation sample The sample used for the simulations of capacity changes contains the population of applications in 2016. This is the same data I use for estimating preferences. I merge this data with national registries to retrieve GPA, courses and grades, residential location, gender and immigrant background.

For some applicants, I lack data on GPA rank, which I use to predict potential outcomes. For most applicants, an alternative GPA is reported in the production data. This GPA is including potential bonuses given due to a short time span between graduating high school or due to raising subjects to a higher level.[14] For foreigners, the GPA conversion follows the guidelines of the Ministry of Higher Education, though as this is done by the program administrators these GPAs are subject to error.[15] The corresponding GPA cohort ranks are imputed by a Random Forrest regression trained on those applicants where both GPA-rank and the alternative GPA are reported. For the subset of applicants where no variant of GPA is observed, I set the value of the GPA rank to 0.5.

To characterize the programs in my simulation sample, I record field and education length

[11]For marginal experiments, one could in principle use the first stages from IV regressions where the instrument is interacted with the next best field as approximating the flow shares as done in the working-paper version of Kirkeboen, Leuven, and Mogstad (2016). However, with discretionary slot allocation and varying program sizes, the definition of a marginal expansion is unclear in practice and first-stages are likely to be noisy.

[12]These simulations are performed for the 2016 admission round, which is the same round that I will perform simulations on in Section 7.

[13]Each program-year combination has a numeric identifier. Though large programs maintain the same identifier throughout the sampling period, many smaller programs change due to institutional merges and changes in program content. To maintain consistency over time I map the identifier to fields and institutions through the register data. If over 95 percent are enrolled in a given program, I link the two. I match the remaining program-year combinations manually. In the case of institutional mergers, I map the programs to the latest institution observed in the data.

[14]Applicants applying within two years of high school graduation can multiply their GPA with 1.08.

[15]For instance, the same applicant can be observed with multiple values of the alternative GPA across applications.

as well as GPA-cutoff in Quota 1.[16] Ripple effects occur because applicants apply to multiple programs. To capture this, I construct a weighted network graph of programs based on the preferences of the 2016 applicants. In this network an edge between two programs exist if the programs are listed by the same applicant. The edges are weighted by the number of occurrences. From this network I compute the weighted eigenvector centrality for each program. I standardize cutoffs and centrality to have mean zero and a standard deviation of 1. I note that the cutoffs along with program field and lengths are public information. Centrality is computed from the applications but does not require knowledge of the allocation mechanisms. Similar characteristics are therefore likely to be know by researchers in other institutional contexts.

Outcomes sample For estimation of outcomes, I take all applications from applicants in the admission data from 1993 to 2015 for whom I observe GPA, a Quota 1 waiting-list number, and positive income in the registers. The income concept is the log of average personal pre-tax income excluding transfers 7 to 9 years after application. With the exception of Medicine, no program is longer than five years, Thus the outcome measure should be interpreted as early career earnings. The combination of register availability and the log specification means that foreign applicants who do not migrate to Denmark are not observed in the data. Additionally, individuals with no income in all three years are excluded from the sample. As a measure of skill, I calculate the GPA rank within graduating cohort. For applicants where I do not observe the GPA (mainly foreign applicants), I impute it using a nearest-neighbor regression on the rank in Quota 1. In the control function specification, I exclude programs that do not have support in the running variable on either side of the cutoff. Further, I exclude programs where the admittance rate is less than 2 percent or higher than 98 percent. I do not impose this restriction in models where I do not control for selection.

4.1 Descriptive statistics

Table 1 presents descriptive statistics for the two samples. The unit of observation is the applicant, and the discrepancy between the number of applications and applicants reflects that on average applicants file 1.9 applications (419,420/217,792) in the outcomes sample. The second column shows that the applicant GPA rank in own cohort is slightly above the average in the applicant cohort with a mean value of 0.54. Around 11 percent enter with an A-level in business form high school, while 68 percent have an A-level in humanities (A-levels are not mutually exclusive). I do not record covariates such as gender and high school performance for 21 percent of the sample.

The following four columns display the corresponding statistics for the samples used for sampling correction. As I will describe in section 7, I estimate four different sets of control functions where I use cutoffs in Quota 1 or Quota 2 and exploit cutoffs in own program and higher prioritized programs. The sample size in terms of applications diminishes reflecting

[16]In programs where there is a single Quota 1 this is defined as the GPA of the last applicant admitted. In case a program has more than one Quota 1, it is recorded as the highest cutoff among these quotas. This implies that some applicants gain access through Quota 1 even if the have a lower cutoff than the one reported by the ministry.

Table 1: Descriptive statistics on samples

		Estimation sample	Program cutoff Q1	Program cutoff Q2	Higher priority cutoff Q1	Higher priority cutoff Q2	Simulation sample
log(Income 8 years after)	Mean	5.47	5.46	5.46	5.46	5.46	-
		(1.01)	(1.02)	(1.00)	(1.02)	(1.00)	-
GPA	Mean	0.54	0.57	0.49	0.54	0.49	0.51
		(0.28)	(0.27)	(0.26)	(0.27)	(0.26)	(0.27)
Female	Mean	0.60	0.63	0.65	0.62	0.66	0.58
		(0.49)	(0.48)	(0.48)	(0.48)	(0.47)	(0.49)
Danish grade	Mean	4.51	4.66	3.89	4.62	3.99	4.62
		(3.47)	(3.51)	(3.40)	(3.40)	(3.37)	(3.68)
Math grade	Mean	3.16	3.25	2.25	3.25	2.33	3.30
		(4.06)	(4.07)	(3.52)	(4.00)	(3.54)	(4.12)
Danish grade missing	Mean	0.02	0.03	0.01	0.03	0.01	0.02
		(0.15)	(0.16)	(0.11)	(0.16)	(0.11)	(0.13)
Math grade missing	Mean	0.22	0.23	0.23	0.23	0.22	0.20
		(0.42)	(0.42)	(0.42)	(0.42)	(0.42)	(0.40)
A-level: Business	Mean	0.11	0.10	0.09	0.09	0.08	0.12
		(0.32)	(0.30)	(0.28)	(0.29)	(0.27)	(0.32)
A-level: Humanities	Mean	0.68	0.70	0.62	0.72	0.64	0.64
		(0.47)	(0.46)	(0.49)	(0.45)	(0.48)	(0.48)
A-level: STEM	Mean	0.36	0.35	0.28	0.37	0.30	0.29
		(0.48)	(0.48)	(0.45)	(0.48)	(0.46)	(0.45)
Missing covars.	Mean	0.21	0.20	0.30	0.18	0.28	0.23
		(0.41)	(0.40)	(0.46)	(0.39)	(0.45)	(0.42)
Year	Mean	-7.31	-7.16	-7.42	-7.11	-7.30	0.00
		(1.73)	(1.77)	(1.83)	(1.75)	(1.82)	(0.00)
N applicants		217,792	141,023	55,063	85,537	30,838	82,794
N applications		419,420	253,517	85,153	156,254	47,110	204,900

The table reports summary statistics. The first five columns report statistics on the samples used for estimating outcome equations while the last column reports the statistics for the simulation sample. Standard deviations are reported in parentheses. Standard deviations are only shown for continuous variables. The year variable is centered at 2016.

that the research design requires binding cutoffs and I restrict attention to applicants close to these cutoffs. The average GPA rank is slightly higher in the Quota 1 samples with values of 0.57 for correction using own program cutoff and 0.54 using higher prioritized cutoffs. The simulation sample contains *all* 83 thousand applicants in 2016. Of these applicants, 43 percent apply to more than one program within the same fields, and, on average, applicants file 2.5 applications. This is higher than in the outcomes sample in the first column due to the lack of sampling restrictions in the simulation sample.

5 Estimation of preferences

As discussed above, applicants do not submit rank-ordered lists containing all programs. Firstly, applicants can only apply to eight programs. Secondly, with even a small (non-pecuniary) application cost, applicants will not list programs with zero perceived probability of admission. The reported rank-ordered lists (ROL) of programs are therefore likely too short to capture the full extent of ripple effects. Instead of using reported ROLs, I model and estimate preferences and use these estimated preferences to rank all programs creating new rank-ordered lists.

I classify applicants according to strata, s, which are constructed as combinations of geographical regions, immigrant background and gender. I assume that applicant i, who belongs to strata s receives the following utility of admission to program p:

$$U_{ip} = V(W_{ip}, \gamma_s) + u_{ip}, \tag{8}$$

where V is a known function of observables, W_{ip}, and an unknown strata-specific vector, γ_s, to be estimated. W_{ip} contains dummies for field, program length and location (administrative region). To capture heterogeneity within strata, program length is interacted with high school GPA and program field is interacted with applicant characteristics such as GPA and dummies for A-levels in STEM, humanities, social studies and business. I assume that u_{ip} follows a standard type-1 extreme value distribution. The outside option is included as a separate program. This specification allows the outside option to vary within strata (which capture geography, gender and ethnicity) according to GPA and A-levels in high school.

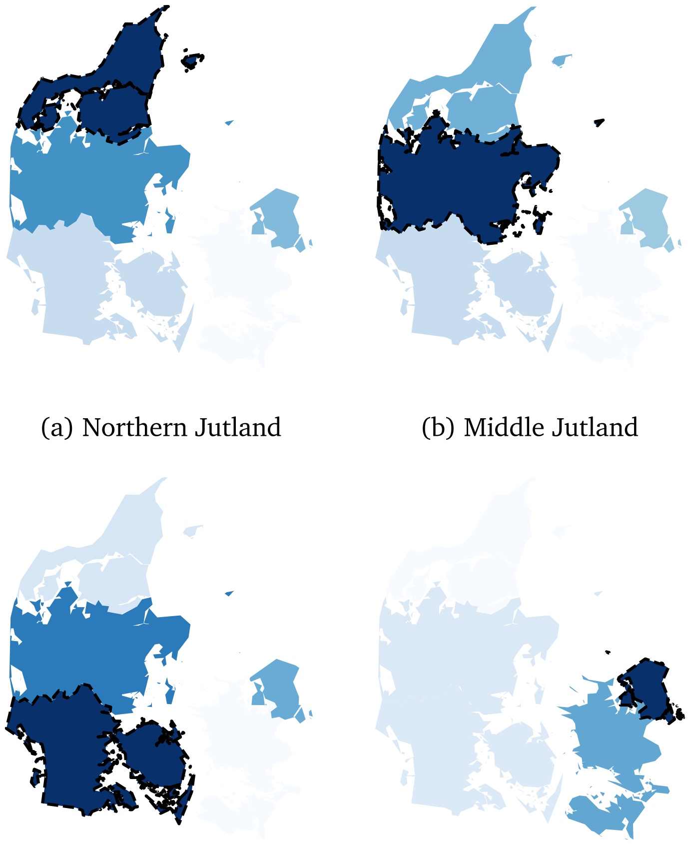
If applicants where truthful, equations (8) could be estimated using rank-ordered logit (also known as exploded logit). However, as applicants likely truncate their lists both from above and below, I instead rely on the stability-based estimator outlined by Fack, Grenet, and He (2019), which assumes that applicants get admitted to their favorite program within a set of feasible programs. The stability-based approach estimates a conditional logit on the ex-post feasible choice set of programs.[17] Chrisander and Bjerre-Nielsen (2023) show that the stability-based approach provides a good approximation of elicited preferences from surveys. To flexibly model preferences, I estimate the model separately within each strata. With the estimated model parameters I predict utilities of all programs for all applicants and construct new rank-ordered lists. I truncate these lists at the position of the outside option as the applicant will never be admitted to a program below this position in the list.[18] Appendix B provides further details on specification and estimation.

Predicted utilities and imputed rank-ordered lists The random utility model in Equation (8) is estimated within strata and the estimated coefficient are reported in a data set in the online appendix.To assess the credibility of a subset of the estimated parameters, Figure 2 plots the coefficients on region fixed effects within each strata for Danish males. The home region is surrounded by dashed lines. I find a strong home bias as the home region has the highest fixed effect within all strata. Neighboring regions also tend to be more attractive than other regions. The Copenhagen metropolitan region (the Eastern-most region) is relatively popular regardless of distance, reflecting that this region is the largest urban region in the country and hosts a large number of educational institutions.

Using the estimated utilities for each applicant-program combination, I construct rank-ordered lists which I truncate at the outside option as described above. The results are shown

[17]Feasible programs are defined as those programs where applicants would have gotten in based on their GPA in quota 1, had they applied. I also include Q2-programs to which the applicant applied. However, the process of ranking in Quota 2 is unknown, see Gandil and Leuven (2022). I therefore do not know the counterfactual rank of an applicant in Quota 2 had she applied.

[18]For some applicants, the predicted utility of the outside option is higher than for the programs in the actual rank-ordered list. If this was true applicants wouldn't apply to these programs. I therefore include these programs in the new rank-ordered lists. The model fails to converge for immigrant women in Copenhagen Metro. For this strata I instead include their original reported preferences.



(a) Northern Jutland (b) Middle Jutland

(c) Southern Denmark (d) Copenhagen Metro

Figure 2: Estimated regional preferences

The maps show coefficients on region fixed effects for Danish males by region of residence. The region of residence is identified by dashed borders. The higher the fixed effect, the darker the color. The coefficients are based on conditional logits with personalized feasible choice sets as described in Section 5. Each map shows the results within a single strata. The island of Bornholm is excluded from the maps. The coefficients for the Sealand region is not shown here due to space constraints. Applicants from Sealand also has a strong home bias but it is rivaled by the Copenhagen metropolitan region which is close and where educational institutions are clustered.

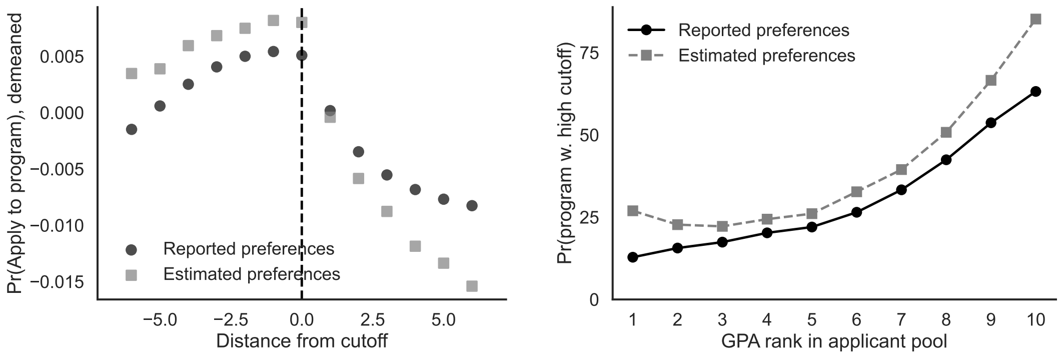
in Table 2. As expected, the imputed rank-ordered lists are longer than the reported ranks. The average length of a list goes from 2.5 programs to 18.5 programs. The allocation mechanism creates an incentive to report the correct (partial) ordering of programs. In the imputed rank-ordered lists, the correct ordering of reported programs is maintained in 58 percent of the cases. This suggests that the random utility modes approximates preferences well and corroborates the findings of Pathak and Shi (2021).

