# Trickle down education - ripple effects in college admissions

Mikkel Høst Gandil[*]

January 2024

**Abstract**

Admitting an applicant to a program will free up a slot in her next-best alternative to be filled. I investigate the significance of such ripple effects for policy evaluation by reconstructing a national allocation mechanism for higher education. Using estimated preferences and outcomes, I simulate thousands of supply policies. On average, 10 added slots affect 18 applicants, in some programs more than 50 applicants. Ignoring ripple effects misses 9 percent of variation in earnings and for 9 percent of programs, ripple effects more than offset the direct effect. The importance of ripple effects for policy likely generalizes to other markets.

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

# 1 Introduction

When a policy-maker 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. These ripple effects potentially create a wedge between the earnings returns as perceived from a local perspective and the broader social returns. Despite their potential importance for policy evaluation, ripple effects within education remain empirically unexplored.

Do ripple effects matter for the evaluation of supply-side policies in education? To answer this question one needs to know how the entire education market clears when capacity changes. Furthermore, one needs to know the expected change in outcomes, not only for the applicants directly affected by the policy but also for those applicants who are part of the ripple effect. To overcome these informational requirements, I exploit the almost ideal setting of centralized admission to higher education in Denmark. First, I reconstruct the centralized allocation mechanism for the entire higher education sector with a very high degree of precision. Using detailed population-level registry data, I then jointly estimate choice models and earnings returns to programs. Combining these two elements allows me to compute realistic counterfactual allocations and predict the effects of supply changes on predicted earnings not only for marginal applicants but for all applicants in the market. I then decompose the aggregate expected policy effects into the direct effects, which are the subject of most policy evaluation, and the ripple effects, which remain largely ignored in the literature. This study is the first attempt to quantify the full extent of ripple effects within a national market for higher education.

I begin by constructing a theoretical framework, wherein I demonstrate that the policy-relevant treatment effects in college matching markets are Local Average Treatment Effects (LATEs) for marginal applicants (Imbens and Angrist, 1994). However, only in a knife-edge case is a single treatment effect sufficient. Whenever applicants move between programs, one needs a treatment effect estimate for each affected program, as well as estimates of the sizes of the applicant flows between programs. The informational demands on a social planner are therefore much larger than generally suggested by the treatment effects literature on returns to education.

With this framework in place, I proceed to the empirical analysis of ripple effects. In a real-life setting applicants likely do not apply to programs where they are highly unlikely to be admitted. However, if chances of admission change so should the probability of applying to competitive programs. Using applicants' reported preferences thus likely leads to an underestimation of the ripple effects of supply changes, as missing applications break the chain of applicants affecting each other. To mitigate this issue, I estimate choice models that account for applicants not applying for infeasible programs and use these models to fill out applicants' rank-ordered lists of programs. With these filled-out lists and the reconstructed allocation model, I then simulate thousands of local supply changes. I find that ripple effects are large. When the capacity in a program is increased by 10 slots, on average 18 applicants are affected. The size of the ripple effects varies across programs, with some programs affecting more than 50

applicants when they increase capacity by 10 slots. These ratios are relatively stable, even as capacity changes become large. However, the average number of programs affected by ripple effects grows substantially with the size of capacity changes. In other words, the larger the change, the more treatment effect estimates are needed to assess the effect of a supply policy on all applicant outcomes. In terms of affected applicants, ripple effects tend to be larger for more oversubscribed programs and for programs that are often listed jointly with other programs. In line with expectations, I find slightly lower, but similar, effects using only the reported preferences of applicants.

I then proceed to evaluate the importance of ripple effects for returns to the simulated supply policies. I measure the returns in terms of expected earnings. In practice, estimating all the required LATEs is infeasible and the estimands from feasible research designs, such as the regression discontinuity design (RDD), cannot capture the effects of nonmarginal changes in supply. To go beyond the local nature of the LATEs, I estimate parametric models for potential applicant×program-specific earnings where I use residual variance from the choice models to control for selection. I then predict the full effect of counterfactual supply changes and compare them to the direct effects ignoring ripple effects. I find that for small supply changes, 9 percent of the variation in effects across simulations is missed by ignoring ripple effects. Overall, ripple effects tend to attenuate the direct effects on earnings, but there are considerable levels of heterogeneity across programs. Despite having similar educational content and estimated gains for marginal applicants, there are large differences in total gains across programs when accounting for ripple effects. A large share of programs have ripple effects that go in the opposite direction of the direct effect. This implies that a policy-maker may face a trade-off where applicants entering a program may gain (lose) but leave slots open in their original program for other applicants who lose (gain). For around 9 percent of the programs, the ripple effects more than offset the direct effects. In these cases, a predicted gain in the direct effect will be associated with an aggregate loss of earnings.

In the main part of the analysis, I focus on cases where a single program changes supply. In an additional analysis, I show that when expanding supply simultaneously in multiple programs, the direct effects on applicants in each program can change relative to a single program expansion. This underscores the importance of the *ceteris paribus* nature of traditional treatment effects and the potential lack of external validity and policy invariance of local estimates. Lastly, for the sake of completeness, I repeat the analysis using non-parametric LATEs from fuzzy regression discontinuity designs. While this approach has severe drawbacks in terms of precision and lack of empirical support, I largely replicate the magnitude of the ripple effects from the main analysis approach.

The qualitative insight that ripple effects matter for policy extends beyond centralized markets and education. An example of a decentralized market is US college admissions, where prestigious or oversubscribed universities may fail to internalize the broader social returns to increasing or decreasing supply. Further, ripple effects will also be relevant for housing and labor markets, where chains of residents or workers play the role of applicants. Ultimately, the ripple effects in matching markets are akin to pecuniary externalities which suggests that ripple

effects matter in a much wider array of markets. My results serve as a specific example of how chains of substitution can either amplify or counteract the direct effects of targeted policies. The findings underscore the importance of appropriately defining the scope of policy evaluations and constructing research designs to capture the relevant policy outcomes. I show that this is feasible in contexts where the market clearing mechanism is known.

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 et al., 2013; Zimmerman, 2014; Arcidiacono et al., 2016; Andrews et al., 2016; Mountjoy and Hickman, 2020), Several European and South-American studies investigate returns to specific fields, see Altonji et al. (2016) for a review. Among these, several studies employ regression discontinuity designs (RDD) to estimate returns to fields of study, either in terms of admission (Hastings et al., 2013) or completion (Kirkeboen et al., 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 et al., 2020; Andersen et al., 2020; Heinesen et al., 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 cutoffs play the role of prices in clearing demand and supply. Abdulkadiroğlu et al. (2017b) and Abdulkadiroglu et al. (2022) use a similar framework to construct research designs for the estimation of causal effects. Several articles exploit the properties of the large market to estimate welfare in centralized mechanisms (Abdulkadiroğlu et al., 2017a; Agarwal and Somaini, 2018; Kapor et al., 2020).[1] A set of papers focuses on outcomes other than welfare. Larroucau and Rios (2022) estimate a dynamic structural model of university applications to investigate dropout and college completion and Agarwal et al. (2020) link the mechanism of assignment of kidneys to survival. Lastly, Kapor et al. (2022) investigate the effect on college outcomes of an expansion of the centralized market. While the procedures of these papers share similarities with my approach to identifying potential outcomes they do not distinguish between direct effects and ripple effects, nor do they connect market clearing to the treatment effect literature.

A few papers investigate applicant flows explicitly. Agarwal (2015, 2017) investigates the effects of changing capacities on allocations in a medical residency market and Bucarey (2018) investigates the crowding out of applicants when introducing financial aid.[2] The importance of ripple effects for policy interventions has been investigated by Manning and Petrongolo (2017)

---

[1]Although my primary focus is not explicitly on applicant welfare, the Danish mechanism enables me to gauge changes in welfare in a straight forward way. 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).

[2]Additionally, Tanaka et al. (2020) provide reduced-form evidence that crowding out occurred when introducing centralized admission to education in Japan.

in the context of local labor markets where markets overlap geographically. In their working paper Kirkeboen et al. (2016) briefly investigate the indirect effects of increasing capacity in science programs using first-stage estimates of complier compositions conditional on the next-best alternative.[3] 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.[4] 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. I show that this conjecture is not generally valid for increases in supply of education.

The paper is structured as follows: Section 2 introduces the conceptualization of ripple effects within a theoretical framework, and 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, and 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 and section 8 concludes.

## 2   Ripple effects and policy evaluation in centralized mechanisms

To investigate policy evaluation with ripple effects I now construct a simple framework and show what the policy-relevant treatment effects are. I employ a stylized model of a centralized matching market as introduced by Azevedo and Leshno (2016).

A finite set of programs, $P = \{1, \ldots, P\}$, indexed by $p$, is matched to a continuum of applicants. The capacity of a program is given as $S = \{S_1, \ldots, S_P\}$ where each element is strictly positive. 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. $Y^\theta = \{Y_1^\theta, \ldots, Y_P^\theta\}$ is a vector of length $P$ of continuous potential outcomes and $X^\theta$ is a vector of length $k$ which contains observable covariates that do not vary across programs. With $\mathcal{R}$ as the set of all preference orderings, the set of all types is $\Theta = \mathcal{R} \times [0,1]^P \times \mathbb{R}^P \times \mathbb{R}^k$ and I define $\eta$ as the probability measure over $\Theta$. The economy is given by $E = (\eta, S)$.

A matching $\mu$ is a function that assigns applicants to programs and programs to applicants, where the outside option of being assigned is treated as a separate program. Formally, $\mu(\theta) = p$ iff $\theta \in \mu(p)$. I only consider stable matchings, where no applicant-program pair can block.[5] Azevedo and Leshno (2016) show that for sufficiently large economies, there exists

---

[3]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 and changes in the composition of compliers. Lastly, I estimate preferences and therefore account for truncated application lists.

[4]Feller et al. (2016) in the same context find differential effects depending on the counterfactual allocation.

[5]This implies that as Azevedo and Leshno (2016), I rule out cases of $(\theta, p)$ where $p \succ^\theta \mu(\theta)$ and either i)

a unique stable matching, which can be found by using Deferred Acceptance algorithm (Gale and Shapley, 1962). A cutoff $C_p$ is the minimum eligibility score required for admission into $p$: $C_p = \inf_{\theta \in \mu(c)} e_p^\theta \in [0,1]$. Define $C = \{C_1, \ldots, C_P\}$ as the vector of market-clearing cutoffs. Facing the vector of cutoffs $C$, the individual demand of type $\theta$, $D^\theta(C) \in P$, is defined as the favorite program among the set of programs where the eligibility score of $\theta$ clears the cutoff. In this matching market, cutoffs therefore play the role of prices in equating demand and supply. Aggregate demand for program $p$ is defined as

$$D_p(C) = \eta\left(\left\{\theta : D^\theta(C) = p\right\}\right),$$

and demands for all programs is contained in a vector $D(C) = \{D_1(C), \ldots, D_P(C)\}$. Azevedo and Leshno (2016) show that for any stable matching, a vector of cutoffs exists such that $\mu(\theta) = D^\theta(C)$. For the cutoffs to clear the market, the following must be satisfied:

$$D_p(C) \leq S_p \forall p$$
$$D_p(C) = S_p, \text{ if } C_p > 0,$$

which simply states that demand cannot exceed capacity and that a positive cutoff can only be observed for programs that are filled. Azevedo and Leshno (2016) show that the cutoffs and demand vary continuously with fundamentals. Finally, a matching will result in a realization of one potential outcome per type:

$$Y^\theta(C) = \sum_{p=1}^{P} \mathbf{1}(D^\theta(C) = p) Y_p^\theta,$$

which corresponds to the switching equation in the potential outcomes framework of Rubin (1974) and Imbens and Angrist (1994). The aggregate realized outcome in the economy is given by integrating over all types in the economy:

$$Y^T = \int_\theta Y^\theta(C) d\eta(\theta)$$

**Change in capacity**   As shown by Azevedo and Leshno (2016), the unique stable matching varies continuously with fundamentals, the distribution of types and capacities. This implies that cutoffs vary continuously when supply is changed. From stability, it follows:

$$\frac{dC_p}{dS_{p'}} \leq 0 \quad \forall p, p' \in (1, \ldots, P)$$

Intuitively, if capacity increases in some programs, applicants can only move to a more preferred program (or stay in their original program). This implies that cutoffs must go down (or remain unchanged). Due to uniqueness, there is a one-to-one mapping between changes in supply and changes in cutoffs, which allows me to express policy effects in terms of cutoffs. As demand is

---

$\eta(\mu(p) < S_p$ or ii) $\theta' \in \mu(p)$ and $e_p^{\theta'} < e_p^\theta$.