To assess whether top truncation of rank-ordered lists is an issue, Figure 3a shows the probability of applying to a program as function of GPA distance to cutoff. A negative distance means that the applicant's GPA is below the cutoff in Quota 1. The black dots show results for reported preferences. Below the cutoff there is a clear drop in the probability of applying. Using estimated preferences, I find a smaller drop. This is suggestive evidence that applicants with low probability of acceptance abstain from applying. To further assess top-truncation, Figure 3b shows the probability of applying for a highly competitive program as a function of GPA. Because original applications are included in the imputed lists, the probability is mechanically

Table 2: Imputed Rank-ordered lists

	Mean	Min	Q25	Q50	Q75	Max
Lenght of reported ROL	2.48	1	1	2	3	8
Lenght of estimated ROL	18.46	1	2	5	5	759
Correct ordering | Length of ROL>2 (mean)	0.58					

The table shows descriptive statistics of reported ranked programs and the estimated rank-ordered lists.



(a) Apply to a program

(b) Apply to competitive program

Figure 3: Evidence of top truncation of preferences

Figure 3a shows the probability of applying to a program as function of GPA distance to cutoff. The black dots represent applications submitted, whereas the gray squares represent applications predicted by random utility models. The two series are demeaned for comparability. Only oversubscribed programs with Quota 1 are included. Figure 3b shows the share of applicants with a GPA-decile who report a preference for a program with a high-GPA cutoff using reported ranked programs and the estimated rank-ordered lists. Because all actual applications are included in the lists based on utility estimates, the share is mechanically higher.

higher. However, the figure shows a smaller gradient for imputed lists than for the reported lists. Especially low-GPA applicants are more likely to apply for competitive programs than reported preferences would suggest.

6 Ripple effect in flows of applicants

I now proceed to investigate the ripple effects by simulating counterfactual assignments. I do this with the reported rank-ordered lists and the filled-out rank-ordered lists based on estimated preferences. In this section, I focus on ripple effects in terms of flows of applicants moving between programs (and the outside-option). In section 7 I link these flows to returns in earnings.

6.1 Definition of experiments

I simulate a change in the supply of Quota 1 slots for programs with more than 50 slots in Quota 1 in the application round of 2016. For each program, I manipulate slots to be between 95 percent smaller or larger than baseline. To avoid an overflow of applicants from Quota 2 into Quota 1, I disable transfers of slots between quotas within the same program. Applicant can however still move between quotas within the same program. In total, I perform 11,210 simulations for 295 programs. I do this twice: once with reported preferences and once with estimated preferences. To keep comparisons consistent, I simulate two baseline allocations, one for each set of preferences.

6.2 Flows of applicants

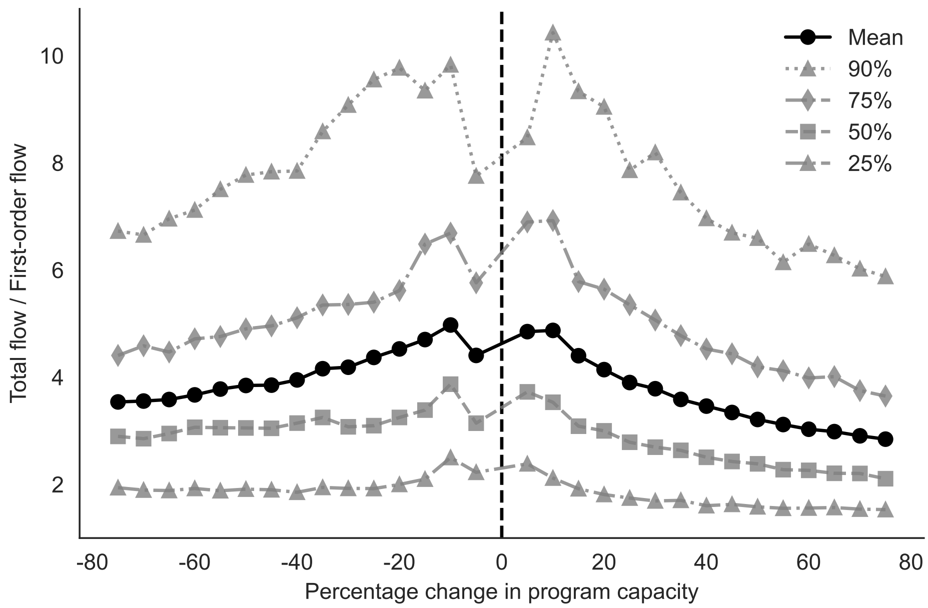
For each experiment, I distinguish first-order moves and higher-order moves. First-order moves are applicants moving into or out of the program where supply is altered. Higher-order moves are all moves that do not occur at this margin.

I begin by computing the ratio of total applicants affected relative to applicants affected on the margin of the altered program for each simulation. Figure 4a shows the distribution of the number of applicants affected relative to applicants shifted on the program margin. In average across the experiments, 36 applicants are affected in total every time 10 applicants are affected on the program margin. In other words, for each 10 applicants let into a program, 26 applicants are affected elsewhere, an externality of 260 percent. The dispersion is large, with some experiments inducing more than ten times as many indirect shifts than direct shifts. In other words, depending on the program, a change in supply may induce a sizable externality on other programs. The magnitude of the ripple effects tapers off slightly as the capacity changes become larger, reflecting that programs throughout the system eventually run out of applicants.

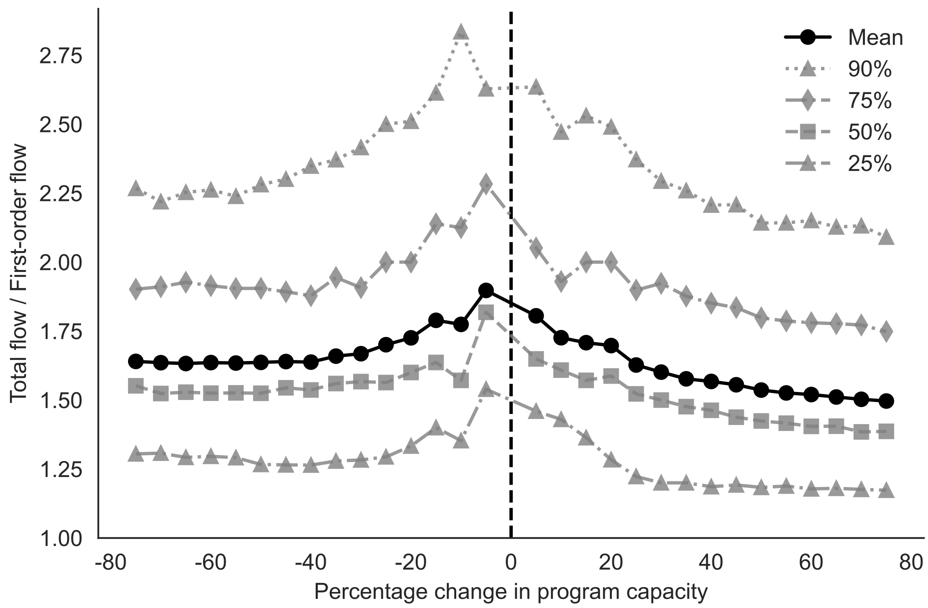
Figure 4b shows the same graph but with simulations based on reported preferences. Though the patterns are similar, magnitudes are generally smaller. On average, when a counterfactual supply induces 10 applicants to change program, this is associated with 15.7 applicants changing allocations. This is expected as reported rank-orded lists are generally shorter then the imputed lists. Holding applications constant thus may severely underestimate the magnitude of the ripple effects in terms of affected applicants.

Popular and oversubscribed programs, such as Business programs at Copenhagen Business School (CBS) and Medicine and Political Science at the University of Copenhagen (KU), tend to create significant externalities. Figure 5 provides examples of these programs. This aligns with expectations, as applicants to popular programs often include a safety program in their ranked lists, which triggers the ripple effect. The extent to which a program is considered a safety option depends on the applicant's GPA. Therefore, what may be a safety program for one applicant can be a highly competitive "reach" program for another. As the supply of a program decreases, marginal applicants are pushed towards their safety program, displacing applicants from the safety program and causing the ripple effect. In cases of supply increases, the dynamics are similar, with applicants transitioning from their safety program to higher-priority choices.

The results demonstrate that ripple effects are more pronounced when rank-ordered lists are filled in compared to using the reported rank-ordered lists. To compare these two approaches, I regress flow ratios (Total/First-order flow) obtained from estimated preferences against those obtained from reported preferences. The findings are presented in Figure 6. I observe a strong positive association of 0.69, and the conditional expectation seems to be well-approximated by a linear function. This suggests that using reported preferences provides a good indication of the relative magnitude of ripple effects between programs when applicants are allowed to adjust their behavior. However, as depicted by the dashed gray line, ripple effects are much larger with estimated preferences.



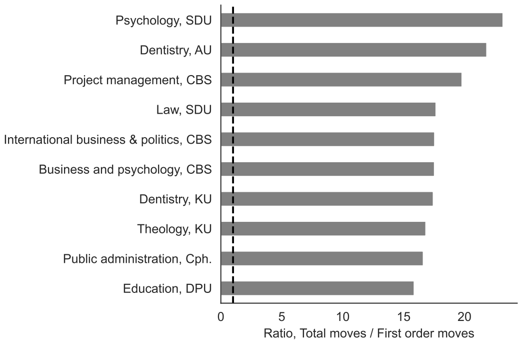
(a) Estimated preferences



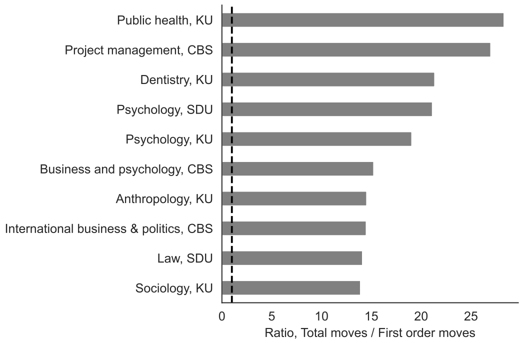
(b) Reported preferences

Figure 4: Total flow relative to first-order flow

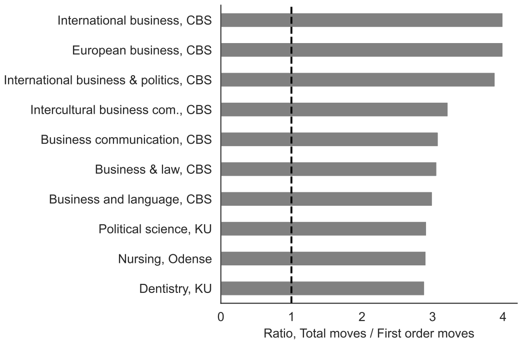
The figure plots the ratio of the total number of movers relative to the first-order movers as a function of percentage change in baseline capacity. The unit of analysis is a simulation of a program change. Simulations where less than 10 applicants are affected on the first-order margin are left out due to anonymity requirements.



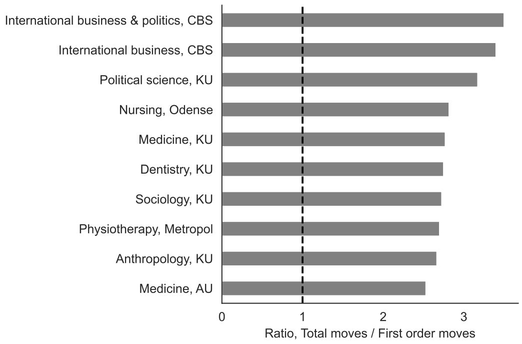
(a) Estimated pref: 10% reduction in capacity



(b) Estimated pref: 10% increase in capacity



(c) Reported pref: 10% reduction in capacity



(d) Reported pref: 10% increase in capacity

Figure 5: Programs with the largest ripple effects measured by flows of applicants

Figures 5a)-d) plots the ratio of the total number of movers relative to the first-order movers for a ten percent change in capacity. The programs shown are the ones with the largest ripple effects and which are sufficiently large to fulfill anonymity requirements set by Statistics Denmark. Panels a) and b) show the results based on reported preferences, whereas panels c) and d) are based on estimated preferences.

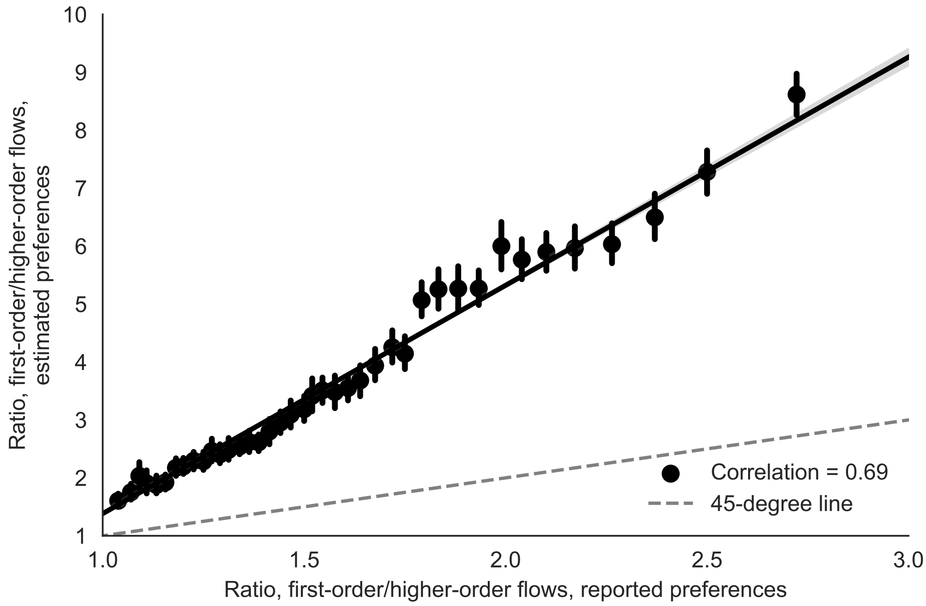


Figure 6: Comparing ripple effects with and without estimated preferences

The figure plots the ratios of Total flow relative to first-order flow of applicants on a program level. The value on the x-axis is using reported preferences, and the value on the y-axis is using estimated preferences. Observations are binned on the x-axis, while the regression line is fitted on the raw data. Confidence-intervals are on 95%-level.