a continuous function of cutoffs, the gradient exists:

$$\frac{dD}{dC} = \begin{bmatrix} \frac{\partial D_1}{\partial C_1} & \cdots & \frac{\partial D_1}{\partial C_P} \\ \vdots & \ddots & \vdots \\ \frac{\partial D_1}{\partial C_P} & \cdots & \frac{\partial D_P}{\partial C_P} \end{bmatrix}. \tag{1}$$

From the properties of monotonicity and gross substitutes, it follows that the elements in the diagonal of (1) are weakly negative while the off-diagonal elements are non-negative.[6] The change in demand for a specific program following a change in supply can be expressed as a function of the cutoff vector:

$$dD_p = \left[ \underbrace{\eta\left(\left\{\theta : D^\theta(C+dC) = p, D^\theta(C) \neq p\right\}\right)}_{\text{Inflow}} - \underbrace{\eta\left(\left\{\theta : D^\theta(C+dC) \neq p, D^\theta(C) = p\right\}\right)}_{\text{Outflow}} \right] dC$$

If supply in program $p$ is unchanged and $C_p + dC_p > 0$, i.e. a cutoff exists after the change in capacity in another program, demand is unchanged as inflow is exactly offset by outflow. However, the composition of admitted applicants will shift. Assuming that $dC_p < 0$, the set of applicants moving into the program, i.e. inflow can alternatively be formulated as follows:

$$I_p = \left\{\theta : D^\theta(C+dC) = p, D^\theta(C) \neq p\right\} \tag{2}$$

$$= \left\{\theta : e_p^\theta \in [C_p + dC_p, C_p], e_{p'}^\theta \notin [C_{p'} + dC_{p'}, C_{p'}] \forall p' \succ^\theta p\right\}, \tag{3}$$

where I suppress the dependence of $I_p$ on $C$ and $dC$. The formulation on the right-hand side shows that applicants in $I_p$ are marginal on $p$ and non-marginal for more preferred programs.[7]

The change in outcomes for the applicants in the inflow to program $p$ is computed by comparing potential outcomes under the two matchings. The change in aggregate outcomes is given by

$$dY^T = Y^T(C+dC) - Y^T(C) = \int_\theta Y^\theta(C+dC) - Y^\theta(C)d\eta(\theta) \tag{4}$$

$$= \sum_p \int_{\theta \in I_p} Y^\theta(C+dC) - Y^\theta(C)d\eta(\theta), \tag{5}$$

which states that the aggregate change can be calculated as the sum of changes for applicants

---

[6]The outside option will never have a positive cutoff and the gradient for this notion is therefore zero.

[7]The set of applicants in the outflow can be similarly defined;

$$O_p = \left\{\theta : D^\theta(C+dC) \neq p, D^\theta(C) = p\right\}$$

$$= \left\{\theta : e_p^\theta > C_p, e_{p'}^\theta < C_{p'} \forall p' \succ^\theta p, \exists p' \succ^\theta p : e_{p'}^\theta \in [C_{p'} + dC_{p'}, C_{p'}]\right\}.$$

Note that as an outgoing applicant must go somewhere else, the set of applicants in the outflow is fully contained in the union set of all inflows, $O_p \in \{I_1, \ldots, I_P\}$ and vice versa, $I_p \in \{O_1, \ldots, O_P\}$. In other words, when looking at flows, to avoid double counting, one may either look at inflows or outflows. If $dC_p > 0$, i.e. when supply is constrained, the definition of the outflow and inflow are reversed.

who switch into programs.[8]

## 2.1 Linking theory and empirical policy-evaluation with treatment effects

To link the above framework to empirical policy evaluation it is fruitful to take a step back and consider a traditional policy evaluation framework where a policy is implemented and a policymaker wants to know what the effect on a given outcome for a particular population is. Assume that the policy, i.e. treatment, is binary. The effect of the policy is often summarized by a treatment effect which is some weighted average of the difference between the potential outcomes when treated and when untreated. For example, what is the effect of being admitted to an educational or training program? Estimating treatment effects implies an implicit stance on the population of interest. In the example, an average treatment effect among the potentially admitted may be sufficient if all the policymaker cares about is the population of eligible applicants. However, as outlined above, due to ripple effects, some applicants can be affected even though they never will enter a program that changes capacity.

A policymaker in government will in many cases care about more than the marginal applicants to a particular program. This would for example be the case if education is regulated by the government or if an institution provides multiple programs. In such cases, looking only at the applicants on the margin of a single program will be insufficient to assess the returns to a given policy.

Assume that a policymaker considers the whole economy and either has uniform welfare weights or cares only about output. Then equation (5) contains the necessary elements for policy evaluation of capacity changes. However, by their nature, only one potential outcome is ever observed, and thus the unrealized outcomes need to be estimated using available data. I provide two ways of doing this. First I show that in principle, the relevant information can be summarized by local average treatment effects and flows of applicants which puts minimal assumptions on the heterogeneity of program-specific returns. However, in most applications, this method will likely be infeasible, and I therefore provide a second method that imposes more structure on the potential outcomes.

### 2.1.1 Policy evaluation using non-parametric treatment effects

A common approach to estimating returns to education is to exploit the discontinuous rise in admission probability around admission cutoffs, see Hastings et al. (2013); Kirkeboen et al. (2016). In a fuzzy RDD framework with admission as the endogenous treatment variable and cutoff crossing as an instrument, the estimand will be a local average treatment effect (LATE) for applicants located close to the cutoff who are admitted when exceeding the cutoff. This set is traditionally called compliers. This set of compliers is similar to the inflow set, $I_p$, defined in equation (3). Assume that only the cutoff of $p$ changes marginally, while all other cutoffs remain constant. The set of incoming students to $p$, $I_p$, will in this case equal the set of compliers from the RDD design. The LATE from a program-specific fuzzy RDD can in this case be equivalently

---

[8]Non-assignment is here treated as a program, such that movement along the extensive margin is captured. Note that the expression for the aggregate change can similarly be written in terms of outflows.

be defined as a function of types and cutoffs:

$$LATE_p = \frac{1}{\eta(I_p)} \int_{\theta \in I_p} \left[ Y^\theta(C + dC) - Y^\theta(C) \right] d\eta(\theta).$$

The effect on aggregate outcomes of $dS_p > 0$ is now given as

$$dY^T = Y^T(C + dC) - Y^T(C) = \sum_p LATE_p \times \eta(I_p). \tag{6}$$

The formulation in (6) suggests a way forward using estimates of LATEs and motivates the literature that seeks to estimate such estimands. Assuming that $LATE_p$ is constant for different, but small, changes in cutoffs, the change in aggregate outcomes can then be approximated by estimating program-specific LATEs and the size of the complier set, i.e. the number of students on the margin. Equation (6) shows that to compute the total effect on outcomes from a supply change, one must obtain estimates of LATE for each program where applicants move into, i.e. where $\eta(I_p) > 0$. In other words, only in the case where the alternative of all marginal applicants is unrestricted is a single LATE sufficient for assessing returns to program expansions.[9]

**Ripple effects and defiers**   The above argument assumes that only a single cutoff changes. However, this can only be the case if the alternatives of the marginal applicants are unrestricted. If this is not the case, ripple effects will occur and multiple cutoffs must change. When more than one cutoff changes the set of incoming compliers may no longer equal the compliers from the discontinuity design. Even in this case, however, the estimated LATE from a discontinuity design may be a good approximation of the LATE for the relevant set of incoming applicants. However, in addition to changing the estimand, ripple effects threaten the identification of the LATE in a discontinuity design more broadly.

The identification of relevant LATEs is based on the monotonicity assumption, which states that a decrease in a cutoff makes applicants weakly more likely to be admitted, i.e. there can be no defiers. If this assumption does not hold, the estimate will be a mix of the treatment effect for compliers and defiers, who do not enroll in the program. Monotonicity is likely to be invalid when programs have separate eligibility scores or multiple quotas. As an example of the former, assume that there are two programs, $p$ and $p'$, which rank applicants differently. In addition, consider two applicants, $A$ and $B$, who differ in both preferences and program-specific eligibility. Suppose $p \succ_A p'$, and $p' \succ_B p$ and that $e_p^A < C_p < e_p^B$ and $e_{p'}^B < C_{p'} < e_{p'}^A$. In this case, $A$ is admitted to $p'$, and $B$ is admitted to $p'$ though both would prefer to trade place. Thus the matching is stable but not Pareto-efficient from the perspective of applicants. Now suppose the supply in $p$ is expanded sufficiently for $A$ to clear the new cutoff in $p$ which implies that she is admitted. This leaves room in $p'$ which causes the cutoff in program $p'$ to fall sufficiently for

---

[9]On the other hand, equation (6) also highlights what is *not* needed for assessing the effects of supply changes. As long as one estimates LATEs using a fully stratified design, one does not need to know the alternative of the compliers, thus sidestepping the issue of multiple treatment effects investigated by Kirkeboen et al. (2016); Heinesen et al. (2022). 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 et al. (2008).

$B$ to be admitted in $p'$. In the LATE framework, this means that $B$ is a *defier* in program $p$; a marginal expansion of program $p$ causes her to not be admitted to $p$. A similar example can be constructed where programs use quotas. I show that such circles occur in practice in the Danish context in section 7.2.

The focus on ripple effects clarifies that defiers may theoretically be present in allocation systems where programs use different eligibility criteria or multiple quotas. This implies that with heterogeneous treatment effects, non-parametric LATEs are not readily identifiable in frameworks such as the ones proposed by Abdulkadiroglu et al. (2022). Few, if any, admissions systems, centralized or not, use a single tie-breaker *and* a single quota per program, both of which are necessary to prevent circles and thus the presence of defiers. However, for defiers to be a problem for discontinuity design, they must be located at the cutoff. To what degree this is the case is an open question.[10]

**Additional practical limitations in the discontinuity design**    Estimating the relevant LATEs comes with additional practical complications. Programs are often small resulting in noisy and unstable estimates, especially with small compiler shares. The design requires a first stage, which is only available for oversubscribed programs. This implies that if a policy would cause some programs to become oversubscribed, one could not estimate the relevant LATE. A solution to this issue, and to increase efficiency, is to estimate returns to fields rather than programs as is common in the literature (e.g. Kirkeboen et al. (2016)). In such approaches, samples for different programs are pooled and the resulting field-specific LATE is for a mix of compliers for different programs. For policy analysis, polling across programs introduces two distinct challenges. Firstly, for a specific set of compliers, returns might differ across programs. Secondly, for a given policy, the set of compliers might differ, both across programs and for different sizes of supply changes.

### 2.1.2   Policy evaluation using models of potential outcomes

Even abstracting from the issues outlined above, the link between supply changes in a school choice market and the fuzzy RDD is only valid for *marginal* changes in capacities and hence cutoffs. For large changes, this link breaks down. The set of incoming applicants will no longer be located at the cutoff and the change in realized outcomes will therefore not be identified in the discontinuity design. In addition, with large changes in cutoffs, the probability that applicants are "marginal" for multiple programs rises. This motivates an alternative approach to estimating the relevant outcomes for policy evaluation.

In a realized economy with knowledge of preferences, eligibility scores, and supply, one can identify members of the program-specific set of incoming students by clearing the market for counterfactual supply. If one knew the potential outcomes, one could compute the aggregate

---

[10]Abdulkadiroglu et al. (2022) sidesteps this issue by assuming that program effects are homogeneous.

effects:

$$dY^T = \sum_p \sum_{p' \neq p} \int_\theta \left[ Y_p^\theta - Y_{p'}^\theta \right] \mathbf{1}\{\theta : D^\theta(C + dC) = p, D^\theta(C) = p'\}d\eta(\theta). \tag{7}$$

Equation (7) provides an alternative approach to characterize effects of supply-side policies, even when such policies result in non-marginal changes in cutoffs. This approach uses simulation of the assignment mechanism to compute the composition of the sets of margin-specific movers with very few restrictions. However, to do that, one must know potential outcomes or have an estimate thereof. In this paper, I assume that the potential outcomes are log-linear in observables:

$$Y_p^\theta = \exp(X^\theta \beta_p), \tag{8}$$

where $X^\theta$ is a vector of observable characteristics of type $\theta$ and $\beta_p$ is a program-specific vector of coefficients. I discuss the estimation of $\beta_p$ in section 7 where I also outline how I control for selection. The assumption in equation (8) along with knowledge of preferences, $\succ^\theta$, allow me to compute the full counterfactual effect of program expansions in (7).[11] Relative to the formulation with LATEs outlined in equation (6), 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. On the other hand, by simulating the assignment mechanism, I do not impose assumptions on the composition of compliers. Given knowledge of preferences, eligibility, capacities, and estimates of $\beta_p$, I can calculate all the empirical equivalents of the elements of equation (7), which I do in section 7.