The analysis conducted above leverages the production data obtained from the centralized allocation system, which provides a rare richness of information compared to other contexts. However, it is important to note that ripple effects can occur irrespective of whether the system is centralized or not. Therefore, it becomes crucial to explore the predictability of ripple effects and determine which programs are more likely to exhibit them. To investigate this predictability, I regress the ripple effects on various program characteristics. As the dependent variable, I once again use the flow-ratio, averaging it within each program across simulations. I then regress this average ratio on the eigenvector centrality and the GPA cutoff in Quota 1. The results of these regressions are presented in Table 3. Panel a) displays estimates based on ratios derived from simulations with estimated preferences, while panel b) presents estimates based on reported preferences.

Both centrality and cutoffs show a strong positive association with larger ripple effects. As evidenced in column (1) of panel a), an increase of one standard deviation in centrality is associated with an increase in the ratio of 1.22 corresponding to 35 percent increase relative to the average. An increase of the GPA cutoff by a standard deviation predicts an increase of the ratio by 53 percent (1.8/3.39) as evidenced in column 3. Centrality is less predictive than cutoffs as measured by R-squared. These conclusions holds even when including institution and program fixed effects as seen in columns 2 and 4. When including the two measures jointly, the predictiveness of centrality attenuates, while the parameters on the GPA cutoff remain essentially unchanged.[19] Panel b) shows the results using reported preferences. The magnitudes are

[19]The correlation between the two measures as shown in Appendix Figure A.8.

Table 3: Predicting ripple effects in flows

	Panel a) Flow ratios, estimated preferences					
	(1)	(2)	(3)	(4)	(5)	(6)
Eigenvector centrality	1.22	0.66			0.41	0.02
	(0.10)	(0.13)			(0.12)	(0.20)
GPA cutoff			1.80	1.21	1.59	1.20
			(0.31)	(0.10)	(0.12)	(0.20)
Constant	3.40	3.40	3.39	3.39	3.39	3.39
	(0.27)	(0.02)	(0.37)	(0.02)	(0.10)	(0.03)
Observations	293	292	293	292	293	292
R2	0.23	0.62	0.50	0.70	0.52	0.70
Field FE		X		X		X
Institution FE		X		X		X

	Panel b) Flow ratios, reported preferences					
	(1)	(2)	(3)	(4)	(5)	(6)
Eigenvector centrality	0.20	0.16			0.02	0.00
	(0.03)	(0.03)			(0.02)	(0.02)
GPA cutoff			0.36	0.30	0.35	0.30
			(0.03)	(0.03)	(0.02)	(0.03)
Constant	1.51	1.52	1.51	1.51	1.51	1.51
	(0.04)	(0.00)	(0.03)	(0.00)	(0.01)	(0.00)
Observations	295	294	295	294	295	294
R2	0.21	0.60	0.70	0.79	0.70	0.79
Field FE		X		X		X
Institution FE		X		X		X

This table show estimated parameters for OLS regressions. The dependent variable is the average ratio of total affected applicants to applicants affected at the margin of the manipulated program. The ratio is average within manipulated program across simulations. The top panel uses ratios from simulations based on estimated preferences. The lower panel uses simulations based on reported preferences. Both Eigen-vector centrality and cutoffs are standardized to have a mean of zero and standard deviation of 1 across programs. Standard errors are in parentheses and clustered on the level of the fixed effects

generally smaller, but the overall conclusion the cutoffs are good predictors is maintained. The functional relationships in Table 3 are shown in Appendix Figure A.2, where it seems that the expectation is well-modeled by a linear function.

7 Program evaluation with ripple effects

In this section, I assess the importance of ripple effects for policy evaluation. To accomplish this, I compute estimates of the local average treatment effects (LATEs) and predicted potential outcomes. These estimates are then used to calculate changes in earnings as a result of local supply changes and the ensuring ripple effects.

7.1 Estimating LATEs for policy evaluation

7.1.1 Non-parametric LATEs

As outlined in Section 2, the relevant treatment effects can be estimated in a Fuzzy Regression Discontinuity Design using eligibility score as a running variable. The endogenous treatment is a program offer. In the data, I observe the ranking of each applicant in each quota. For each program in present in the 2016 round where I have sufficient empirical support around the cutoff in my outcomes sample, I compute program level LATEs using a standard Fuzzy Regression Discontinuity Design.

$$Y_{ip} = \alpha_0 + \delta_p D_i + f(e_{ip}) + X_i \gamma_p + \epsilon_{ip}, \tag{9}$$

where Y_{ip} is average earnings (in DKK) and e_{ip} is the ranking within a quota centered around the cutoff. The function $f(\cdot)$ is a linear spline with a not at zero. D_i is a dummy for admission and it is instrumented by crossing the cutoff in the running variable. X_i is a vector of individual level covariates including year, high school GPA rank within graduating cohort, math and Danish grades, dummies for A-levels and a gender dummy.[20]

In the presence of heterogeneous treatment effects, program specific LATEs, represents the earnings gain for applicants at the threshold who do not get admitted into a more preferred program. I estimate (9) for each program in my sample which has a cutoff. Though restrictive, the linear spline allows me to estimate LATEs for small programs. However, estimates are still noisy and I therefore shrink the estimates using Empirical Bayes, where I shrink the estimates towards the global mean.[21]

As discussed in section 2, for a given capacity change, all applicants either move (weakly) up or down their rank-ordered list. When capacity is increased, this implies that marginal applicants will be located around the cutoff of the program. However, when capacity is decreased, applicants move down their priority list. The marginal applicant is now located at the cutoff of their higher priority and this is the group for which I need to estimate a LATE. I estimate the program-specific LATEs for program contractions using variation in cutoffs for higher ranked alternatives, similarly to Humlum and Meyer (2020).For under-subscribed programs I cannot estimate a LATE at the programs own cutoff and I instead set it to zero. However, most of these programs will have a LATE for program contractions due to the presence of more preferred programs on the applicants' rank-ordered lists.

7.1.2 Estimating potential outcomes

As shown in Section 2, an alternative to using LATEs and applicant flows, is to model potential outcomes as a function of observables and extract the counterfactual composition of admitted applicants from the simulations. I now present the methods used to obtain estimates of the parameters governing the expected potential outcomes. As in Equation (6) I assume that the potential outcomes are linked to programs via the following log-linear model:

$$\log(Y_{ip}) = y_{ip} = X_i \beta_p + \varepsilon_{ip}, \tag{10}$$

[20] These covariates are not necessary for estimating the LATE but increase precision. They are included to maintain comparability with the control function approach which I outline below.

[21] This Empirical Bayes procedure does not account for heterogeneity in complier composition. In case of treatment effect heterogeneity this may be an issue. It is not obvious what the target of shrinkage is, when complier compositions vary across programs. As an alternative to this form of shrinkage, I circumvent this issue with the control functions, where I model heterogenity as a function of observables.

where y_{ip} is the log of a positive scalar outcome of interest of individual i admitted in program p and X_i is a vector of observable and predetermined covariates while ε_{ip} is unobserved. The vector of observables, X_i, are the same as in (9) and contains a second-degree polynomial in the GPA-rank in the applicant's high-school cohort, A-levels and a linear year trend. The outcome y_{ip} is only observed for applicants admitted to p. Taking the expectation of equation (10) for those admitted yields:

$$E[y_{ip}|X_i, i \in p] = X_i\beta_p + E[\varepsilon_{ip}|X_i, i \in p]. \tag{11}$$

I estimate β under different assumptions on the unobserved component, $E[\varepsilon_{ip}|X_i, i \in p]$.

If applicants sort into programs based on the unobservable part of the potential outcomes such as in the Roy model (Roy, 1951), Ordinary Least Squares (OLS) estimates of β_p will be biased. To control for selection I once again exploit the variation around the cutoffs for admission. As an instrument for admission in program p, I define a binary instrument for crossing the cutoff in a quota q in program p, $Z_i^{pq} = \mathbf{1}\{e_{iq} \geq 0\}$, where I center the running variable, e_{iq} around the cutoff in quota q in the year. Conditional on the running variable, e_{iq}, Z_i^{pq} is a valid instrument for D_i^p. I specify the following selection equation:

$$D_i^p = \mathbf{1}\left\{\gamma Z_i^{pq} + f\left(e_{iq}\right) + X_i\eta + \epsilon_{iq} > 0\right\}, \tag{12}$$

where $f(\cdot)$ is a linear spline with a knot at zero. The inner function on the right-hand side is identical to the first stage in fuzzy RDD specification presented in Equation (9).

Assuming that $(\varepsilon_{ip}, \epsilon_{iq})$ are jointly normal-distributed, the model in Equation (11) becomes a standard Probit selection model:

$$E[y_{ip}|X_i, i \in p] = X_i\beta_p + \psi_p\lambda\left(\hat{\gamma} Z_i^{pq} + \hat{f}\left(e_{iq}\right) + X_i\hat{\eta}\right), \tag{13}$$

where $\lambda(\cdot)$ is the inverse Mills-ratio and $\psi_p = corr(\varepsilon_{ip}, \epsilon_{iq})\sigma_\epsilon$.

This model embeds the identifying variation from the regression discontinuity design in a switching regression framework and imposes structure through equation (11). To include the instrument, Z_i^{pq} in equation (13) and avoid identifying off functional form and the running variable alone, there needs to be applicants on both sides of the cutoff. Similarly to the fuzzy RDD-approach, this implies that the control function methods are only feasible for oversubscribed programs or in cases where applicants are restricted from admission into higher prioritized programs.[22]

Due to the multi-quota nature of the admission system and the specification of the instrument, there is two-sided non-compliance; Firstly, applicants may have $Z_i^p = 0$ but be over the threshold in another quota within p. This implies that applicants below the threshold gain admission.[23] Secondly, applicants above the threshold might also be eligible to a higher ranked

[22]As the running variable is just a rescaling of GPA (In Quota 1) within the cohort of applicants, I exclude the running variable from the structural equation. For each program, I stack application years where the program is oversubscribed.

[23]This contrasts with the theoretical model outlined in Section 2. The theoretical outline did not allow for always-takers because each program had one quota. In the empirical section I observe always-takers because I only model

program and thus get admitted there. The two-sided non-compliance imply a correspondence between the selection model and the LATE estimates from a fuzzy RDD. To validate the selection model, following Kline and Walters (2019), I calculate the LATE of being admitted to p in the selection models and compare it to the LATEs from the fuzzy RDD. The methods for calculating the LATE in the control function approach is outlined in Online Appendix C.

As in the fuzzy regression discontinuity design, I estimate the control functions using variation around the programs own cutoff as well as cutoffs for higher-ranked alternatives. I do this for cutoffs in Quota 1 (GPA) and Quota 2 (other criteria).

For the programs where I lack support for estimating the control functions, I instead estimate the necessary parameters under the assumption that the expected value of the unobserved component conditional on covariates is zero, $E[\varepsilon_{ip}|X_i, i \in p] = 0$:

$$E[y_{ip}|X_i, i \in p] = X_i\beta_p \tag{14}$$

Under this conditional independence assumption, an unbiased estimate of β_p can be obtained by OLS.

For each program, I estimate the potential outcomes of those admitted. For a subset of programs, it is not feasible to estimate program-specific parameters in the estimation sample as these programs are too new for applicants to realize outcomes. For predicting these programs I estimate field-specific versions of equation (14). Standard errors for the control functions are obtained from the two-step variance estimator derived by Heckman (1979). For the OLS estimates I compute heteroskedasticity-robust standard errors.

Joint distribution of parameters, shrinkage, and prediction Under appropriate assumptions, the two sets of estimates of β_p are unbiased but noisy measures of the true parameters. This noise is especially an issue with the control function method on small programs with correspondingly small samples. As the estimates are needed for predictions, I follow Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) and apply an empirical Bayes shrinkage estimator to the estimates, thereby reducing the mean squared error of the predicted outcomes (Robbins, 1992; Morris, 1983). I assume the following hierarchical model:

$$\hat{\beta}_p|\beta_p \sim \mathcal{N}(\beta_p, \Omega_p) \tag{15}$$

$$\beta_p \sim \mathcal{N}(\mu_\beta, \Sigma_\beta) \tag{16}$$

Estimates of μ_β and Σ_β are obtained through maximum likelihood estimation of (15) and (16), where Ω_p is approximated by the estimated covariance matrices of the estimation methods described above. For the control function estimates, I include ψ_p in the vector of parameters. The resulting estimates of the hyperparameters, μ_β, and $\hat{\Sigma}_\beta$, are in turn used to obtain empirical

one type of quota at a time.

Bayes posterior means for β_p:

$$\beta_p^* = \left(\hat{\Omega}_p^{-1} + \hat{\Sigma}_\beta^{-1}\right)^{-1} \left(\hat{\Omega}_p^{-1}\hat{\beta}_p + \hat{\Sigma}_\beta^{-1}\hat{\mu}_\beta\right),$$

where $\hat{\Omega}_p$ is approximated by the covariance matrix of $\hat{\beta}_p$. In essence, this procedure shrinks imprecise estimates towards the mean of the estimated coefficients. The procedure is described in detail in Appendix D.

The predicted outcomes in the simulation sample are calculated as

$$\hat{Y}_{ip} = \exp\left(X_i\beta_p^*\right). \tag{17}$$

For each program, I also obtain a predicted potential outcome if not admitted. For the last program on the rank-ordered list of an applicant, this predicted outcome is interpreted as the outcome in the case of not being admitted to any program. In addition to improving prediction, the hyperparameters are also informative on the joint distribution of parameters in the educational production function and by comparing the hyperparameters across the estimation methods, I can investigate the importance of correcting for selection.