**Caveats and implicit assumptions**  My outline 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 and certainly not in the Danish context (see the next section for institutional context). Applicants often only list a subset of 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. Secondly, in practice, I cannot know $Y^\theta$ but only estimate the expected outcomes for the applicants in the market given the assumption of log-linearity. I assume that the mapping from applicant observables to program-specific potential outcomes is fixed through the coefficients, $\{\beta_p\}_P$. This implies stable skill prices and rules out the 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 the 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

---

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

flows, as applicants can only affect each other in the cross-section.

These assumptions and limitations constrain the nature of my set of counterfactual equilibria and I do not claim to capture all general equilibrium effects, neither in terms of affected applicants nor changes in outcomes. Rather, the ripple effects analyzed in this paper are observed in the cross-sectional dimension and the simulations 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 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 are eligible in Q1 if they fulfill course requirements set by the program. 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 ministry only observes the final ranking in Q2, see Gandil and Leuven (2022) for details.

Each applicant provides a rank-ordered list 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 to be assessed. 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.[12] In the mechanism, along with this modified

---

[12]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

rank-ordered list, an eligibility score is observed for each applicant in each relevant quota.

Program admission officers 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. Before 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.[13]

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 et al. (2010); Fack et al. (2019), applicants might leave out unrealistic programs at the top of the rank-ordered list. Further, applicants might leave out programs below a program that they perceive as 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 the 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 shows that cutoffs in Quota 1 tend to be stable from year to year, which allows applicants to infer admission probabilities from previous admission rounds.[14]

In contrast to settings such as the United States, it is very common in Denmark to take at least one gap year between graduating high school and enrolling in higher education. In 2012, 70 percent of the graduating high school cohort chose to take at least one year off before entering higher education (DST, 2023). Further, dropouts are common. Around 16 percent of first-year students drop out, but half of these applicants typically enroll in another program the year after. Of those applicants enrolling in a bachelor program, 85 percent have obtained a degree in higher education approximately six years later (UFM, 2018). Thus, dropouts that result in no education are rare and a large share of those applicants who are not assigned in a given year will likely reapply in later years.

## 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 applicants correctly in numerous years. To illustrate the simulation model, Figure 1 presents

---

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.

[13]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.

[14]Cutoffs in Q2 are not informative as programs vary in their ranking function which is not generally known – nor in most cases formalized.
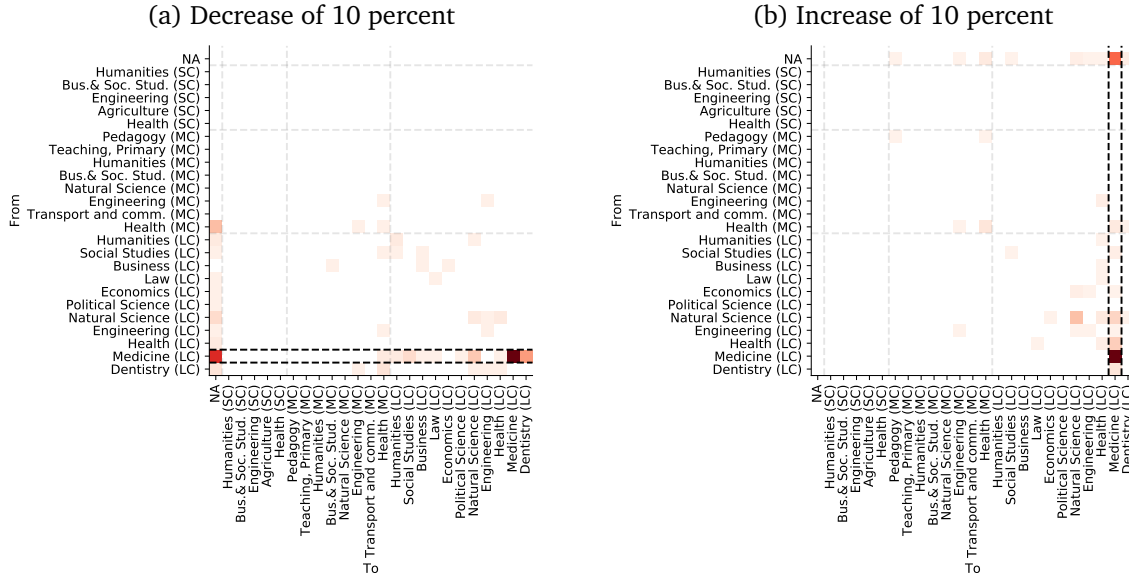
Figure 1: Applicant flows following a change in capacity for Medicine at KU

**Note:** 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.

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. 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 the 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 reflects that many 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. Fortunately, the simulation model keeps track of these flows which accurately reflect what would have occurred had the supply been changed while holding demand constant.[15] Again, holding demand constant is a restrictive assumption, which I do not make in the main analysis.

---

[15]These simulations are performed for the 2016 admission round, which is the same round that I will perform simulations on in Section 6.

# 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. I construct two separate samples which I describe below. The first sample consists of multiple years of data which I use to estimate preferences and potential outcomes. The second sample consists of the population of applications in a single admission round and is used for the simulation of counterfactual allocations.

**Estimation sample**   For estimation of outcomes, I take all applications from applicants in the admission data from 2006 to 2011 for whom I observe GPA, a Quota 1 waiting-list number, and positive income within three years in the registers. From the production data, I retrieve rank-ordered lists. I merge this data with national registries to retrieve GPA, courses and grades, residential location, gender, and immigrant background. As a measure of skill, I calculate the GPA rank within the 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.

   The income concept is the log of average personal pre-tax income excluding transfers 7 to 9 years after application. Most graduates from academic programs go on to earn a 2-year master's degree. Except for 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.

   In the production data, the program identifier is not stable over time. However, in the registers, I observe enrollment on a separate time-stable identifier. I create a crosswalk from the production data using dominant enrollment outcomes for admitted applicants and manually assigning programs with less than 95 percent of admitted applicants in the dominant program identified in the registers.

**Simulation sample**   The sample used for the simulations of capacity changes contains the population of applications in 2016. For applicants in the registers, I retrieve the same variables as in the estimation sample. 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 includes potential bonuses given due to a short period between graduating high school or due to raising high school subjects to a higher level.[16] For foreigners, the GPA conversion follows the guidelines of the Ministry of Higher Education. The conversion is done by the

---

[16]In this period, applicants applying within two years of high school graduation could multiply their GPA with 1.08.

program administrators and these GPAs are subject to error.[17] 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 a small 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 as well as GPA-cutoff in Quota 1.[18] Ripple effects occur because applicants apply to multiple programs. To capture this, I constructed a weighted network graph of programs based on the preferences of the 2016 applicants. In this network, an edge between two programs exists 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. This centrality measure ranks programs higher if they are themselves connected to higher ranked programs, similar to how Google ranks websites. I standardize cutoffs and centrality to have a mean of zero and a standard deviation of one. 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 known by researchers in other institutional contexts where the data is less rich.

The first column in table 1 shows summary statistics for applicants in the estimation sample, which is used for estimating preferences and potential outcomes. The second column shows the corresponding statistics for the sample used to simulate counterfactual allocations. For the simulation sample, I only present statistics for applicants who are found in the registers, which is 77 percent. The statistics are fairly balanced across samples, though this is no requirement for the validity of the analysis. The average realized income is on average 306 DKK in 2015-prices, roughly equivalent to 45 thousand USD.[19] This is relatively low, which reflects that I measure early-career earnings.

## 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 (Fack et al., 2019). The reported rank-ordered lists (ROLs) 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 fill out the 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

---

[17]For instance, the same applicant can be observed with multiple values of the alternative GPA across applications.
[18]In programs where there is a single Quota 1 this is defined as the GPA of the first rejected applicant. 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 they have a lower cutoff than the one reported by the ministry.
[19]The average exchange rate in 2015 was 6.72 DKK/USD.

Table 1: Descriptive statistics

|  |  | Estimation | Simulation |
|---|---|---|---|
| GPA rank | Mean | 0.55 | 0.51 |
|  |  | (0.28) | (0.28) |
| Female | Mean | 0.64 | 0.57 |
|  |  | (0.48) | (0.49) |
| Danish grade | Mean | 5.91 | 6.11 |
|  |  | (2.75) | (2.97) |
| Math grade | Mean | 5.61 | 5.78 |
|  |  | (3.93) | (3.91) |
| Danish grade missing | Share | 0.03 | 0.02 |
|  |  | - | - |
| Math grade missing | Share | 0.29 | 0.26 |
|  |  | - | - |
| A-level: Business | Share | 0.14 | 0.15 |
|  |  | - | - |
| A-level: Humanities | Share | 0.87 | 0.83 |
|  |  | - | - |
| A-level: STEM | Share | 0.45 | 0.37 |
|  |  | - | - |
| Income 8 years after, DKK | Mean | 306.32 | - |
|  |  | (146.72) | - |
| N applicants |  | 160,848 | 82,794 |
| Share w. missing cov |  | 0.00 | 0.23 |

**Note:** The table reports summary statistics. The first column reports statistics on the sample used for estimating outcome equations and preferences. Observations are on an applicant level. The last column reports same the statistics for the simulation sample. Missing variables are ignored in computing means and standard deviations. Income is not recorded for the simulation sample. Standard deviations for continuous variables are reported in parentheses.

to strata $s$ receives the following utility of admission to program $p$:

$$U_{ip} = V(W_{ip}, \gamma_s) + \nu_{ip}, \tag{9}$$

where $V$ is a known function of observables, $W_{ip}$, and an unknown strata-specific vector, $\gamma_s$, to be estimated. $W_{ip}$ contains dummies for field, program length, and location (administrative region). To capture heterogeneity within strata, I interact program length with high school GPA, and program field 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 were truthful, equations (9) could be estimated using rank-ordered logit (i.e. 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 et al. (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.[20] 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 stratum. With the estimated model parameters, I predict the 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. I combine the original and the new rank-ordered lists to generate filled-out lists, which I use for my main analysis. Appendix B provides further details on specification and estimation.[21] Relative to the theoretical framework in section 2, the empirical section is framed in terms of individuals, $i$ rather than types $\theta$. In the empirical analysis, for given strata, $s$, $W_i$, and reported preferences, applicants are indistinguishable and these three elements therefore identify the types.

**Predicted utilities and imputed rank-ordered lists**     The random utility model in equation (9) is estimated within strata and the estimated coefficients 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. Dashed lines surround the home region. I find a strong home bias as the home region fixed effect is the highest within all strata. Neighboring regions also tend to be more attractive than other regions. The Copenhagen metropolitan region (the Easternmost 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 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 3.1 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 models approximate preferences well and corroborate the findings of Pathak and Shi (2021).

To assess where truncation of rank-ordered lists may be an issue, figure 3 shows the probability of applying for a highly competitive program as a function of high school GPA. For most applicants, the probability of applying to a competitive program is the same using the original and the filled-out rank-ordered lists. However, for the three top deciles, the filled-out lists contain more competitive programs. Appendix figure A.2 suggests that this increased propensity to list programs is largely explained by filling in competitive programs *below* the existing application. As these safety programs are competitive, this filling out is likely to amplify ripple

---

[20]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 Quota 2 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 potential rank of an applicant in Quota 2 had she applied, if she did not in my data.

[21]The model fails to converge for immigrant min in Southern Jutland. For this stratum, I instead include their original reported preferences.

(a) Northern Jutland        (b) Middle Jutland

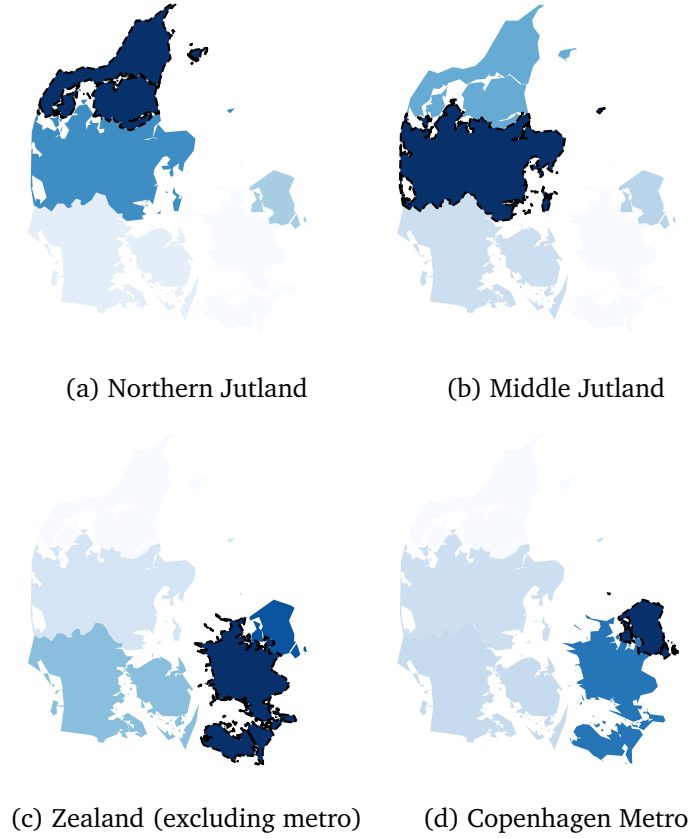(c) Zealand (excluding metro)    (d) Copenhagen Metro

Figure 2: Estimated regional preferences

**Note:** 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 stratum. The island of Bornholm is excluded from the maps. The model for Southern Jutland region for immigrant males is not shown as it does not converge.

effects from contractions in supply. Even with filled-out rank-ordered lists, the lists remain short due to a high value of the outside option. This partly reflects the institutional setting, where high-school graduates often postpone higher education by a year or more and where there are list (social) cost to taking a gap year. Thus the limited length of lists reflects that application to higher education is a dynamic problem (Larroucau and Rios, 2022), and counterfactual simulations will not capture dynamic adjustments of application behavior. In the analysis that follows I will show results using the filled-out rank-ordered lists. Results for the reported lists are relegated to the appendix.

# 6 Ripple effect in flows of applicants

I now proceed to investigate the ripple effects by simulating counterfactual assignments. 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 to assess the influence that ripple effects have on policy evaluation.

Table 2: Imputed Rank-ordered lists

|  | Mean | Min | Q25 | Q50 | Q75 | Max |
|---|---|---|---|---|---|---|
| Lenght of reported ROL | 2.49 | 1 | 1 | 2 | 3 | 8 |
| Lenght of estimated ROL | 3.14 | 1 | 2 | 4 | 4 | 697 |
| Correct ordering \| Length of ROL>2 (mean) | 0.58 | | | | | |

**Note:** The table shows descriptive statistics on the reported list of ranked programs and filled-out lists.
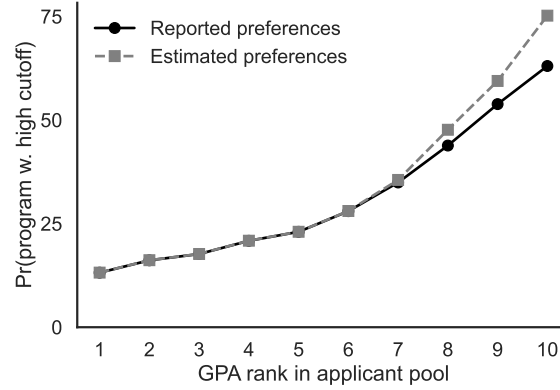


Figure 3: Evidence of truncation of preferences

**Note:** The figure 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.

## 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 the 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 rank-ordered lists and once with the filled-out lists. 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 between direct flows and ripple flows. The direct flow is the number of applicants moving into or out of the program where supply is altered in a given simulation. Ripple flows are all flows that are not direct flows and the total flow is the sum of the direct flows and the ripple flows, i.e. the number of applicants who are admitted to a different program than the baseline (including being not admitted).

Figure 4a shows the distribution of the total flow relative to the direct flow. On average. across the experiments, around 18 applicants are affected in total every time 10 applicants are affected on the program margin. In other words, for every 10 applicants let into a program, 8 are affected elsewhere. The dispersion is large. At the 90th percentile, experiments induce

more than twice as many indirect shifts as direct shifts. The magnitude of the ripple effect in terms of flows is more or less symmetric around zero, which means that it does not depend on whether the ripple effects are due to contractions or expansions of supply. The distribution of ripple effects appears fairly stable regardless of the size of the supply change. This implies that applicants further from the cutoff are not part of longer chains than those closer to the cutoff.

Appendix figure A.3 shows the same graph but with simulations based on reported preferences. Though the patterns are similar, magnitudes are generally somewhat smaller. On average, when a counterfactual supply induces 10 applicants to move in or out of a program, this is associated with 15.7 applicants changing allocations. This is expected as reported rank-ordered lists are generally shorter than the imputed lists. Holding applications constant thus may underestimate the magnitude of the ripple effects in terms of affected applicants.

As discussed in section 2, policy evaluation with ripple effects requires knowledge of treatment effect for all margins, where applicants move. Figure 4b quantifies on how many margins one needs to quantify changes in realized outcomes. The figure shows the distribution of the number of programs affected by a capacity change as a function of the percentage change. As expected, the number of margins is increasing in the size of the capacity change. However, even for small capacity changes, one needs multiple margins. If a program change capacity by 5 percent, then on average 5 programs are affected, with a 90th percentile of 10. For larger changes, the number of affected programs grows to over 30 on average with a 90th percentile of 60 for large expansions.

Programs that create large ripple effects tend to be within long-cycle health programs such as medicine and dentistry as shown in figure 5. Public health creates the largest effect with a ratio of nearly nine. In contrast, when simulations are based on reported preferences, business programs dominate as shown in appendix figure A.4.

**Predicting ripple effects**  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, as discussed above, 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 and present the results in table 3.

Both centrality and cutoffs show a strong positive association with larger ripple effects. As evidenced in column (1) in table 3, an increase of one standard deviation in centrality is associated with an increase in the ratio of 0.25 corresponding to 14 percent increase relative to the average of 1.77. An increase of the GPA cutoff by a standard deviation predicts an increase of the ratio by 21 percent (0.38/1.77) as evidenced in column 3. Centrality is less predictive than cutoffs as measured by R-squared. These conclusions hold 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
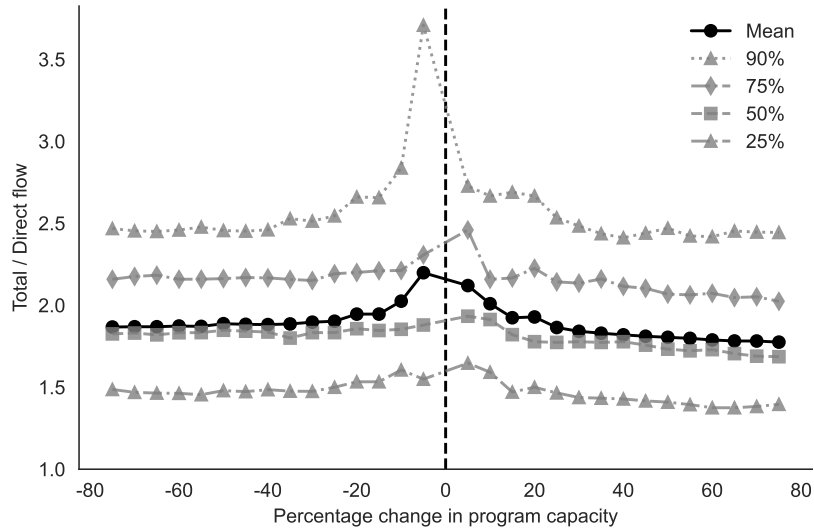
(a) Total flow relative to direct flow
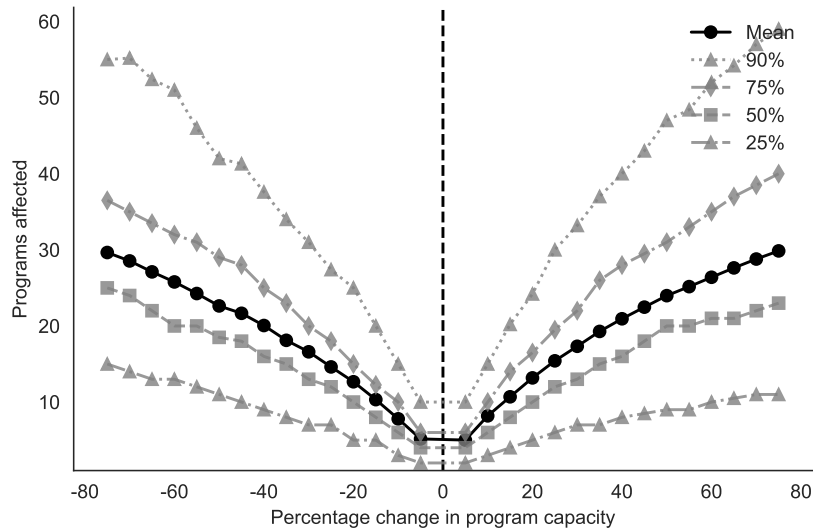


(b) Number of programs affected

Figure 4: Ripple effects as a function of capacity changes

**Note:** Figure 4a plots the ratio of the total number of movers relative to the direct 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 direct margin, are left out due to anonymity requirements. Figure 4b plots the number of programs where applicants either move in or out.

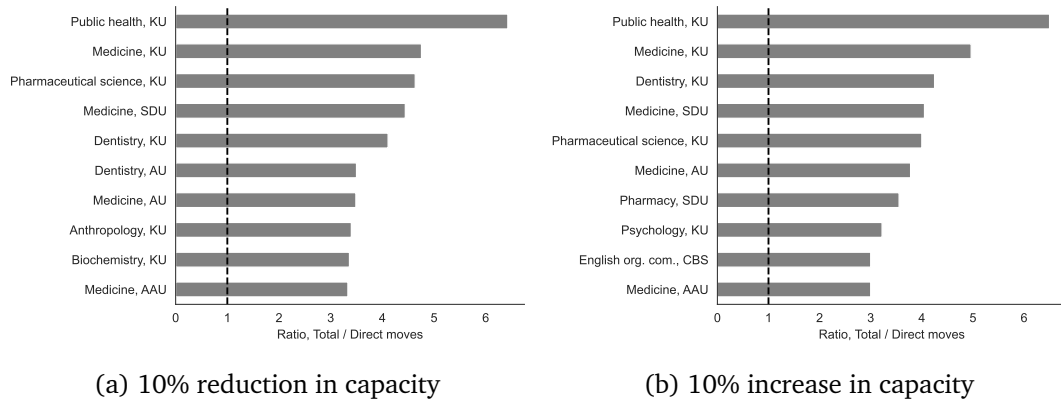(a) 10% reduction in capacity      (b) 10% increase in capacity

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

**Note:** Figures 5a)-b) plots the ratio of the total number of movers relative to the direct 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.4 contains equivalent results using reported preferences, rather than estimated preferences.

Table 3: Predicting ripple effects in flows

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Eigenvector centrality | 0.25 | 0.18 |  |  | 0.07 | 0.04 |
|  | (0.06) | (0.03) |  |  | (0.03) | (0.05) |
| GPA cutoff |  |  | 0.38 | 0.28 | 0.35 | 0.26 |
|  |  |  | (0.07) | (0.02) | (0.03) | (0.04) |
| Constant | 1.77 | 1.77 | 1.77 | 1.77 | 1.77 | 1.77 |
|  | (0.10) | (0.01) | (0.08) | (0.00) | (0.02) | (0.01) |
| Observations | 295 | 294 | 295 | 294 | 295 | 294 |
| R2 | 0.20 | 0.59 | 0.47 | 0.67 | 0.49 | 0.68 |
| Field FE |  | X |  | X |  | X |
| Institution FE |  | X |  | X |  | X |

**Note:** This table shows 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 programs 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 a standard deviation of 1 across programs. Standard errors are in parentheses and clustered on the level of the fixed effects.

cutoff remain essentially unchanged.[22] The functional relationships in Table 3 are shown in Appendix Figure A.5, where it seems that the expectation is well-modeled by a linear function.

The predictability of ripple effects is a function of the market size, clearing, and the preferences of applicants and eligibility criteria. However, I conjecture that popular programs are likely to create larger ripple effects in other contexts as well.

# 7   Program evaluation with ripple effects

In this section, I assess the importance of ripple effects for policy evaluation. To accomplish this, I estimate predicted potential outcomes and use these estimates to calculate changes in expected earnings as a result of local supply changes. I also investigate the effect of changing capacity in multiple programs simultaneously and the implication that such policies might have

---

[22]The correlation between the two measures is shown in Appendix Figure A.6.

for the external validity of local average treatment effects (LATE). As a test of robustness of the magnitude of ripple effects, I also estimate LATEs in a regression discontinuity design, though as I have discussed and will show, this approach comes with considerable drawbacks.