7.2 Estimated returns to programs

Assessment of validity and aggregate effects Figures 7a and 7b show balancing of predetermined variables of parental income and high school GPA. I find no evidence of effects of sorting around the cutoff.[24] The first stage, displayed in Figure 7c, is large at 0.4. In other words, crossing the threshold increases the probability of admittance by 40 percentage points. The level of admittance conditional on crossing the cutoff is around 60 percent. This reflects that applicants will only be admitted to a program if they have not been admitted to a higher prioritized program. The reduced form estimate is 0.031, which means that crossing the threshold is associated with an increase in income of approximately 3.1 percent, though the discontinuity is not visually convincing. A fuzzy regression discontinuity estimate of the application-weighted LATE of admittance is 7 percent. Though this is weak evidence of positive returns to the preferred field, this estimate masks considerable heterogeneity, as I will return to below. Additionally, the gains reflect early-career outcomes and do not take into account that lifetime income profiles may differ substantially across programs.

Appendix Figure A.3 shows the distribution of program specific LATEs estimated as described in the previous section. The spread is large but centered close to zero. Appendix figure A.4 shows examples of programs with high and low LATEs. STEM programs are among the top, however they are also found in the bottom of the distribution. It is important to underscore the that the estimates are for compliers around the cutoff. Cutoffs can be located at many different points in the GPA distribution and estimates are not readily comparable as the programs vary in both quality and complier composition. This motivates estimating potential outcomes using

[24] Because I use the position in the waiting list as the running variable, the densities are by definition uniform. Statistics on possible density manipulation would therefore solely reflect composition effects by stacking programs. I, therefore, do not report these statistics.

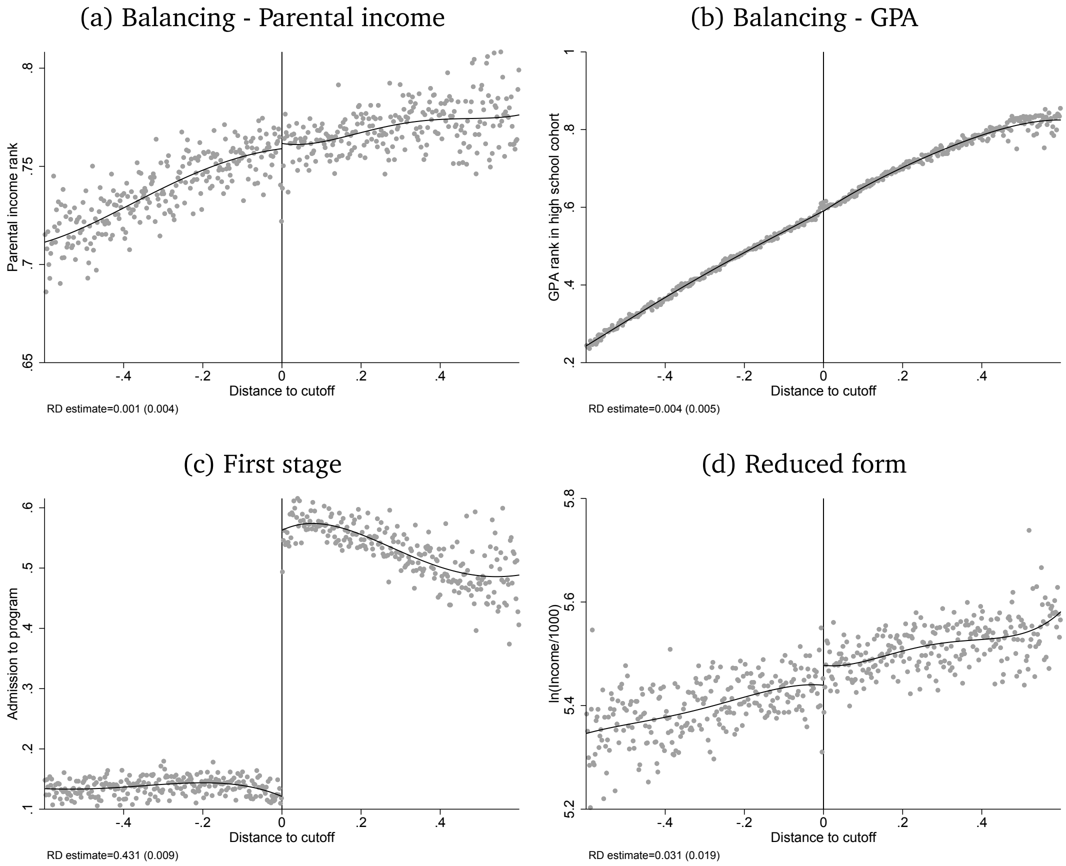


Figure 7: Regression discontinuity plots

Effects of crossing cutoffs are estimated using the rd-robust package in Stata with standard bandwidths. For the graphs, I subtract the program mean by regressing the dependent variable on a program fixed effect. Binned scatter plots are plotted with 250 bins on each side of the cutoff. The effects are estimated in the full sample across all program-year combinations and due to composition effects slopes should be interpreted with caution. For the IV estimation, observations exactly at cutoff are removed from the sample. Standard errors are clustered at the individual level as applicants can appear multiple times both across programs and years.

Table 4: Joint distribution of estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.042	0.035	0.047	5.258	0.038	5.153	0.028	5.412
	(0.077)	(0.092)	(0.071)	(0.321)	(0.105)	(0.611)	(0.071)	(0.339)
GPA	0.189	0.244	0.191	0.061	0.182	0.700	0.195	0.328
	(0.294)	(0.240)	(0.216)	(0.350)	(0.201)	(1.851)	(0.232)	(0.611)
GPA^2	0.242	0.278	0.238	0.050	0.233	-0.501	0.213	-0.215
	(0.262)	(0.259)	(0.271)	(0.388)	(0.276)	(2.047)	(0.280)	(0.598)
Female	0.080	0.087	0.094	-0.116	0.075	-0.151	0.053	-0.170
	(0.209)	(0.265)	(0.199)	(0.084)	(0.168)	(0.121)	(0.185)	(0.093)
Danish grade	-0.027	-0.091	-0.097	-0.005	-0.038	0.007	-0.146	-0.002
	(0.434)	(0.401)	(0.371)	(0.008)	(0.321)	(0.025)	(0.361)	(0.018)
Math grade	0.002	-0.016	-0.020	0.003	-0.010	0.004	-0.012	0.006
	(0.234)	(0.236)	(0.221)	(0.013)	(0.177)	(0.015)	(0.154)	(0.017)
Danish grade missing	-0.004	-0.005	-0.012	-0.099	-0.008	0.021	-0.016	-0.030
	(0.013)	(0.015)	(0.014)	(0.179)	(0.019)	(0.235)	(0.020)	(0.254)
Math grade missing	0.006	0.002	0.009	-0.030	0.011	-0.057	0.014	-0.054
	(0.020)	(0.010)	(0.021)	(0.089)	(0.024)	(0.186)	(0.023)	(0.088)
A-level: Business	0.005	0.005	-0.001	0.118	0.000	0.139	-0.003	0.082
	(0.029)	(0.022)	(0.022)	(0.072)	(0.026)	(0.215)	(0.023)	(0.079)
A-level: Humanities	0.400	0.538	0.309	0.272	0.297	0.162	0.308	0.145
	(0.896)	(0.906)	(0.955)	(0.168)	(0.920)	(0.413)	(0.928)	(0.155)
A-level: STEM	-0.226	-0.279	-0.240	0.039	-0.171	0.071	-0.251	0.010
	(0.881)	(0.786)	(0.895)	(0.051)	(0.876)	(0.197)	(0.891)	(0.058)
Missing covars.	-0.162	-0.162	-0.148	0.168	-0.122	0.192	-0.147	0.085
	(0.095)	(0.105)	(0.065)	(0.197)	(0.077)	(0.280)	(0.073)	(0.229)
Year	5.306	5.140	5.279	0.001	5.290	0.011	5.310	-0.003
	(0.468)	(0.496)	(0.453)	(0.020)	(0.421)	(0.038)	(0.397)	(0.027)
Level	Program	Field	Program	Program	Program	Program	Program	Program
Correction	-	-	-	CF	-	CF	-	CF
Quota	-	-	1	1	1	1	2	2
Cutoff	-	-	Program	Program	Higher	Higher	Program	Program
$E\left[\frac{\partial y}{\partial GPA}\right]$	0.18	0.21	0.18	0.05	0.17	0.18	0.17	0.09

The table displays the estimated joint prior of parameters, μ_β in Equation (16) estimated from the program level estimates for the estimation methods outlined in Section 7.1.2. The dependent variable is the log of average income 7 to 9 years after admission. Standard errors, shown in parentheses below estimates, are the square root of the diagonal part of the estimate prior, Σ_β. Assuming uniform distribution of the ranked GPA, the average slope is calculated as $E\left[\frac{\partial y}{\partial GPA}\right] = \frac{1}{2}\beta_{GPA} + \frac{1}{3}\beta_{GPA^2}$.

parametric selection models.

Earnings functions Table 4 presents the mean and standard deviations (in parenthesis) of the posterior distribution of program-specific parameters.[25] Columns 1 and 2 show baseline estimates using conditional independence. Columns 4, 6 and 8 display control function estimates. To assess the importance of controlling for endogeneity, columns 3, 5 and 7 show corresponding estimates based on conditional independence but with the same subsample of program used in the control function approaches. The last row shows the average effect of GPA rank on earnings for a uniform distribution which corresponds to the distribution of in the population of high school graduates. The return to ability as measured by GPA is positive. However, controlling for selection at the margin of the cutoff lowers the average return from 0.18 to 0.05 (Column 3 vs. column 4), showing that selection is important for estimating returns to over-subscribed

[25]Note that the standard deviations are not standard errors and cannot be used to test significance of the means.

Table 5: Imputation methods for $E[y|admission]$

Level	Selection correction	Quota	Cutoff	Share	Cumulative share
Program	RDD	1	Program	0.55	0.55
Program	RDD	1	Higher priority	0.06	0.61
Program	RDD	2	Program	0.00	0.62
Program	CIA	-	-	0.15	0.77
Field	CIA	-	-	0.10	0.86
-	-	-	-	0.14	1.00

The table show the distribution of imputation methods in the simulation sample. The order of the prediction methods correspond to the priority the the methods.

programs. This is also reflected using Quota 2 applicants. I do not observe the same evidence of selection using cutoffs from more preferred programs.

Appendix Figure A.6 show that estimated program-specific LATEs using the control function approach follow the corresponding IV estimates using the variation from crossing own-program cutoffs. This is supportive evidence that the structure of the outcome equations is not overly restrictive.[26]

Using these sets of program specific parameters, I predict earnings in my simulation sample for all combinations of programs and applicants. Table 5 breaks down how many potential outcomes are computed by each method and the priority of the methods. With control function estimates, I am able to predict 62 percent of potential outcomes. 25 percent of outcomes are predicted with estimates based on conditional independence. The remaining 14 percent is predicted using the mean parameters of column 4 in Table 4.

7.3 Implications of ripple effects for program evaluation

To quantify the importance of ripple effects for program evaluation I exploit the following identity for the changes in predicted earnings relative to the baseline:

$$\underbrace{\Delta^T}_{\text{Total change in expected earnings}} = \underbrace{\Delta^F}_{\text{First-order}} + \underbrace{\Delta^H}_{\text{higher-order}}, \tag{18}$$

where I separate out the earnings changes occurring at the margin of the program with changed capacity. The rest of the changes in earnings is due to ripple effects. I calculate the elements of Equation (18) for each simulation in two ways; i) merging program specific LATEs with simulated applicant flows, and, ii) merging predicted potential outcomes to the allocation of individual applicants.

Using the identity in Equation (18), I first perform a variance decomposition of the simulations based on reported preferences and display the results in Table 6. I perform the decomposition using the raw LATEs (estimated by 2SLS) and for the shrunken estimates. For the raw estimates, the share explained by first-order moves is 85 percent with 15 percent left unexplained. In other words, the aggregate variance is 18 percent (15/85) larger than what

[26]Appendix Figure A.6 plots the model-based LATEs based on the raw estimates. In the remaining part of this paper I use shrunken estimates to reduce variance in the prediction at the cost of bias.

Table 6: Variance decomposition using reported preferences and LATEs

	2SLS	Empirical Bayes
Var(First order)	85.0	78.2
Var(Higher order)	8.0	9.1
2× Cov(First order, Higher order)	7.0	12.7
Var(Total)	100.0	100.0

The table displays the contribution to the total variance in gains from moves across experiments from first-order and higher-order moves (and covariance). First-order moves are defined as moves that occur along the margin of program where capacity is changes prior to the counterfactual clearing. The results are based on reported rank-ordered lists and 2SLS estimates of LATEs which do not account composition changes.

Table 7: Variance decomposition using estimated preferences and predicted outcomes

	Var	Pct.	Pct. cum.
Var(First order)	58,389,480.31	84.52	84.52
Var(Higher order)	5,163,294.33	7.47	92.00
2× Cov(First order, Higher order)	5,528,465.43	8.00	100.00
Var(Total)	69,081,240.08	100.00	

The table displays the contribution to the total variance in gains from moves across experiments from first-order and higher-order moves (and covariance). First-order moves are defined as moves that occur along the margin of program where capacity is changes prior to the counterfactual clearing. The results are based on filled out rank-ordered lists and control function estimates taking composition changes into account.

would be expected if one ignored ripple effects. Using shrunken estimates, the share explained by first-order moves falls to 78.2 percent. Appendix Figure A.5 shows that the share is increasing in absolute size of the percentage change in capacity. The share of variance explained by first-order moves is largest for big program expansions. In Table 7, I perform the same kind of variance decomposition but I exchange reported preferences and LATEs for estimated preferences and predicted potential outcomes. The results are remarkably similar, as around 85 percent of the variance is explained by the first-order moves, again leaving 15 percent of the variance unexplained.

The first-order and the higher-order effects are positively correlated with a Pearson coefficient of 0.15. The positive correlation implies that the sign of a first-order LATE is informative of the sign of the higher-order effect and failing to account for ripple-effects leads to underestimation of the aggregate effects. However, this correlation is far from 1, meaning that the results in Table 7 masks considerable heterogeneity across programs.

Program-level returns to adding slots To investigate program-level heterogeneity further, I compute the marginal return to increasing supply for each program based on estimated preferences and predicted potential outcomes. For each program, I have performed 28 experiments ranging from subtracting 95 percent to adding 95 percent of slots which, depending on the size of the program, correspond to different numbers of slots. To quantify the marginal effect of supply changes, I regress the first-order (F), higher-order (H), and the aggregate total (T) gains of the experiments on the number of slots changed:

$$\Delta_{pe}^{g} = \gamma_{p}^{g}(S_{pe} - S_{p0}) + \varepsilon_{pe}, \quad g \in \{F, H, T\}, \tag{19}$$

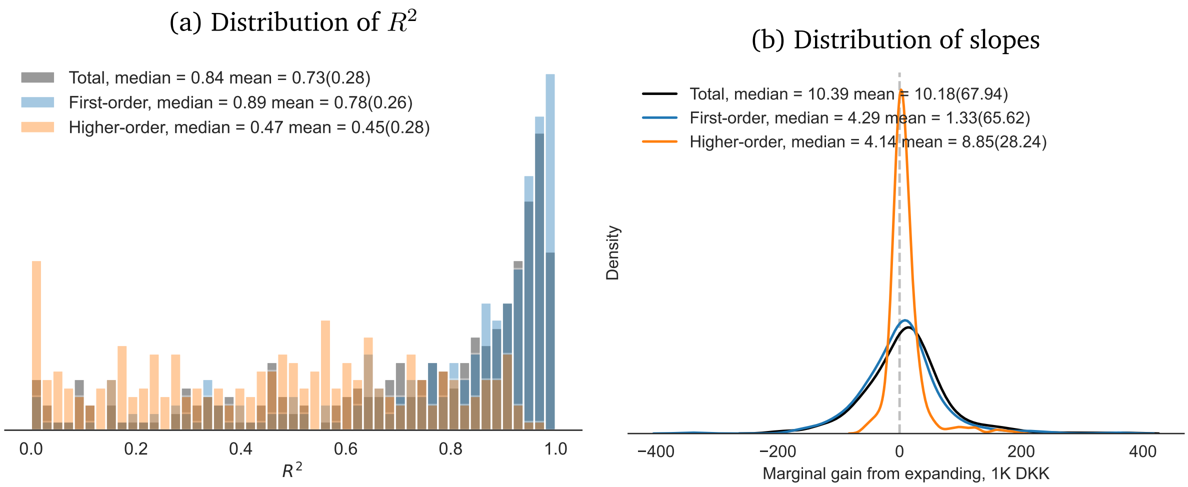


Figure 8: Marginal gain of changing supply

Figure 8a shows the distribution of R^2 from regressing aggregate gains on the number of slots as in Equation (19). Figure 8b shows the distribution of corresponding slopes, γ_p^g.

where Δ_{pe}^g is the gain from a change e to program p, S_{pe} is the number of slots in program p in simulation e and S_{p0} is the number of slots in the baseline scenario. Gains are measured in 1,000 DKK. The set of parameters, $\{\gamma_p^g\}$, can be interpreted as the marginal gain from adding a slot in each of the programs.

The usefulness of these measures depends on the underlying linearity between changes in slots and gains. Linearity could be violated for a number of reasons: Firstly, the applicant pool might not be large enough to fill up a program after an expansion, meaning that at some point the marginal gain from expansion drops to zero. Secondly, if the composition of the complier groups depends on the magnitude of supply change, the marginal effect of supply changes might change as a function of changing composition. To gauge the validity of the slope as a metric of marginal gains, Figure 8a plots the R^2 from the regressions in Equation (19). Total and first-order gains exhibit a very large degree of linearity with a median value of R^2 of 0.84 and 0.89 respectively. This implies that the first-order complier population does not change dramatically with the magnitude of the supply change and lends credence to the external validity of LATEs for first-order gains. The higher-order gains, however, exhibit less, though still substantial, linearity with a median R^2 of 0.47. The high degree of linearity implies that the sets of γ_p^g provides valuable information on the effects of changes in supply on earnings.

The distribution of the slopes, γ_p^g, is shown in Figure 8b. The median slope for total gains is 10.4, meaning that the median marginal gain of adding 1 slot to a program is an anual gain of 10.400 DKK ($\approx$ 1500 USD). However, there are large differences between programs, as some programs yield as much as 200 thousand DKK per added slot ($\approx$ 30K USD). The magnitude of first-order gains is generally larger than higher-order gains as evidenced by the standard deviations (65.62 and 28.24 respectively).[27]

[27] The program-specific marginal gains are somewhat sensitive to parametrization of the potential outcomes. This can be seen by inspecting an earlier version of this paper (Gandil, 2021). However, the results presented here are the ones where I obtain the maximum share of variation explained by the first-order moves.

Table 8: Marginal gains of program expansion

Length	Field		First-order	Higher-order	Total	N
Long-cycle	Health	Mean	-17.5	-12.7	-30.2	10
		Std	(116.5)	(32.8)	(108.6)	
	Humanities and communication	Mean	7.0	9.9	16.8	55
		Std	(42.4)	(13.9)	(41.0)	
	STEM	Mean	13.7	9.2	22.9	55
		Std	(48.6)	(15.7)	(52.9)	
	Social science and business	Mean	-2.0	31.9	29.9	53
		Std	(103.7)	(50.5)	(102.9)	
Medium-cycle	Health	Mean	-5.0	5.2	0.2	23
		Std	(60.9)	(19.3)	(63.1)	
	STEM	Mean	0.4	1.1	1.5	12
		Std	(69.4)	(6.7)	(71.8)	
	Social science and business	Mean	-8.0	2.2	-5.8	19
		Std	(61.8)	(19.3)	(69.9)	
	Teaching	Mean	-5.4	-0.3	-5.7	39
		Std	(32.4)	(7.7)	(34.2)	
Short-cycle	STEM	Mean	-18.0	-1.6	-19.6	6
		Std	(48.9)	(9.5)	(56.6)	
	Social science and business	Mean	5.0	-6.6	-1.6	23
		Std	(61.7)	(16.7)	(64.4)	

The table provides means (and standard deviations in parentheses) of the estimated slopes within fields using the model in equation (19).

Table 8 breaks down the distribution of marginal gains by program length and field. On average, long-cycle health programs yield large first-order losses, but these only account for around 60 percent of the aggregate losses (-17.5 versus -30.2). STEM programs, on the other hand, have on average both first-order and higher-order gains. Long-cycle social science in an example where the averages of first-order and higher-order effects counteract. A small marginal loss is offset by a relatively large higher-order gain. Notice, however, that the standard deviations are large implying heterogeneity *within* fields. This is important as applicants substitute between similar programs.

Though Table 8 shows heterogeneity in returns to program expansion, it is uninformative on the interdependence between first-order and second-order gains within fields. To investigate how higher-order effects might change policy recommendations, Figure 9 plots the higher-order gains against first-order gains for each program. The dashed black line represents break-even points where a marginal first-order gain is completely offset by a higher-order loss (or vice-versa). The programs are colored according to total gain and all programs to the right of the dashed line have positive total marginal returns.

The figure provides a number of takeaways. Firstly, the externality of investing in the added supply of slots is in general positive (as seen by the mass above zero on the y-axis). Secondly, a number of programs, such as Business at Aarhus University (AU) and Economics at University of Copenhagen (KU), yield large marginal first-order returns to program expansion while having small second-order effects. These programs typically have low entry barriers and therefore do not create long chains of applicants. Thirdly, programs with similar first-order marginal gains can exhibit very different second-order gains, exemplified with the Business program at Copenhagen Business School (CBS) which despite its similarity in content with the Aarhus program yields larger higher-order gains. Finally, even though a program yields negative first-order

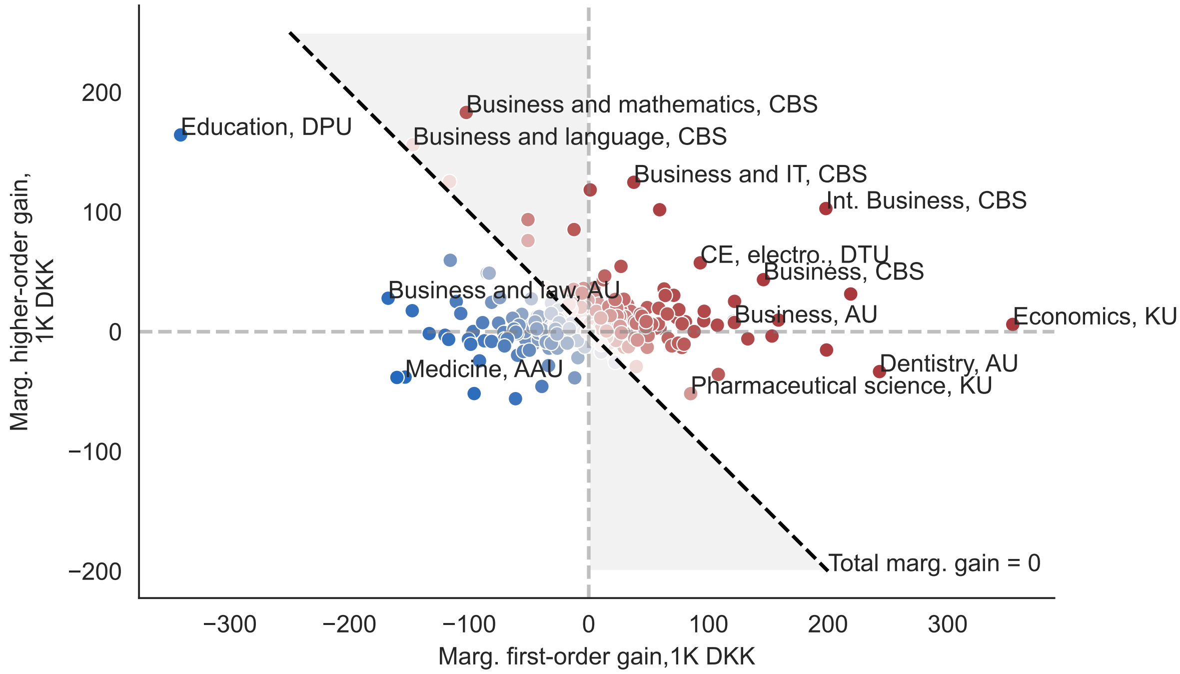


Figure 9: First-order and higher-order marginal gains

The figure plots estimated higher-order marginal gains against first-order marginal gains. The programs are colored according to their aggregate gains, which can be read off the distance to the dashed line. Select programs are annotated, CE and BE are abbreviations for civil and bachelor engineering respectively.

marginal returns, this may be redeemed by higher-order gains. These programs are displayed in the upper shaded triangle. An example of these is Business and Mathematics at Copenhagen Business School.[28] However, the majority of programs are outside this triangle, which implies that negative first-order gains are not, in general, fully mitigated by higher-order gains.

To quantify the relationship in Figure 9, I regress higher-order gains (the y-axis) and total gains (the colors) on the first-order gains (the x-axis) and present results in Table 9. Panel a) presents results using the marginal higher-order gain as dependent variable. Column 1 shows that the first-order gain is not predictive of higher-order gains. However, this masks a non-linearity. In column 2, I interact the first-order gain with its own sign. For negative gains I find a *negative* coefficient 0f -0.18, which is marginally statistically significant. This implies that negative first-order gains are on average mitigated by higher-order gains. In columns 3 and 4 I regress the marginal higher-order gains on centrality and cutoffs. The coefficients are positive and significant, which indicate that more central and oversubscribed programs generally yield higher higher-order gains. This holds up when I include first-order gains as seen in columns 5, 6, 8 and 9. When I include both measures of program popularity jointly, the parameter on centrality attenuates and becomes highly insignificant.

In panel b) of Table 9, I exchange the dependent variable for the total marginal gain which includes both first-order and higher-order gains. Unsurprisingly, the coefficient on the first-order

[28]Note that as a policy, it is not necessary to incur the first-order loss. Instead one can choose to expand capacity in the programs where marginal applicants come from.

Table 9: Predicting ripple effects in gains

	Panel a) Dependent variable: Higher-order marginal gain, γ_p^H									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FO gain	-0.06				-0.06	-0.07	-0.07			
	(0.05)				(0.05)	(0.05)	(0.05)			
FO gain$\leq 0 \times$ FO gain		-0.18						-0.16	-0.15	-0.15
		(0.10)						(0.10)	(0.10)	(0.10)
FO gain$>0 \times$ FO gain		0.06						0.04	0.00	0.00
		(0.05)						(0.05)	(0.05)	(0.05)
Eigenvector centrality			4.48		4.57		-0.83	3.43		-1.38
			(1.22)		(1.25)		(1.66)	(1.14)		(1.60)
GPA cutoff				9.72		10.13	10.56		9.30	9.97
				(1.77)		(1.87)	(2.33)		(1.75)	(2.23)
Constant	8.85	3.56	8.85	8.85	8.85	8.85	8.85	4.30	5.59	5.44
	(1.63)	(2.49)	(1.63)	(1.55)	(1.61)	(1.53)	(1.53)	(2.55)	(2.62)	(2.63)
Observations	295	295	295	295	295	295	295	295	295	295
R2	0.02	0.06	0.03	0.12	0.04	0.15	0.15	0.07	0.16	0.16

	Panel b) Dependent variable: Total marginal gain, γ_p^T									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FO gain	0.94				0.94	0.93	0.93			
	(0.05)				(0.05)	(0.05)	(0.05)			
FO gain$\leq 0 \times$ FO gain		0.82						0.84	0.85	0.85
		(0.10)						(0.10)	(0.10)	(0.10)
FO gain$>0 \times$ FO gain		1.06						1.04	1.00	1.00
		(0.05)						(0.05)	(0.05)	(0.05)
Eigenvector centrality			5.88		4.57		-0.83	3.43		-1.38
			(3.98)		(1.25)		(1.66)	(1.14)		(1.60)
GPA cutoff				15.54		10.13	10.56		9.30	9.97
				(4.27)		(1.87)	(2.33)		(1.75)	(2.23)
Constant	10.18	4.89	10.18	10.18	10.18	10.18	10.18	5.62	6.92	6.77
	(1.63)	(2.49)	(3.95)	(3.86)	(1.61)	(1.53)	(1.53)	(2.55)	(2.62)	(2.63)
Observations	295	295	295	295	295	295	295	295	295	295
R2	0.83	0.84	0.01	0.05	0.83	0.85	0.85	0.84	0.85	0.86

This table show estimated parameters for OLS regressions. The dependent variable in panel a) is the marginal higher-order gain, γ_p^H as defined in Equation 19. In panel b) the dependent variable is the marginal total gain, γ_p^T. Both eigenvector centrality and cutoffs are standardized to have a mean of zero and standard deviation of 1 across programs. Robust standard errors are in parentheses.

gain is close to one and highly significant. This mirrors the variance decomposition in Table 7 where first-order responses explain the majority of the variance. Centrality of the program is now much less predictive relative to panel a) as seen in column 3 of panel b). This reflects that centrality drives higher-order responses but not first-order responses. Correspondingly, when including first-order gains in column 5, the parameter is equal to column 5 in panel a).[29]

The results in Table 9 suggest that popularity of programs, either measured by centrality or GPA cutoffs are good predictors of higher-order effects. These results are confirmed when I exchange the dependent variables in Table 9 with indicators for positive marginal gains as seen in Appendix Table A2. To the extend that these results carry over to other institutional contexts, this implies that policymakers may reasonably conjecture that the returns to expanding popular

[29]This holds for panel b columns 5-10 where it is only the parameters on first-order gains which can change.

programs are underestimated.