As outlined in section 2, I capture the full effect of a capacity change on earnings through predicted potential earnings. To the extent that observables correctly model heterogeneity in outcomes, this approach allows me to handle changes in earnings due to the composition of flows across programs. The mapping depends on parameters, which need to be estimated while accounting for selection bias. I model the expected log of earnings, $y_{ip}$, conditional on being admitted to program $p$ for individual $i$ is a linear function of observables, $X_i$, and an idiosyncratic component. Let $\nu_i = \{\nu_{i1}, \ldots, \nu_{iP}\}$ be the vector of individual$\times$program-specific residuals from the preference model in (9). I assume that conditioning on the residuals from the preference model is sufficient to control for endogenous selection:

$$E[y_{ip}|X_i, \nu_i] = X_i\beta_p + E[u_{ip}|\nu_i], \tag{10}$$

$$E[u_{ip}|\nu_i] = \sum_{p\in P} \delta_p(\nu_{ip} - \bar{\nu}_p), \tag{11}$$

where $\bar{\nu}_p = E[\nu_{ip}]$. As the preference model is a conditional logit, $\nu_{ip}$ are Gumbel distributed, and the setting and the selection terms are functions of the predicted probabilities as shown by Dubin and McFadden (1984) and Bourguignon et al. (2007). Such restrictions on outcomes are common in settings with a large set of treatments(Abdulkadiroğlu et al., 2020).

Identification is achieved by excluding part of the conditioning set in the preference model, $W_{ip}$, from the set of covariates used to predict outcomes, $X_i$. Similarly to Dahl (2002), I assume that the region of origin does not affect outcomes and I exclude immigrant status from the outcome equations. I also exclude indicators for A-levels except for STEM due to a lack of support in smaller programs. Due to the stability assumption, applicants only select within a set of feasible programs, which adds a source of identification as cutoffs change from year to year.

The set of $\beta_p$ is estimated in a single regression by interacting $X_i$ with program dummies and including the non-interacted correction terms.[23] To account for the uncertainty due to the estimation of the selection terms in the preference models, I apply the score bootstrap procedure developed by Abdulkadiroğlu et al. (2020). To investigate the role of selection bias, I also estimate the set of $\beta_p$ under the assumption that $E[u_{ip}|X] = 0$, which is a standard conditional independence assumption. Some programs have too small pools of admitted applicants to estimate $\beta_p$. To fill in potential outcomes for applicants in these programs, I estimate field-

---

[23]As shown by Dubin and McFadden (1984), the correction terms take the following values:

$$E[\nu_{ip} - \bar{\nu}_p | i \in p, W_i, \gamma] = -\log(P_{ip})$$

$$E[\nu_{ip'} - \bar{\nu}_{p'} | i \in p, W_i, \gamma] = \frac{P_{ip'}\log(P_{ip'})}{1 - P_{ip'}},$$

where $W_i$ is the set of $W_{ip}$ for all programs, and $\gamma$ varies by strata. I set the probability of infeasible programs to zero and note that the limit of the selection terms for non-chosen alternatives is zero: $\lim_{P\to 0} \frac{P\log(P)}{1-P} = 0$. Estimating fully stratified models is not feasible. The number of selection terms will cause the design matrix not to have full rank for most programs. Identification is thus obtained by assuming that $\delta_p$ is the same across programs.

level versions of equation (10).

**Joint distribution of parameters, shrinkage, and prediction**  Under the above-stated assumptions, the sets of estimates of $\beta_p$ are unbiased but noisy measures of the true parameters. As the estimates are needed for predictions, I follow Abdulkadiroğlu et al. (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:

$$\hat{\beta}_p|\beta_p \sim \mathcal{N}\left(\beta_p, \Omega_p\right) \tag{12}$$

$$\beta_p \sim \mathcal{N}\left(\mu_\beta, \Sigma_\beta\right). \tag{13}$$

Estimates of $\mu_\beta$ and $\Sigma_\beta$ are obtained through maximum likelihood estimation of (12) and (13), where $\Omega_p$ is approximated by the estimated covariance matrices of the estimation methods described above. The resulting estimates of the hyperparameters, $\mu_\beta$, and $\hat{\Sigma}_\beta$, are in turn used to obtain empirical Bayes posterior means for $\beta_p$:

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

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 C.

The predicted outcomes in the simulation sample are calculated as

$$\hat{Y}_{ip} = \exp\left(X_i\beta_p^*\right). \tag{14}$$

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. With the predicted potential outcomes and the estimated preferences, I now have the empirical equivalents of the type-specific fundamentals in section 2.

**Estimated returns to programs**  Table 4 presents the mean and standard deviations (in parenthesis) of the posterior distribution of program-specific parameters.[24]  Columns 1 shows estimates on a program level not controlling for preferences, while column 2 shows the results when including the selection terms. The averages are broadly similar except for a lower average parameter for GPA when controlling for selection and a higher parameter for having an A-level in STEM. The same pattern is seen when estimating the models on a field level in column 3 and 4, though the magnitudes are slightly different. Using the sets of program-specific parameters, I predict earnings in my simulation sample for all combinations of programs and applicants. For

---

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

Table 4: Joint distribution of estimates

|  | (1) log($Y$) | (2) log($Y$) | (3) log($Y$) | (4) log($Y$) |
|---|---|---|---|---|
| Constant | 5.289 | 5.690 | 5.500 | 5.393 |
|  | (0.727) | (0.278) | (3.039) | (0.354) |
| GPA | 0.328 | 0.077 | 0.288 | 0.215 |
|  | (0.517) | (0.094) | (0.601) | (1.282) |
| Female | -0.105 | -0.114 | -0.140 | -0.101 |
|  | (0.375) | (0.100) | (0.353) | (0.109) |
| Danish grade | -0.003 | -0.003 | -0.005 | -0.001 |
|  | (0.036) | (0.013) | (0.083) | (0.003) |
| Math grade | 0.001 | 0.002 | 0.007 | 0.004 |
|  | (0.034) | (0.010) | (0.022) | (0.020) |
| Danish grade missing | 0.004 | 0.018 | 0.015 | 0.004 |
|  | (0.569) | (0.358) | (0.537) | (0.240) |
| Math grade missing | 0.054 | 0.001 | 0.063 | 0.029 |
|  | (0.285) | (0.081) | (0.101) | (0.034) |
| A-level: STEM | 0.145 | 0.209 | 0.116 | 0.174 |
|  | (0.283) | (0.122) | (0.379) | (0.153) |
| Year | 0.009 | 0.057 | 0.019 | 0.011 |
|  | (0.064) | (0.010) | (0.263) | (0.049) |
| Level | Program | Program | Field | Field |
| Selection correction | - | CF | - | CF |
| N | 382 | 382 | 26 | 26 |

**Note:** The table displays the estimated joint prior of parameters, $\mu_\beta$ in equation (13) estimated from the program level estimates. The dependent variable is the log of average income 7 to 9 years after admission. Standard deviations of the distribution, shown in parentheses below estimates, are the square root of the diagonal part of $\Sigma_\beta$ from equation (13) and cannot be used for hypothesis testing.

the programs in the simulation sample without sufficient support in the estimation sample, I use field-specific estimates to predict earnings.

To assess the credibility of the predicted outcomes I construct a setup where I mimic a regression discontinuity design in figure 6. I rank applications by distance to the cutoff and center the data at the program-specific cutoff in the realized allocation. As I predict all potential outcomes, I can plot the average of the outcomes conditional on being admitted to the first-priority program and the average of the next-best alternative. Reassuringly, predicted outcomes are increasing in the ranking, which reflects the average positive slope on GPA in table 4. The vertical distance between points represents the predicted treatment effects of going from the next-best alternative to the preferred program. The differences are small and negative, indicating that applicants tend to list programs with similar predicted outcomes. There is, however, substantive heterogeneity in individual returns. Further, these predicted treatment effects do not correspond to a real change in capacity. Thus figure 6 only serves as a validation exercise. I now proceed to investigate realistic supply changes where applicants can through their entire rank-ordered list as well as remain unaffected by a supply change.

## 7.1 Policy evaluation with ripple effects

To quantify the expected change in earnings when capacity is changed, I sum the predicted outcomes from a counterfactual allocation and compare it to the outcomes in baseline allocation. Let $D_{is} \in P$ be the assignment of application $i$ in allocation $s$ and let $b$ be the baseline
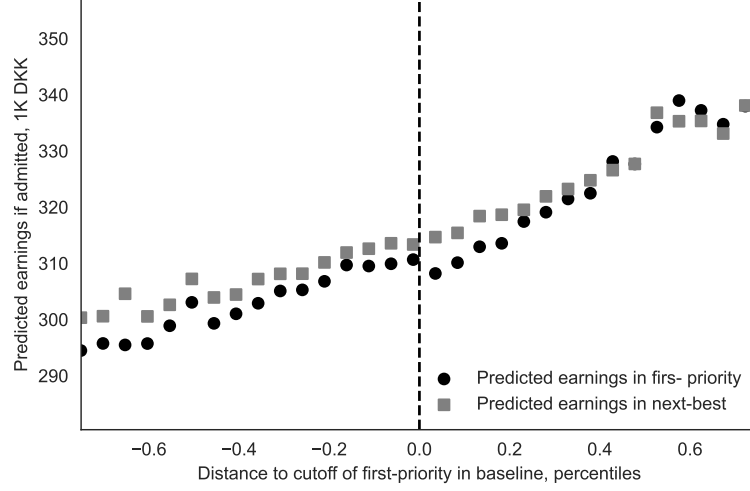
Figure 6: Predicted earnings as a function of eligibility score

**Note:** The figure shows the predicted outcomes of the first-priority program and the predicted outcome of the next-best program or the outside option. The sample is balanced such that the vertical distances between the black and the gray points represent the average predicted treatment effect of being admitted into the first-priority program conditional on location in the quota relative to the cutoff. Only applications in oversubscribed programs are included. Both Quota 1 (GPA-based) and Quota 2 (holistic assessment) are used for the graph.

allocation. Then the total effect of a simulation relative to the baseline is

$$dY_s^T = \sum_i \left[ \sum_p \hat{Y}_{ip} \times \mathbf{1}(D_{is} = p) - \hat{Y}_{ib} \right] = \sum_i \Delta_{is},$$

where $\Delta_{is}$ is the predicted individual "treatment effect" of going from the baseline allocation to allocation $s$.

In the main analysis, I investigate changes to a single program capacity at a time. This implies that I can distinguish between direct and ripple effects. In this case, $s$ identifies that manipulated program. Suppose that in simulation $s$ supply is changed in program $p$ and define $I_s$ as the set of applicants moving into or out of $p$, $I_s = \{i : D_{is} = p, D_{ib} \neq p \vee D_{is} \neq p, D_{ib} = p\}$. I then use the following identity for the changes in predicted earnings relative to the baseline:

$$\underbrace{dY_s^T}_{\text{Total effect}} = \underbrace{dY_s^D}_{\text{Direct effect}} + \underbrace{dY_s^R}_{\text{Ripple effect}}, \tag{15}$$

where

$$dY_s^D = \sum_{i \in I_s} \Delta_{is} \quad \text{and} \quad dY_s^R = \sum_{i \notin I_s} \Delta_{is}. \tag{16}$$

In this formulation, $dY^D$ captures the earnings changes occurring at the margin of the program with changed capacity. The rest of the changes in earnings are due to ripple effects captured by $dY^R$.
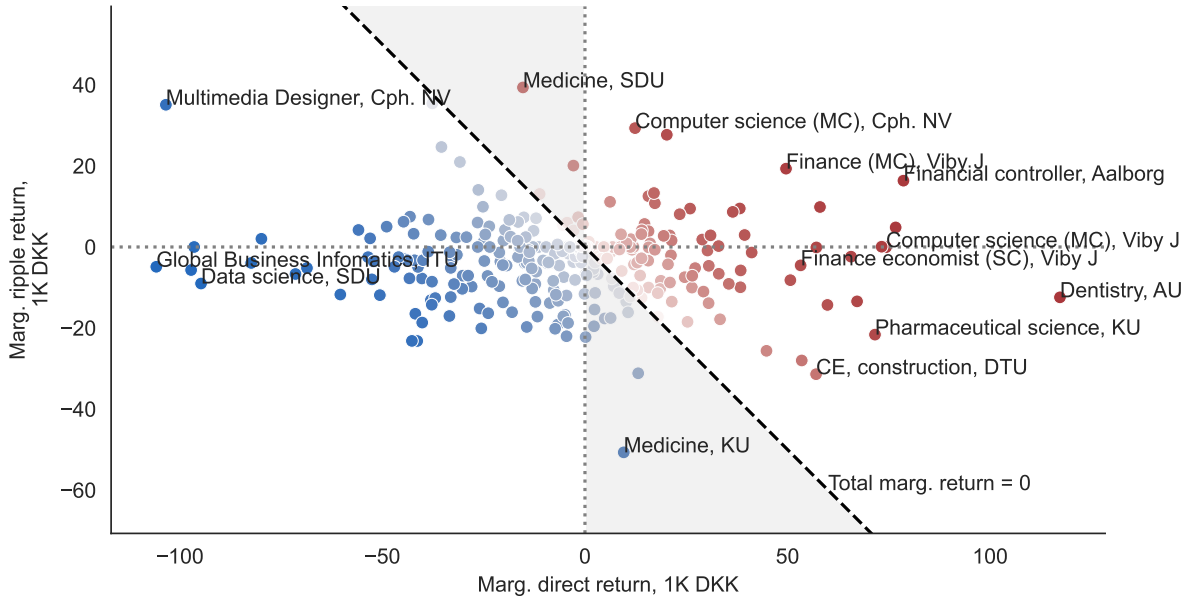
Figure 7: Direct returns to program changes and ripple effects, ≈ 30 slots added