8 Conclusion

Ripple effects in the educational sector is a critical factor that can either amplify or counterbalance the effects of local supply changes. Neglecting the externalities imposed on other programs through ripple effects can lead to severely misguided policy recommendations.

The findings of this paper reveal two important aspects of ripple effects for policy evaluation. Firstly, even applicants who do not apply to a particular program are influenced by ripple effects, underscoring the importance of considering the welfare of these individuals. Secondly, ripple effects introduce a divergence between the perceived benefits from the perspective of a program's admission office and the overall societal implications. The results indicate that ripple effects generally make program expansions more attractive, potentially justifying further program subsidies. However, it is crucial to recognize the high degree of heterogeneity among programs, requiring careful consideration when selecting programs for subsidization.

I approach the implications of ripple effects for policy evaluation in a number of ways. Regardless of whether reported or estimated preferences are employed, or whether non-parametric LATEs or structured potential outcomes are used, I consistently arrives at similar qualitative conclusions. A sizable fraction of the relevant variation in outcomes is lost when ignoring ripple effects and policy recommendations may change once ripple effects are taken into account.

Chains of applicants are significantly shorter when using reported preferences, primarily due to the shorter application lists. However, it is important to note that ripple effects are intrinsically tied to applicants' decision-making process, and they still occur even if not explicitly observed in admission data. Therefore, it is likely that the magnitude of ripple effects aligns more closely with simulations that incorporate complete rank-ordered lists of programs based on preference estimation.

This paper underscores the significance of considering the equilibrium effects of changing the supply of education. The computed equilibria in this paper are subject to important constraints. Future research could explore the time dynamics as potential applicants may adjust their application timing in response to supply changes. Additionally, incorporating potential changes in labor market returns to skills resulting from shifts in the composition of the workforce is another crucial dimension to explore. Nevertheless, the study provides evidence supporting the incorporation of ripple effects into policymaking.

Although the results are specific to the Danish context of centralized assignment, the qualitative finding that ripple effects matter extends beyond this particular context or the centralized nature of the market. Educational slots are inherently indivisible, and applicants typically have unit demand, making ripple effects significant whenever supply is limited. These findings can however also be generalized to other markets. Interpreting cutoffs as prices, the results highlight the importance of pecuniary externalities even without distributional concerns. In other words, when formulating policies, it is imperative to complement well-defined and policy-relevant treatment effects with an understanding of how the relevant market clears.

References

ABDULKADIROĞLU, A., N. AGARWAL, AND P. A. PATHAK (2017): "The welfare effects of coordinated assignment: Evidence from the New York City high school match," *American Economic Review*, 107(12), 3635–89.

ABDULKADIROĞLU, A., J. D. ANGRIST, Y. NARITA, AND P. A. PATHAK (2017): "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*, 85(5), 1373–1432.

ABDULKADIROGLU, A., J. D. ANGRIST, Y. NARITA, AND P. A. PATHAK (2022): "Breaking ties: Regression discontinuity design meets market design," *Econometrica*.

ABDULKADIROĞLU, A., AND T. SÖNMEZ (2003): "School choice: A mechanism design approach," *American economic review*, 93(3), 729–747.

ABDULKADIROĞLU, A., P. A. PATHAK, J. SCHELLENBERG, AND C. R. WALTERS (2020): "Do Parents Value School Effectiveness?," *American Economic Review*, 110(5), 1502–39.

AGARWAL, N. (2015): "An empirical model of the medical match," *American Economic Review*, 105(7), 1939–78.

——— (2017): "Policy analysis in matching markets," *American Economic Review*, 107(5), 246–50.

AGARWAL, N., AND P. SOMAINI (2018): "Demand analysis using strategic reports: An application to a school choice mechanism," *Econometrica*, 86(2), 391–444.

ALTONJI, J. G., P. ARCIDIACONO, AND A. MAUREL (2016): "The analysis of field choice in college and graduate school: Determinants and wage effects," in *Handbook of the Economics of Education*, vol. 5, pp. 305–396. Elsevier.

ANDERSEN, M. B., S. HØRLÜCK, AND C. L. SØRENSEN (2020): "Effekt af optag på videregående uddannelser for marginale studerende," Discussion paper, Working Paper.

ANDREWS, R. J., J. LI, AND M. F. LOVENHEIM (2016): "Quantile treatment effects of college quality on earnings," *Journal of Human Resources*, 51(1), 200–238.

ARCIDIACONO, P., E. M. AUCEJO, AND V. J. HOTZ (2016): "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California," *American Economic Review*, 106(3), 525–62.

AZEVEDO, E. M., AND J. D. LESHNO (2016): "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 124(5), 1235–1268.

BUCAREY, A. (2018): "Who pays for free college? Crowding out on campus," *Job market paper*.

CALSAMIGLIA, C., G. HAERINGER, AND F. KLIJN (2010): "Constrained school choice: An experimental study," *American Economic Review*, 100(4), 1860–74.

CHRISANDER, E., AND A. BJERRE-NIELSEN (2023): "Why Do Students Lie and Should We Worry? An Analysis of Non-truthful Reporting," *arXiv preprint arXiv:2302.13718*.

DALE, S. B., AND A. B. KRUEGER (2002): "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, 117(4), 1491–1527.

DALY, M., M. F. JENSEN, AND D. LE MAIRE (2020): "University admission and preferred field of study," Discussion paper, Working Paper.

FACK, G., J. GRENET, AND Y. HE (2019): "Beyond truth-telling: Preference estimation with centralized school choice and college admissions," *American Economic Review*, 109(4), 1486–1529.

FELLER, A., T. GRINDAL, L. MIRATRIX, L. C. PAGE, ET AL. (2016): "Compared to what? Variation in the impacts of early childhood education by alternative care type," *Annals of Applied Statistics*, 10(3), 1245–1285.

GALE, D., AND L. S. SHAPLEY (1962): "College admissions and the stability of marriage," *The American Mathematical Monthly*, 69(1), 9–15.

GANDIL, M., AND E. LEUVEN (2022): "College Admission as a Screening and Sorting Device," .

GANDIL, M. H. (2021): "Substitution effects in college admissions," Discussion paper, Memorandum.

HASTINGS, J. S., C. A. NEILSON, AND S. D. ZIMMERMAN (2013): "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," Discussion paper, National Bureau of Economic Research.

HECKMAN, J. J. (1979): "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, pp. 153–161.

HECKMAN, J. J., S. URZUA, AND E. VYTLACIL (2008): "Instrumental variables in models with multiple outcomes: The general unordered case," *Annales d'Economie et de Statistique*, pp. 151–174.

HECKMAN, J. J., AND E. VYTLACIL (2001): "Policy-relevant treatment effects," *American Economic Review*, 91(2), 107–111.

HEINESEN, E., AND C. HVID (2019): "Returns To Field Of Study In The Medium Term – Instrumental Variables Estimates Based On Admission Thresholds," Discussion paper, Working Paper.

HEINESEN, E., C. HVID, L. J. KIRKEBØEN, E. LEUVEN, AND M. MOGSTAD (2022): "Instrumental variables with unordered treatments: Theory and evidence from returns to fields of study," Discussion paper, National Bureau of Economic Research.

HUMLUM, A., AND B. B. MEYER (2020): "Artificial Intelligence and College Majors," Discussion paper, Working Paper.

IMBENS, G. W., AND J. D. ANGRIST (1994): “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62(2), 467–475.

KAPOR, A. J., C. A. NEILSON, AND S. D. ZIMMERMAN (2020): “Heterogeneous beliefs and school choice mechanisms,” *American Economic Review*, 110(5), 1274–1315.

KIRKEBOEN, L. J., E. LEUVEN, AND M. MOGSTAD (2016): “Field of study, earnings, and self-selection,” *The Quarterly Journal of Economics*, 131(3), 1057–1111.

KLINE, P., AND C. R. WALTERS (2016): “Evaluating public programs with close substitutes: The case of Head Start,” *The Quarterly Journal of Economics*, 131(4), 1795–1848.

——— (2019): “On heckits, late, and numerical equivalence,” *Econometrica*, 87(2), 677–696.

MANNING, A., AND B. PETRONGOLO (2017): “How local are labor markets? Evidence from a spatial job search model,” *American Economic Review*, 107(10), 2877–2907.

MORRIS, C. N. (1983): “Parametric empirical Bayes inference: theory and applications,” *Journal of the American statistical Association*, 78(381), 47–55.

MOUNTJOY, J., AND B. HICKMAN (2020): “The Returns to College (s): Estimating Value-Added and Match Effects in Higher Education,” *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 1(2020-08).

PATHAK, P. A., AND P. SHI (2021): “How well do structural demand models work? Counterfactual predictions in school choice,” *Journal of Econometrics*, 222(1), 161–195.

ROBBINS, H. E. (1992): “An empirical Bayes approach to statistics,” in *Breakthroughs in statistics*, pp. 388–394. Springer.

ROY, A. D. (1951): “Some thoughts on the distribution of earnings,” *Oxford economic papers*, 3(2), 135–146.

SANDMO, A. (2015): “Chapter 1 - The Principal Problem in Political Economy: Income Distribution in the History of Economic Thought,” in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 2 of *Handbook of Income Distribution*, pp. 3–65. Elsevier.

TANAKA, M., Y. NARITA, AND C. MORIGUCHI (2020): “Meritocracy and Its Discontent: Long-run Effects of Repeated School Admission Reforms,” Discussion paper, Research Institute of Economy, Trade and Industry (RIETI).

ZIMMERMAN, S. D. (2014): “The returns to college admission for academically marginal students,” *Journal of Labor Economics*, 32(4), 711–754.

A Appendix - Figures and tables

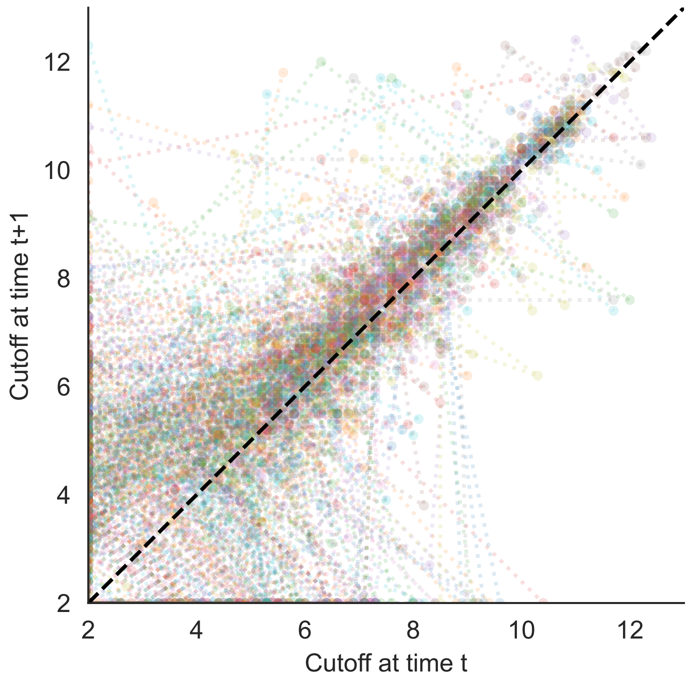


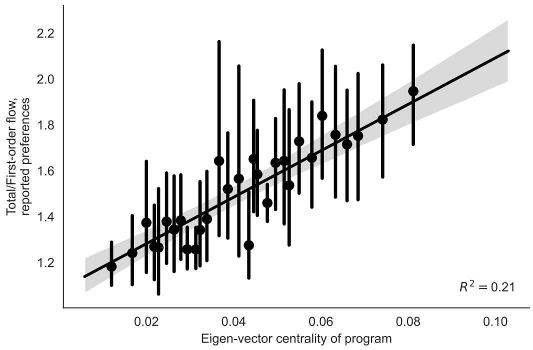
Figure A.1: Cutoff change from year to year

The figure plots Quota 1 cutoffs as a function of cutoff from last year. Mass around the 45 degree line indicate stability of cutoffs from year to year.

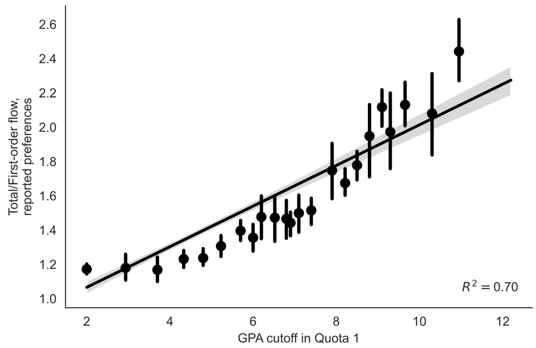
Table A1: Variance decomposition with reported preferences and estimated outcomes

	Var	Pct.	Pct. cum.
Var(First order)	58,524,445.55	78.41	78.41
Var(Higher order)	5,846,831.40	7.83	86.25
2× Cov(First order, Higher order)	10,264,999.28	13.75	100.00
Var(Total)	74,636,276.23	100.00	

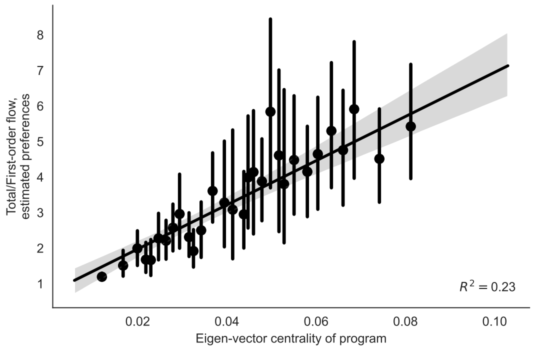
The table displays the contribution to the total variance in gains from moves across experiments from first-order and higher-order moves (and covariance). First-order moves are defined as moves that occur along the margin of program where capacity is changes prior to the counterfactual clearing. The results are based on reported rank-ordered lists and control function estimates taking composition changes into account.



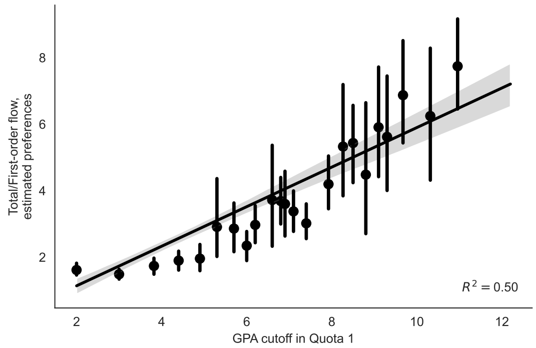
(a) Centrality with rep. pref.



(b) Cutoffs with rep. pref.



(c) Centrality with est. pref.



(d) Cutoffs with est. pref.

Figure A.2: Predicting ripple effects in flows

The figure plots the linear relationship between ripple effects in terms of flows of applicants and measures of popularity of programs. Panel a) and b) shows the results based on simulation with reported preferences. Panel c) and d) shows the results based on simulation with estiamted preferences. Centrality is measured as the Eigen-vector centrality of a weighted graph constructed from the applications for the round of 2016. An edge weigth is computed as the number of times the two programs appear on the same rank-ordered list. The GPA cutoffs are the cutoffs in Quota 1 and are publicly available on the ministry website. Confidence-intervals are on 95%-level.

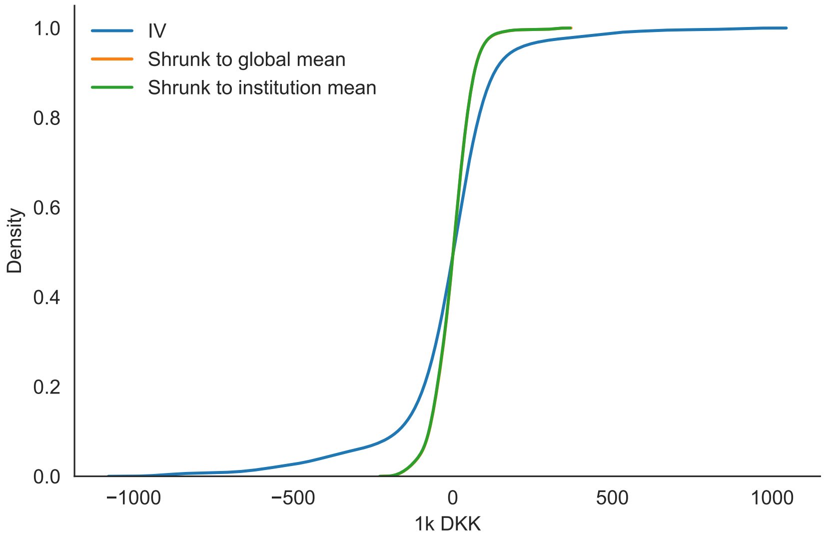
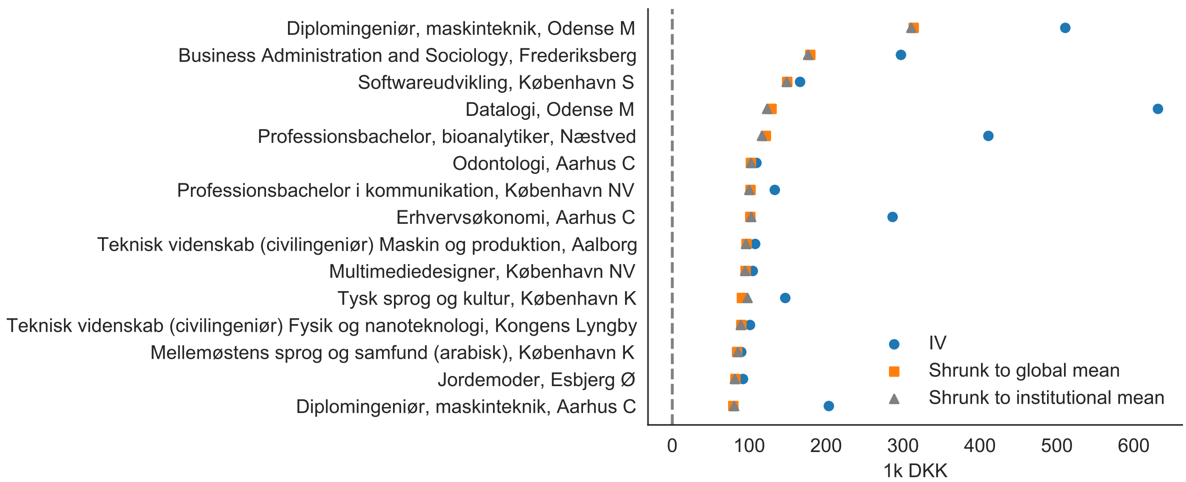
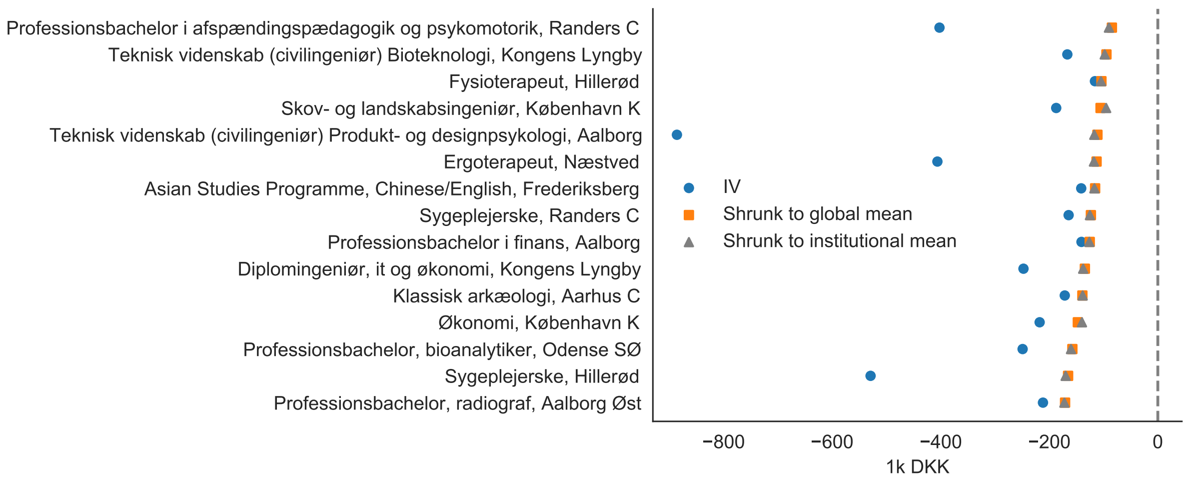


Figure A.3: Distribution of program-specific LATEs

The figure shows the smoothed cumulative distribution function of the program specific LATEs. The raw estimates are shrunken towards the institutional mean. It makes little difference whether I shrink towards the global or the institutional mean. I therefore leave out shrinkage towards the institutional mean.



(a) Top programs



(b) Bottom programs

Figure A.4: Programs in the top and bottom of the distribution of LATEs

The figures shows the programs with highest and lowest Empirical Bayes LATEs. The estimation approach is outlined in Section 7.1.1

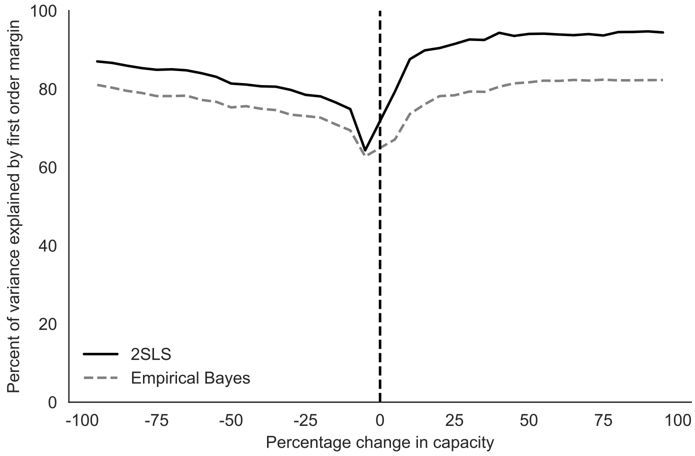
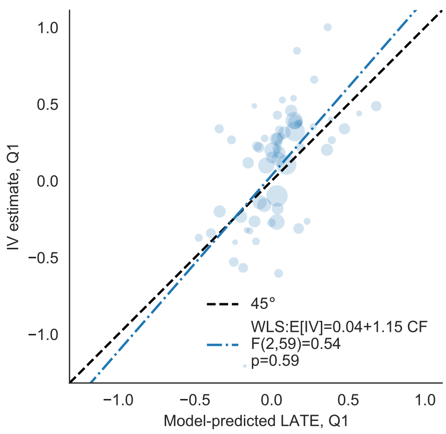


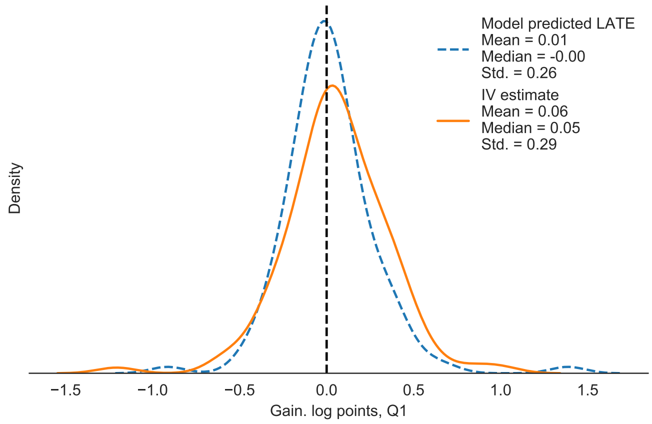
Figure A.5: Variance explained by first order margin as a function of size of program expansion

The figure shows the share variance in the earnings change in simulations explained by movement of applicants on the margin of the changed program.

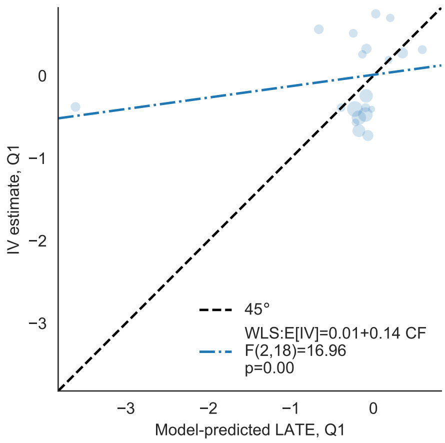
(a) Q1, program, correlation



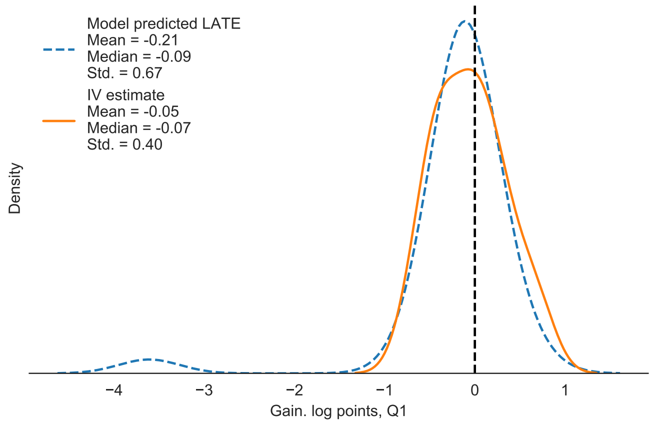
(b) Q1, program, distribution



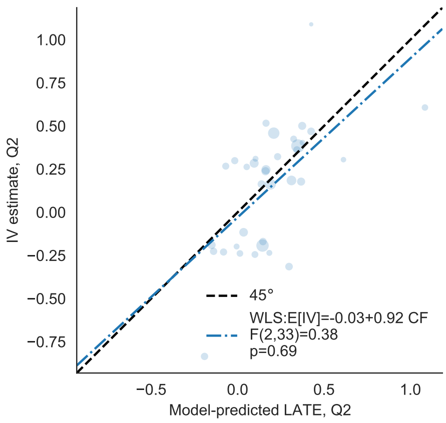
(c) Q1, higher, correlation



(d) Q1, higher, distribution



(e) Q2, program, correlation



(f) Q2, program, distribution

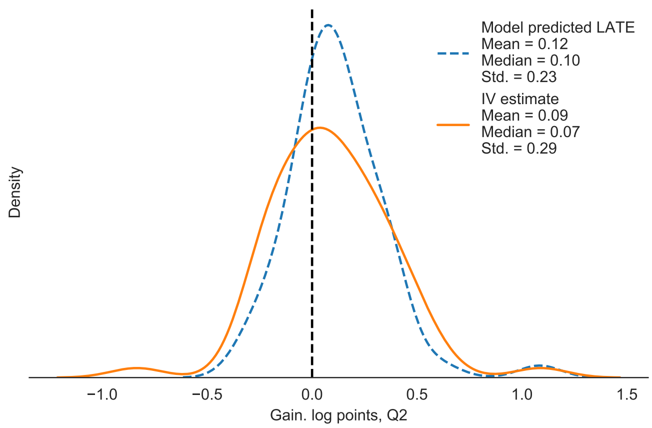
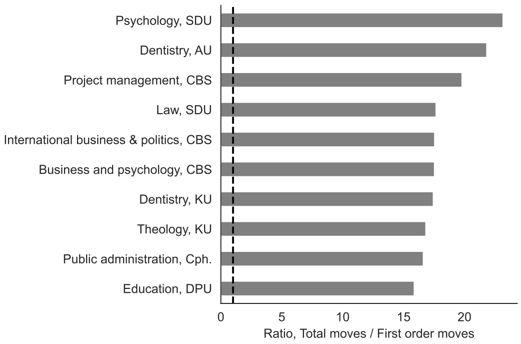
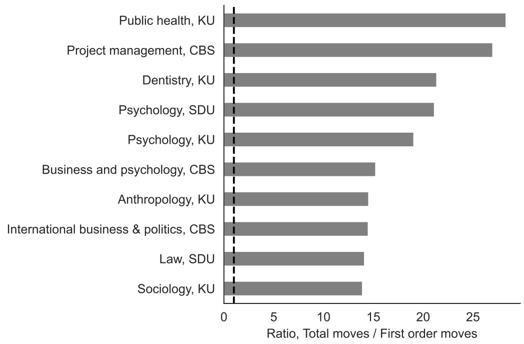


Figure A.6: Model-based and IV estimates of LATE

The figure shows LATE estimates calculated from the selection correction model in Equation (13) and IV estimates of a standard fuzzy regression discontinuity design implemented with two stage least squares. The source of variation is crossing own program cutoff in quota 1. Figures on the left shows the model-based estimates on the x-axis and the corresponding IV estimate on the y-axis. The data points are weighted by sample size. The legend shows the result of weighted least squares and the corresponding F-test for whether the points are on the 45-degree line up to sampling noise. 2SLS estimates with an F-statistic lower than 10 or a p-value larger than 0.7 have been excluded. Figures on the right shows the marginal distribution of LATEs using IV and control functions.