**Note:** The figure shows the direct returns and ripple effects of adding one slot to a program in the GPA-based quota (Quota 1). On the horizontal axis is the direct effect, meaning the changes in predicted incomes for applicants moving into or out of the program. On the vertical axis is the ripple effect; the aggregate changes in predicted outcomes for applicants who change admission status but do not move in or out of the predicted program. The aggregate effect, the sum of direct and ripple effects, is illustrated by the colors. Red means positive aggregate effect, while blue means negative. The color coding is centered at zero, which coincides with the dashed line, which indicates where direct and ripple effects exactly offset each other. The shaded triangles show the programs where the direct effect has a different sign than the total effect. The predicted program× applicant-specific outcomes are based on the estimates of equation (10) which have been shrunk using the Empirical Bayes procedure described in section 7. For each program, I pick the simulation closest to an expansion of 30 slots and divide the return by the number of slots changed.

**Effects of expanding program capacity** I begin by illustrating the ripple effects for a subset of simulations, where programs are expanded with approximately 30 slots. Figure 7 shows a scatter plot of direct effects on the horizontal axis and ripple effects on the vertical axis. The dashed line indicates the points where direct effects are exactly offset by the ripple effects. The Euclidean distance from this line indicates the total return to the capacity change, which is shown using colors, where red indicates positive total returns, while blue indicates negative returns.

Figure 7 provides several takeaways. Firstly, there is a larger dispersion in the direct effects than in the ripple effects, which implies that the direct effect in general drives the total effect. Secondly, there is little correlation between the direct effect and the ripple effect. In other words, it is difficult to predict the externality from ripple effects measured in earnings, even when one has a qualified guess of the direct effect. Thirdly, there are cases where the ripple effects may cause the aggregate effect to have a different sign than the direct effect, even for similar programs. In figure 7 this is seen by the location of the medicine programs of the University of Copenhagen, (KU) and the University of Southern Denmark (SDU). While an expansion of the medicine program at KU is predicted to cause increases for applicants moving into this program, the applicants moving into the slots previously occupied by these marginal applicants
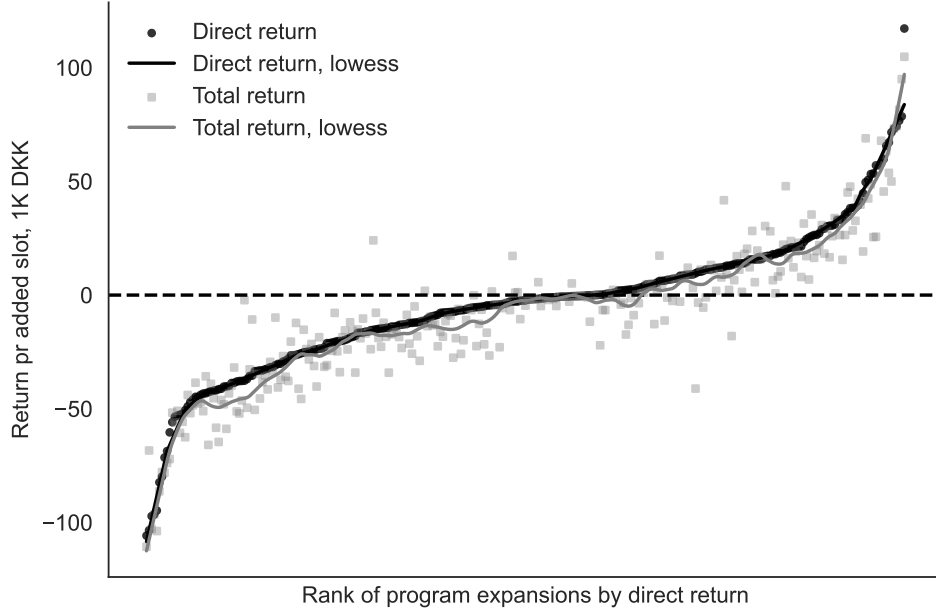
Figure 8: Direct and total returns to program changes, $\approx 30$ slots added

**Note:** The figure shows the programs ranked by their direct returns per slot added on the horizontal axis. On the vertical axis are the direct returns in black and the total returns in gray. The lines are fitted using Lowess smoothing. Refer to the figure 7 for details on computation. For each program, I pick the simulation closest to an expansion of 30 slots, and the effect divided by the number of slots changed.

(and their "descendants") have sufficiently lower predicted earnings to more than offset the direct return resulting in a loss in aggregate earnings. The opposite is the case for the medicine program at SDU, where the ripple effect offsets the direct loss for applicants moving into the medicine program. The comparison of the two programs illustrates the importance of ripple effects for cost-benefit analysis. Assuming that the costs of providing a slot are equal in the two medicine programs, a policy maker deciding where to allocate additional funds for expanding the medicine program would invest in the wrong program if she ignored ripple effects.[25] I return to the presence of sign-flipping between direct and total effects below. In appendix figure A.7 I report similar scatter plots for larger and smaller expansions and contractions.

To further illustrate the discrepancy between direct returns and total returns, I rank programs by the estimated direct returns and plot the direct returns in black and the total returns in gray in Figure 8. The distance between the black and gray points is the ripple effect. I fit the average returns conditional on rank using a Lowess smoother. On average the total effect is smaller than the direct effects, which indicates that the social returns to program expansions are lower than when ignoring smoothing. In other words, without taking into account ripple effects, the returns to program expansions will on average appear too attractive for policy-makers who seek to increase earnings. However, as is evident from figure 8 there is a lot of variation around the average conditional on rank.

---

[25]The heterogeneity of both the direct effects and ripple effects in the medicine programs are important for the literature on estimating returns to education where it is common to aggregate programs into fields, such as Kirkeboen et al. (2016); Humlum and Meyer (2020) and others.

**Characterizing the magnitude and sign of ripple effects** Figure 7 provided examples of the importance of ripple effects for policy analysis for a given capacity change. I now proceed to the more general analysis of ripple effects and how they depend on the size of capacity changes. To calculate the importance of ripple effects for policy analysis in terms of returns, I use the $R^2$ from a regression of the direct change, $dY^D$, on the total change, $dY^T$, which corresponds to the square of the correlation coefficient. For a given subset of simulations, the $R^2$ measures the share of variance explained by applicants moving in or out of the program where capacity is manipulated. To ease interpretation, I report $1 - R^2$, which is a direct measure of the ripple effect:

$$1 - R^2 = 1 - \frac{Cov(dY^D, dY^T)^2}{Var(dY^D)Var(dY^T)} \tag{17}$$

Further, I investigate the predictiveness of the direct returns on the ripple effects. I do this by estimating the slope of a regression of direct returns on the ripple effects:

$$\psi = \frac{Cov(dY^D, dY^R)}{Var(dY^D)}, \tag{18}$$

where I compute corresponding 95-percent heteroskedasticity robust confidence intervals.[26] I compute the two measures for subsets of simulations, where similar amounts of slots are changed, and plot the results in figure 9. Figure 9a shows that on average 9 percent of the variation in total returns is explained by ripple effects. I do not find that this share varies systematically with the number of slots that are subtracted or added. For small changes, the ripple effects appear to counteract direct effects, as shown in Figure 9b, where the coefficient is negative and significant for small changes. Around 3 percent of direct returns are lost on average due to ripple effects. However, as shown above in figure 8, the average effect masks considerable heterogeneity as ripple effects can go in both directions. This is also illustrated by the wide confidence intervals in figure 9b for the larger set of simulations.

**False classification of policy when ignoring ripple effects** As outlined above, ripple effects in returns do not consistently go in one or the other direction leaving little overall guidance for policy-makers. An important consideration is if and when the ripple effects may cause the aggregate returns to flip sign relative to the direct returns. Table 5 summarizes the ripple effects by the sign of the direct and total returns. The ripple effect is measured as the percentage difference between the total return and the direct return, $(dY^T - dY^D)/dY^D = dY^R/dY^D$ according to the identity defined in equation (15). From the off-diagonal part of this table, it is evident that for almost 9 percent of the simulations (4.65+4,27) the ripple effects cause the

---

[26]Note, that $R^2$ as defined in equation (17) is mechanically linked to $\psi$ through the following identity:

$$\psi = \sqrt{R^2 \frac{Var(dY^T)}{Var(dY^D)}} - 1.$$

The two measures thus describe the same features of the joint distribution of direct returns, ripple effects, and total returns.
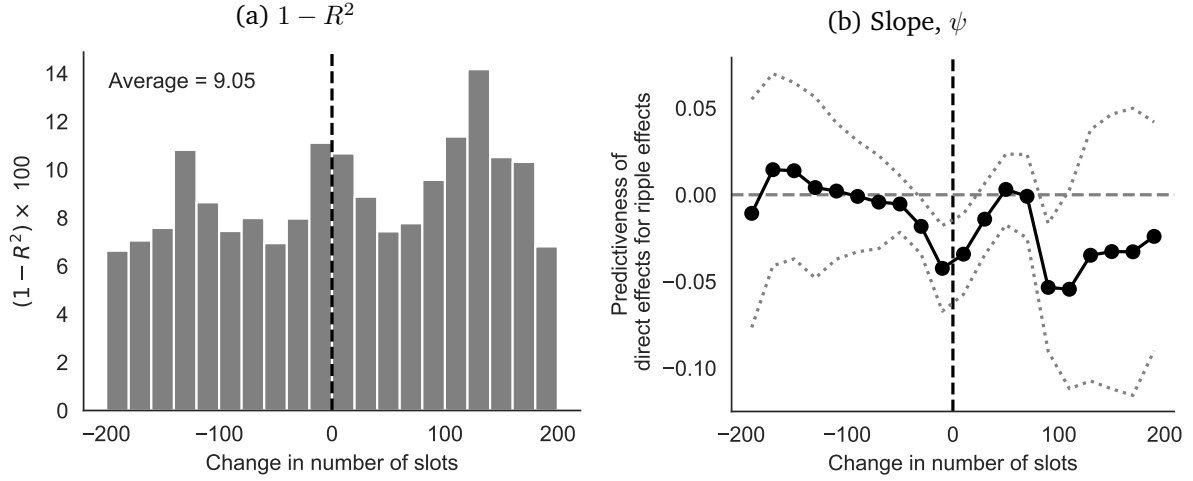
(a) $1 - R^2$      (b) Slope, $\psi$

Figure 9: Predictiveness of direct effects

**Note:** Figure 9a shows $(1 - R^2)$ from a regression of direct returns on total returns for each bin of simulations as described in equation (17). This is a measure of the variance in total returns not accounted for by the direct returns. Figure 9b presents estimates of the slope of a regression of direct returns on ripple effects, as described in equation (18). Each point represents an estimate within simulations where a certain number of slots are added. The dashed lines show the 95 percent confidence interval based on robust standard errors.

sign of the return to flip. Appendix figure A.8 shows that the share of simulations where the sign flip is over 12 percent for small changes.

For simulations where the sign does not flip, the median ripple effect in returns is 3 percent and 5 percent for negative and positive returns respectively. In the off-diagonal, the ripple effects are necessarily larger around 200 percent. As evidenced by the difference between the median and mean, the distributions are highly skewed, with some simulations creating very large ripple effects. Appendix figure A.9 shows the distributions underlying Table 5.[27]

## 7.2 Policy evaluation with simultaneous supply changes in multiple programs

The results above are based on simulations where the capacity of a single program is changed. It thus captures externalities for a policy-maker who considers changing capacity in a single program. Policy-makers in systems where supply is governed centrally often consider multiple programs at a time. For example, an increase of slots in STEM or as in the following example medicine. However, programs within fields are likely to be each other's substitutes as seen in the example in section 3. As applicants can only be admitted in one place, simultaneous changes in capacity likely alter the composition of incoming applicants to each program and thus returns.