(a) 10 percent reduction in capacity



(b) 10 percent increase in capacity

Figure A.7: Programs with the largest ripple effects measured by flows of applicants, estimated preferences

The figure plots the ratio of the total number of movers relative to the first order movers for a ten percent change in capacity. The programs shown are the ones with the largest ripple effects and which are sufficiently large to fulfill anonymity requirements set by Statistics Denmark.

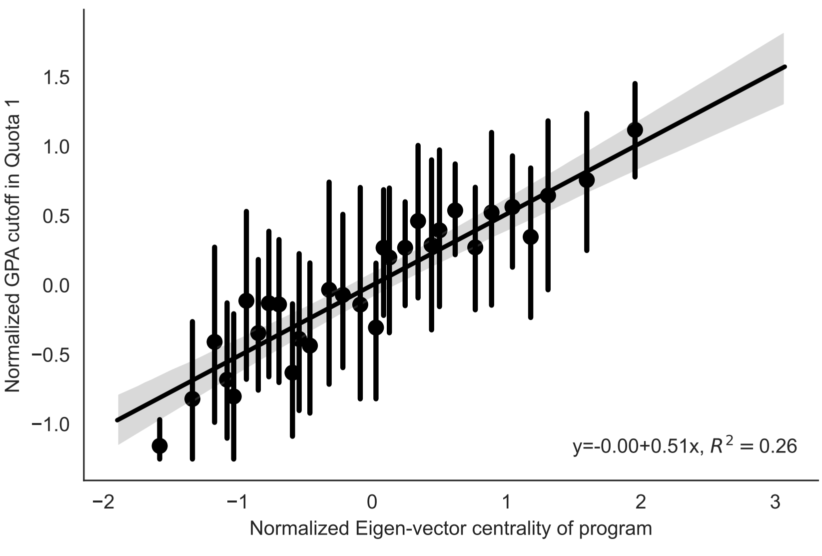


Figure A.8: Predicting ripple effects in ratio of absolute gains

The figure plots the linear relationship between GPA cutoffs and weighted eigenvector centrality in the applications for the 2016 cohort.

Table A2: Predicting ripple effects in gains

	Panel a) Dependent variable: Positive higher-order marginal gain, $\gamma_p^H > 0$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FO gain	0.001				0.001	0.000	0.000			
	(0.000)				(0.000)	(0.000)	(0.000)			
FO gain≤ 0 $\times$ FO gain		0.001						0.001	0.001	0.001
		(0.001)						(0.001)	(0.001)	(0.001)
FO gain>0 $\times$ FO gain		0.001						-0.000	-0.000	-0.001
		(0.001)						(0.001)	(0.001)	(0.001)
Eigenvector centrality			0.126		0.125		0.063	0.133		0.070
			(0.026)		(0.026)		(0.030)	(0.026)		(0.031)
GPA cutoff				0.156		0.154	0.121		0.163	0.129
				(0.026)		(0.026)	(0.031)		(0.026)	(0.031)
Constant	0.620	0.621	0.620	0.620	0.620	0.620	0.620	0.650	0.657	0.664
	(0.028)	(0.039)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.037)	(0.036)	(0.036)
Observations	295	295	295	295	295	295	295	295	295	295
R2	0.006	0.006	0.067	0.103	0.072	0.105	0.118	0.076	0.111	0.126

	Panel b) Dependent variable: Positive total marginal gain, $\gamma_p^T > 0$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FO gain	0.005				0.005	0.005	0.005			
	(0.001)				(0.001)	(0.001)	(0.001)			
FO gain≤ 0 $\times$ FO gain		0.006						0.006	0.006	0.006
		(0.001)						(0.001)	(0.001)	(0.001)
FO gain>0 $\times$ FO gain		0.003						0.003	0.003	0.003
		(0.001)						(0.001)	(0.001)	(0.001)
Eigenvector centrality			0.037		0.031		0.005	0.050		0.018
			(0.029)		(0.023)		(0.026)	(0.022)		(0.024)
GPA cutoff				0.081		0.054	0.052		0.075	0.066
				(0.029)		(0.024)	(0.027)		(0.023)	(0.026)
Constant	0.600	0.664	0.600	0.600	0.600	0.600	0.600	0.675	0.680	0.682
	(0.022)	(0.042)	(0.029)	(0.028)	(0.022)	(0.022)	(0.022)	(0.041)	(0.042)	(0.042)
Observations	295	295	295	295	295	295	295	295	295	295
R2	0.385	0.404	0.006	0.027	0.389	0.397	0.397	0.414	0.426	0.427

This table show estimated parameters for OLS regressions. The dependent variable in panel a) is an indicator for a positive marginal higher order gain, $\gamma_p^H > 0$, as defined in Equation 19. In panel b) the dependent variable is an indicator for positive marginal total gain, γ_p^T. Both eigenvector centrality and cutoffs are standardized to have a mean of zero and standard deviation of 1 across programs. Robust standard errors are in parentheses.

B Preference estimation

As described in Section 5 I employ the stability-based estimator introduced by Fack, Grenet, and He (2019). I estimate the models within stata and ignore strata subscripts in the followning outline.

I maintain notation as in Section 5 and assume that applicant i receives the following utility of admission to program p:

$$U_{ip} = V_{ip}(W_{ip}, \gamma) + u_{ip}, \tag{20}$$

where V_ip is a known function of applicant and school characteristics and a strata-specific vector of coefficients. I assume that u_{ip} follows a standard type-1 extreme value distribution.

Let $e_i = \{e_p\}$ be the vector of eligibility scores. Let μ be a realized matching applicants and programs and let $C(\mu)$ be the resulting vector of cutoffs. The feasible set of programs for applicant i is then $\mathcal{S}(e_i, C(\mu))$, where the outside option of no-admittance is always included.

The log-likelihood function is then given by

$$\ln L\left(\gamma | W_{ip}, \mathcal{S}(e_i, C(\mu))\right) = \sum_i^N \sum_p^P V_{ip} \times \mathbf{1}(i \in p) - \sum_i^N \ln\left(\sum_{p' \in \mathcal{S}(e_i, C(\mu))} \exp(V_{ip'})\right),$$

which is a standard conditional logit model with personalized choice set.

Specification of utility The non-stochastic part of utility, V_ip is modeled as a linear function of program level covariates interacted with applicant characteristics. Each program, p maps into a geograpci region, $r(p)$, program length, $l(p)$, and field, $f(p)$. Again suppressing strata subscripts on parameters, I specify the function of V_{ip}:

$$\begin{aligned} V_{ip} &= \mu_{r(p)} + \eta_{l(p)} + \lambda_{f(p)} + \sum_{l'} \mathbf{1}(l(p) = l') \times \gamma_{gpa,l} GPA_i \\ &+ \mathbf{1}(CV_i = 1) \sum_{f'} \mathbf{1}(f(p) = f') \times (\gamma_{gpa,f} GPA_i + \alpha_{1,l} STEM_i + \alpha_{2,l} SOC_i + \alpha_{3,l} HUM_i + \alpha_{4,l} BUS_i) \\ &+ \mathbf{1}(CV_i = 0) \sum_{f'} \delta_f \mathbf{1}(f(p) = f') \end{aligned}$$

, where $\mu_{r(p)}, \eta_{l(p)}, \lambda_{f(p)}$ are region, length and field fixed effects. GPA_i is the high school GPA. CV_I is a dummy for whether covariates are available in the registers. The dummy variables $STEM_i, SOC_i, HUM_i, BUS_i$ indicate the A-levels in high school of the applicants. The outside option is treated as it's own program and heterogeneity in the value of the outside option is thus captured by the same interactions as for the other programs.

C Calculation of the Local Average Treatment Effect

This appendix briefly describe how the Local Average Treatment Effect (LATE) is calculated in the control function approach and how it is calculated in a standard fuzzy discontinuity design.

Control function estimation of LATE The control function estimates are estimated on a program level using a standard two-step switching regression framework. To simplify notation I suppress the program and quota indices used in the main text. Let D_i be a binary indicator for receiving offer. Again, e_i is the eligibility score centered around the cutoff and $Z_i = \mathbf{1}(e_i > 0)$. For each program with associated sample of applicants I estimate by probit the selection equation with a linear spline in the running variable:

$$D_i = \mathbf{1}\left\{\gamma_1 Z_i + \gamma_2 e_i + \gamma_3 Z_i \times e_i + \epsilon_i > 0\right\}.$$

Let Φ be the standard normal distribution and ϕ the associated density. Using the estimated parameters in the selection equation, I construct the inverse Mill's ratio and estimate the following regressions for treated and untreated respectively by OLS:

$$\begin{aligned} Y_i &= X_i\beta_1 + \psi_1 \frac{\phi(\hat{\gamma}_1 Z_i + \hat{\gamma}_2 e_i + \hat{\gamma}_3 Z_i \times e_i)}{\Phi(\hat{\gamma}_1 Z_i + \hat{\gamma}_2 e_i + \hat{\gamma}_3 Z_i \times e_i)} + u_i, \text{ if } D_i = 1 \\ Y_i &= X_i\beta_0 + \psi_0 \frac{-\phi(\hat{\gamma}_1 Z_i + \hat{\gamma}_2 e_i + \hat{\gamma}_3 Z_i \times e_i)}{1 - \Phi(\hat{\gamma}_1 Z_i + \hat{\gamma}_2 e_i + \hat{\gamma}_3 Z_i \times e_i)} + v_i, \text{ if } D_i = 0 \end{aligned}$$

Again, using the estimated values of the parameters from the selection model I calculate the following individual correction term:

$$\Gamma_i(e_i) = \frac{\phi\left(\hat{\gamma}_1 + (\hat{\gamma}_2 + \hat{\gamma}_3)e_i\right) - \phi\left(\hat{\gamma}_2 e_i\right)}{\Phi\left(\hat{\gamma}_1 + (\hat{\gamma}_2 + \hat{\gamma}_3)e_i\right) - \Phi\left(\hat{\gamma}_2 e_i\right)}.$$

I then proceed to construct expected outcomes for $d \in (0, 1)$ using the estimates of β_d. The LATE in the control function approach, τ_{CF}, is calculated as the sample analog to following expectation over the full sample:

$$\tau_{CF} = E[X_i(\hat{\beta}_1 - \hat{\beta}_0) + (\hat{\psi}_1 - \hat{\psi}_0)\Gamma_i(e_i)],$$

Instrumental variable estimation of LATE The IV estimate of the LATE in the fuzzy regression discontinuity framework is obtained using the following first stage:

$$D = \alpha_1 Z_i + \alpha_2 e_i + \alpha_3 Z_i \times e_i + \epsilon_i$$

The predicted treatment value is then used in the following second stage:

$$Y_i = \tau_{IV}\hat{D}_i + \eta_2 e_i + \eta_3 Z_i \times e_i + u_i,$$

where the estimate of τ_{IV} is the LATE in the fuzzy regression discontinuity design.

D Empirical Bayes shrinkage

In this section I provide details on the Empirical Bayes shrinkage procedure outlined in Section 7.1.2. As mentioned in the main text, this procedure follows Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) closely. The estimates based on conditional independence and the control function approach each return a set of program specific estimates, $\left\{\hat{\beta}_p\right\}_{p=1}^{P}$, where for the control function I include the parameter on the correction term, $\hat{\psi}_p \in \hat{\beta}_p$.

Let K be the length of the vector of parameters. Under the hierarchical model outlined in Equations (15) and (16), the likelihood of the estimates for program p conditional on the unobserved parameters, β_p, and the associated covariance matrix, Ω_p, is

$$\mathcal{L}\left(\hat{\beta}_p|\beta_p,\Omega_p\right) = (2\pi)^{-K/2}\left|\Omega_p\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\left(\hat{\beta}_p-\beta^p\right)'\Omega_p^{-1}\left(\hat{\beta}_p-\beta^p\right)\right)$$

Assuming that my estimates of Ω_p are accurately approximated, the integrated likelihood function conditioning only on hyperparameters is then

$$\begin{aligned}\mathcal{L}^I\left(\hat{\beta}_p|\mu_\beta,\Sigma_\beta,\Omega_p\right) &= \int \mathcal{L}\left(\hat{\beta}_p|\beta_p,\Omega_p\right)dF\left(\beta_p|\mu_\beta\Sigma_\beta\right)\\ &= (2\pi)^{-K/2}\left|\Omega_p+\Sigma_\beta\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\left(\hat{\beta}_p-\mu_p\right)'\left(\Omega_p+\Sigma_p\right)^{-1}\left(\hat{\beta}_p-\mu_p\right)\right).\end{aligned}$$

Empirical Bayes estimates of the hyperparameters are obtained by maximizing the integrated log likelihood function where I plug in estimates, $\hat{\Omega}_p$, for Ω_p:

$$(\mu_p,\Sigma_p) = \arg\max\sum_p \log\mathcal{L}^I\left(\hat{\beta}_p|\mu_\beta\Sigma_\beta,\hat{\Omega}_p\right).$$

In Table 4, I report the square root of the diagonal elements of $\hat{\Sigma}_p$ under the parameters in $\hat{\mu}_p$. Using the estimates, $\left(\hat{\mu}_p,\hat{\Sigma}_p,\hat{\Omega}_p\right)$, posteriors of β_p and Ω_p are obtained as follows:

$$\begin{aligned}\beta_p^* &= \left(\hat{\Omega}_p^{-1}+\hat{\Sigma}_\beta^{-1}\right)^{-1}\left(\hat{\Omega}_p^{-1}\hat{\beta}_p+\hat{\Sigma}_\beta^{-1}\hat{\mu}_\beta\right)\\ \Omega_p^* &= \left(\hat{\Omega}_p^{-1}+\hat{\Sigma}_\beta^{-1}\right)^{-1}.\end{aligned}$$

The procedure is implemented in Python and code is provided in the data appendix.