To illustrate the complications that deciding over multiple programs poses for decision-makers, I simulate a joint supply change in the two largest medicine programs in Denmark, at the University of Copenhagen and Aarhus University. I perform three simulations. First, I expand the capacity by 200 slots at each program holding the capacity in the other fixed. I then add 200 slots to both programs simultaneously. I present the resulting flows relative to the

---

[27]The mean ripple effect of -.36.25 does not coincide with the reported mean in figure 9b, which reports simulations where no more than 200 slots are added or subtracted.

Table 5: Classification of capacity changes according to direct and total returns

| Direct return | Total return | Negative | Positive | All |
|---|---|---|---|---|
| Negative | N | 4,686 | 476 | 5,162 |
| | Share, percent | 45.73 | 4.65 | 50.38 |
| | Ripple effect, Median | 0.03 | -2.06 | -0.00 |
| | Ripple effect, Mean | 1.96 | -5.26 | 1.30 |
| | Ripple effect, Std. | 63.14 | 12.54 | 60.31 |
| Positive | N | 438 | 4,646 | 5,084 |
| | Share, percent | 4.27 | 45.34 | 49.62 |
| | Ripple effect, Median | -2.11 | 0.05 | 0.02 |
| | Ripple effect, Mean | -920.92 | 5.21 | -74.57 |
| | Ripple effect, Std. | 19,112.53 | 286.64 | 5,616.72 |
| All | N | 5,124 | 5,122 | 10,246 |
| | Share, percent | 50.01 | 49.99 | 100.00 |
| | Ripple effect, Median | -0.00 | 0.01 | 0.00 |
| | Ripple effect, Mean | -76.92 | 4.24 | -36.35 |
| | Ripple effect, Std. | 5,588.38 | 273.04 | 3,956.50 |

**Note:** This table summarizes the ripple effects by the sign of the direct and total returns. The ripple effect is measured as the percentage difference between the total return and the direct return, $(dY^T - dY^D)/dY^D = dY^R/dY^D$ according to the identity defined in equation (16). Simulations with direct returns of zero are excluded from the table. Histograms of the distribution are in appendix figure A.9.

baseline allocation graphically in figure 10. The number of applicants affected is shown in the top panel of table 6.

As seen in figures 10a and 10b, the largest inflows come from other (non-medicine) programs, but there is a large inflow from other medicine programs as well. There is also a large flow on the extensive margin, but these are mostly flows into non-medicine programs, which is possible due to the ripple effects. However, the Aarhus program attracts more applicants on the extensive margin as seen in figure 10b relative to the Copenhagen program. When 200 slots are added to both programs simultaneously, the largest single flowis from non-assignment into non-medicine programs. Again these flows come due to slots left vacant by applicants switching to medicine. However, these applicants are now split into the two medicine programs. Further, now applicants previously admitted to the Copenhagen program go to Aarhus and vice-versa. This indicates that the compiler population for the single-program expansions in figures 10a and 10b are likely to differ from the composition in figure 10c.[28]

To assess whether heterogeneity of the complier population matters for policy evaluation, I calculate the average gain for incoming applicants in the three simulations. The results are presented in table 6. Column 1 shows the effect of an increase in capacity in the GPA-based quota (Quota 1) in the medicine program at the University of Copenhagen. The average loss for the 212 incoming applicants is 29 thousand DKK (4.3 thousand USD). The increase in Copenhagen implies that 18 applicants gain entry to the Aarhus program, with an average gain of

---

[28]Note, that cycles occur. When the Copenhagen program is expanded, in figure 10a, some applicants *leave* the program for other non-medicine programs. This is due to the quota system with multiple tie-breakers as discussed in section 2.
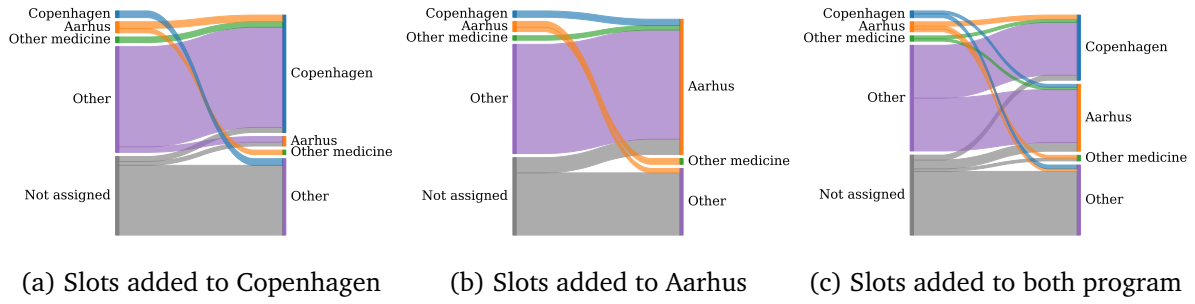
(a) Slots added to Copenhagen  (b) Slots added to Aarhus  (c) Slots added to both program

Figure 10: Flow of applicants when expanding medicine programs

**Note:** The figures the flows between medicine programs and other programs when medicine programs are expanded. Flows of less than five applicants are not shown due to anonymity requirements imposed by Statistics Denmark. Figure 10a shows the flows when the medicine program at the University of Copenhagen is expanded by 200 slots. Figure 10b shows the corresponding flows when 200 slots are added to the program at Aarhus University. Figure 10c shows the resulting flows when 100 slots are added to each of the two programs simultaneously. Sankey diagrams are produced using a modified version of code developed by Golob (2018).

Table 6: Simultaneous changes in the capacity of Medicine programs

| Simulation | (1) | (2) | (3) |
| Change to: | Copenhagen | Aarhus | Both |
|---|---:|---:|---:|
| N entrants to Copenhagen | 212 | 13 | 223 |
| N entrants to Aarhus | 18 | 213 | 229 |
| N affected | 1,358 | 1,040 | 1,952 |
| Avg. gain, entrants to Copenhagen | -28.88 | -24.39 | -29.21 |
| Avg. gain, entrants to Aarhus | 17.16 | 2.03 | 4.54 |
| Avg. gain, all affected | -5.33 | 7.71 | -0.07 |
| Aggregate gain | -4,979.26 | 4,749.98 | -112.54 |

**Note:** The table displays the sizes of flows and gains from three experiments. Columns 1 and 2 show the result of adding 200 slots to the Medicine program at the University of Copenhagen and Aarhus University respectively. Column 3 shows the results of adding 100 slots to each Medicine program at the same time. In all simulations, slots are added to the GPA-based quota. The simulations also form the basis of Figure 10 The conversion rate from USD to DKK used in this paper is 6.72.

17 thousand DKK (2.6 thousand USD). In column 2, the slots are instead added to the Aarhus program. Now, the average gain for the incoming applicants to the Aarhus program is merely 2 thousand DKK (300 USD), which is consistent with diminishing returns for applicants further from the cutoff shown in figure 6. In column 3 the two medicine programs are increased by 200 slots each. The average gain for admitted applicants to the Copenhagen program is now more negative while the average for entrants into the Aarhus program is larger and positive. This is a clear indication that the complier compositions change due to ripple effects.

The change in the marginal return to the Aarhus program in the three experiments shows that the LATE cannot be interpreted as a policy-invariant parameter. While this lack of policy-invariance is well recognized in the literature, the simulation exercises show the practical limitations of local treatment effect estimates for informing policies of sector-wide changes within higher education.

## 7.3 Using non-parametric LATEs

As outlined in Section 2, the relevant treatment effects can in principle be estimated in a fuzzy regression discontinuity design using eligibility score as a running variable. Relative to the

parametric approach above, this approach has the advantage of not imposing restrictions on the heterogeneity of potential outcomes. Further, a stratified design retrieves an estimate which is an integral over the relevant alternatives in the distribution of compliers, which avoids the conditioning issues highlighted by Kirkeboen et al. (2016). Finally, one only needs aggregate inflows and outflows to quantify the ripple effect. The apparent attractiveness of this approach is however severely limited by the locality of the LATE estimands, low precision, and lack of empirical support. For the sake of completion, this section performs the analysis of ripple effects using LATEs and flows.

In a fuzzy RD design, I define the endogenous treatment as admission. For each program with sufficient empirical support around the cutoff in Quota 1 in my estimation sample, I compute program-level LATEs using the following specification:

$$Y_{ip} = \delta_p D_i + f(e_{ip}) + \epsilon_{ip}, \tag{19}$$

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 flexibly specified spline with a knot at zero. $D_i$ is a dummy for admission and it is instrumented by crossing the cutoff in the running variable. In the presence of heterogeneous treatment effects, the program-specific LATE represents the earnings gain for applicants at the threshold who do not get admitted into a more preferred program. I estimate (19) for each program in my sample which has a cutoff using the *rdrobust* package in Stata (Calonico et al., 2017), with heteroskedasticity robust standard errors. Assuming that there are no applicants at the cutoffs who are part of circles discussed in section 2.1.1, crossing a cutoff leads to weakly increasing the probability of admittance (ie. there are no defiers for the instrument of crossing thresholds). Thus, following Andrews and Armstrong (2017); Angrist and Kolesár (2023), I exclude programs with an estimated negative first-stage.

As outlined in section 2, the standard fuzzy RD design retrieves the policy-relevant treatment effect for marginal capacity increases. For marginal capacity decreases, however, the flow is better captured by the reverse design, where crossing the threshold from above, pushes compliers into their next-best program (including the option of no admission). This design has previously been employed by Humlum and Meyer (2020). To capture heterogeneity in outside options, I estimate an additional set of program-specific LATEs, for the application where no next-best program listed. In other words, moving into non-admission has a different effect for each program margin.[29]

**Estimates of local average treatment effects**  Appendix figure A.10 shows that predetermined covariates are balancing around the threshold.[30] I begin by showing the aggregate fuzzy RDD results where programs are pooled and centered at cutoffs. The first stage, displayed in

---

[29]For both the original and the reverse design I exploit data from multiple admission rounds. This implies that the set of compliers is going to be located at various places on the GPA scale depending on each year's cutoff. Thus, implicit in the design, the resulting LATE is an average on the range, where program cutoffs are located in the data.

[30]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.

(a) Aggregate first stage       (b) Aggregate reduced form

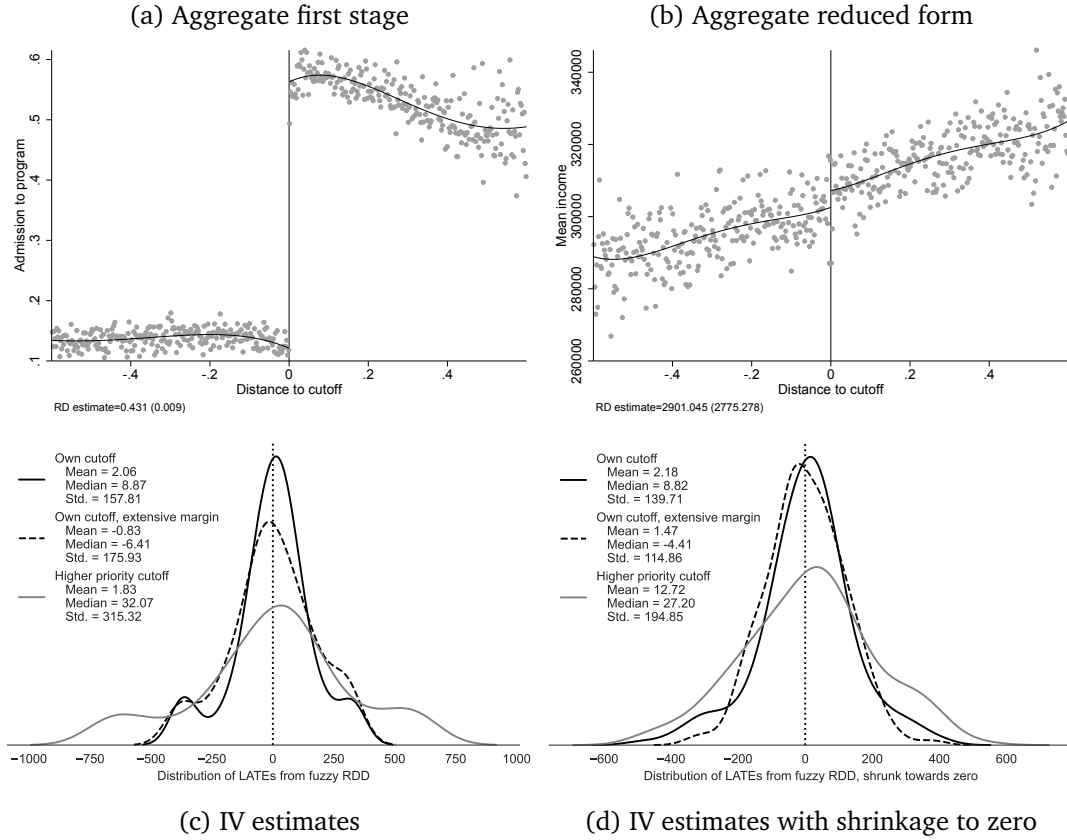(c) IV estimates       (d) IV estimates with shrinkage to zero

Figure 11: Regression discontinuity plots and distributions of LATEs

**Note:** Figures 11a and 11b show the first stage and reduced form for an aggregate fuzzy RDD specification. 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 across programs and years. Figures 11c and 11d show the distribution of LATEs as estimated in the program-level fuzzy RD design outlined in equation (19). The solid line shows the standard design, where the running variable is Quota 1-ranking within the program. The dashed gray line is for the "reverse design, where the running variable is the reversed priority in Quota 1 for the higher prioritized program. The densities are computed using kernel-density estimation.

figure 11a, 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, illustrated in figure 11b is 3,000 DKK equivalent to 450 USD in 2015, and is not graphically convincing. Scaling the reduced form by the first stage produces a LATE of 6,730 DKK (app. 1000 USD) in yearly income gain for compliers.

Figure 11c presents the distribution of estimated program-level LATEs. The median LATE using the standard design is 8.87 thousand DKK (1,300 USD) with a large dispersion in the distribution.[31] The median effect of getting into a less preferred program is a gain of 32 thousand DKK (4,700 USD). The dispersion of effects is large, with some programs yielding large gains or

---

[31]The very large LATEs in terms of absolute value partly reflect measurement noise. Thus I consider a LATE of 1 million DKK implausibly large.

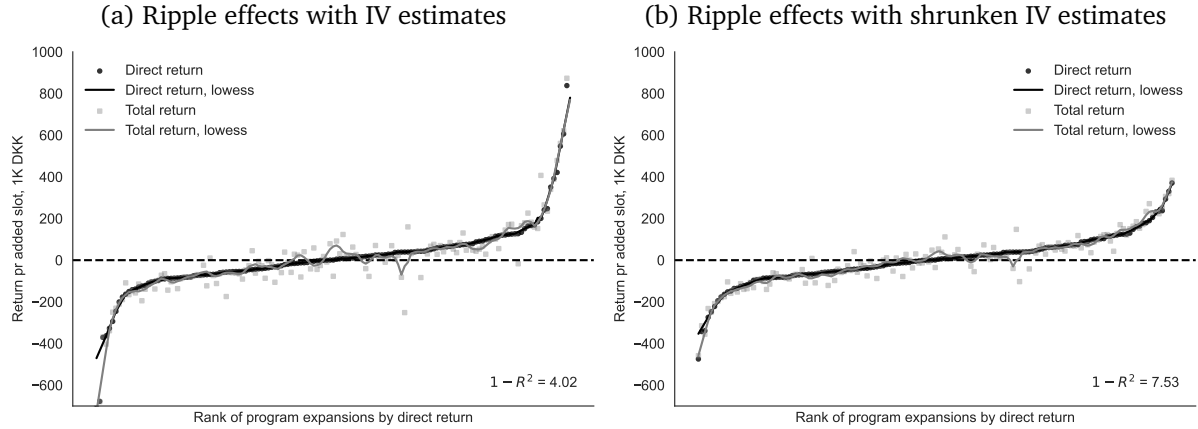| (a) Ripple effects with IV estimates | (b) Ripple effects with shrunken IV estimates |

Figure 12: Ripple effects with IV estimates of returns

**Note:** The figures show the programs ranked by direct returns on the horizontal axis. On the vertical axis are the returns from ripple effects in black and the total returns in gray. The lines are fitted using Lowess smoothing. Refer to the figure 7 for details on computation. For each program, I pick the simulation closest to an expansion of 30 slots and divide by the number of slots changed. In figure 12a I use returns estimated by IV, and in figure 12b I use estimates that are shrunk towards zero using the procedure outlined by Armstrong et al. (2020). I multiply $(1 - R^2)$ by 100.

losses for compliers. To assess the importance of the lack of precision, I also shrink the estimates towards zero using the procedure outlined by Armstrong et al. (2020) and present the results in figure 11d. As expected, the magnitudes are lower, especially for the design with the higher priority cutoff.

**Ripple effects based on LATEs**   I merge the estimated LATEs onto the simulations and again calculate direct returns, ripple effects, and total returns. An important caveat is that I can only estimate LATEs where I have sufficient support around a cutoff. Appendix figure A.11 shows that I do not have support to estimate a large share of the required LATEs, which implies that in the median simulation, around 40 percent of the applicants who move do not have an associated LATE. The gains are in this case set to zero, and this underlines the problems of the non-parametric approach.

Figure 12 reproduces figure 8 using the LATEs rather than the potential outcomes from the control function method. Figure 12a shows the results from using the raw IV estimates. The scale is much larger than using the control function estimates, which is largely due to the impression of the non-parametric RDD in small samples. However, the overall magnitude of ripple effects is similar to the finding using my main control function approach. In this case, 4 percent of the variation in total returns is due to ripple effects. However, this relatively lower share is likely due to two factors; impression and lack of support. While I cannot improve on the latter, the former can be mitigated by applying an Empirical Bayes procedure, where I shrink estimates toward zero. As evidenced by figure 12b, using shrunken estimates, ripple effects now explain 7.5 percent of the variation, which is close to the share that I find using the control function procedure. Once again, it must be stressed that the IV estimates are imprecise which is especially problematic in the context of ripple effects where one needs a potentially large set of IV estimates as shown in figure 4b.

# 8 Conclusion

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 outcomes for these individuals. Secondly, ripple effects introduce a divergence between the perceived benefits from the perspective of a program's admission office and the overall social implications.

I approached the implications of ripple effects for policy evaluation in several ways. Regardless of whether I use reported or estimated preferences, or whether I use structured potential outcomes or non-parametric LATEs, I consistently arrive at similar qualitative conclusions. On average, 10 percent of the relevant variation in outcomes is lost when ignoring ripple effects and policy recommendations may change dramatically once ripple effects are taken into account.

This paper underscores the significance of considering the equilibrium effects of changing the supply of education. However, 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 important dimension to explore. Lastly, the returns are measured early in the career. As different fields may have very different earnings profiles, both direct returns and ripple effects may change when investigating lifetime earnings.

Although the quantitative results are specific to the Danish context of centralized assignment, the qualitative finding that ripple effects matter extends beyond this particular context and the centralized nature of the market. Educational slots are inherently indivisible, and applicants typically have unit demand, which creates ripple effects 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 market clearing determines the relevant population.

# References

Abdulkadiroğlu, A., Agarwal, N., Pathak, P. A. 2017a. The welfare effects of coordinated assignment: Evidence from the new york city high school match. American Economic Review, 107, 3635–89.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., Pathak, P. A. 2017b. Research design meets market design: Using centralized assignment for impact evaluation. Econometrica, 85, 1373–1432.

Abdulkadiroglu, A., Angrist, J. D., Narita, Y., Pathak, P. A. 2022. Breaking ties: Regression discontinuity design meets market design. Econometrica.

Abdulkadiroğlu, A., Sönmez, T. 2003. School choice: A mechanism design approach. American economic review, 93, 729–747.

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J., Walters, C. R. 2020. Do parents value school effectiveness? American Economic Review, 110, 1502–39.

Agarwal, N. 2015. An empirical model of the medical match. American Economic Review, 105, 1939–78.

Agarwal, N. 2017. Policy analysis in matching markets. American Economic Review, 107, 246–50.

Agarwal, N., Hodgson, C., Somaini, P. 2020. Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys.Technical report, National Bureau of Economic Research.

Agarwal, N., Somaini, P. 2018. Demand analysis using strategic reports: An application to a school choice mechanism. Econometrica, 86, 391–444.

Altonji, J. G., Arcidiacono, P., Maurel, A. 2016. The analysis of field choice in college and graduate school: Determinants and wage effects. In Handbook of the Economics of Education, 5, Elsevier, 305–396.

Andersen, M. B., Hørlück, S., Sørensen, C. L. 2020. Effekt af optag på videregående uddannelser for marginale studerende.Technical report, Working Paper.

Andrews, I., Armstrong, T. B. 2017. Unbiased instrumental variables estimation under known first-stage sign. Quantitative Economics, 8, 479–503.

Andrews, R. J., Li, J., Lovenheim, M. F. 2016. Quantile treatment effects of college quality on earnings. Journal of Human Resources, 51, 200–238.

Angrist, J., Kolesár, M. 2023. One instrument to rule them all: The bias and coverage of just-id iv. Journal of Econometrics.

Arcidiacono, P., Aucejo, E. M., Hotz, V. J. 2016. University differences in the graduation of minorities in stem fields: Evidence from california. American Economic Review, 106, 525–62.

Armstrong, T. B., Kolesár, M., Plagborg-Møller, M. 2020. Robust empirical bayes confidence intervals. arXiv preprint arXiv:2004.03448.

Azevedo, E. M., Leshno, J. D. 2016. A supply and demand framework for two-sided matching markets. Journal of Political Economy, 124, 1235–1268.

Bourguignon, F., Fournier, M., Gurgand, M. 2007. Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. Journal of Economic surveys, 21, 174–205.

Bucarey, A. 2018. Who pays for free college? crowding out on campus. Job market paper.

Calonico, S., Cattaneo, M. D., Farrell, M. H., Titiunik, R. 2017. rdrobust: Software for regression-discontinuity designs. The Stata Journal, 17, 372–404.

Calsamiglia, C., Haeringer, G., Klijn, F. 2010. Constrained school choice: An experimental study. American Economic Review, 100, 1860–74.

Chrisander, E., Bjerre-Nielsen, A. 2023. Why do students lie and should we worry? an analysis of non-truthful reporting. arXiv preprint arXiv:2302.13718.

Dahl, G. B. 2002. Mobility and the return to education: Testing a roy model with multiple markets. Econometrica, 70, 2367–2420.

Dale, S. B., Krueger, A. B. 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. The Quarterly Journal of Economics, 117, 1491–1527.

Daly, M., Jensen, M. F., Le Maire, D. 2020. University admission and preferred field of study.Technical report, Working Paper.

DST. 2023. Nyt fra Danmarks Statistik - nr. 286 - 2023. Accessed on 2023-11-30.

Dubin, J. A., McFadden, D. L. 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica: Journal of the Econometric Society, 345–362.

Fack, G., Grenet, J., He, Y. 2019. Beyond truth-telling: Preference estimation with centralized school choice and college admissions. American Economic Review, 109, 1486–1529.

Feller, A., Grindal, T., Miratrix, L., Page, L. C. et al. 2016. Compared to what? variation in the impacts of early childhood education by alternative care type. Annals of Applied Statistics, 10, 1245–1285.

Gale, D., Shapley, L. S. 1962. College admissions and the stability of marriage. The American Mathematical Monthly, 69, 9–15.

Gandil, M., Leuven, E. 2022. College admission as a screening and sorting device.

Golob, A. 2018. pySankey. https://github.com/anazalea/pySankey.

Hastings, J. S., Neilson, C. A., Zimmerman, S. D. 2013. Are some degrees worth more than others? evidence from college admission cutoffs in chile.Technical report, National Bureau of Economic Research.

Heckman, J. J., Urzua, S., Vytlacil, E. 2008. Instrumental variables in models with multiple outcomes: The general unordered case. Annales d'Economie et de Statistique, 151–174.

Heinesen, E., Hvid, C. 2019. Returns to field of study in the medium term – instrumental variables estimates based on admission thresholds.Technical report, Working Paper.

Heinesen, E., Hvid, C., Kirkebøen, L. J., Leuven, E., Mogstad, M. 2022. Instrumental variables with unordered treatments: Theory and evidence from returns to fields of study.Technical report, National Bureau of Economic Research.

Humlum, A., Meyer, B. B. 2020. Artificial intelligence and college majors.Technical report, Working Paper.

Imbens, G. W., Angrist, J. D. 1994. Identification and estimation of local average treatment effects. Econometrica, 62, 467–475.

Kapor, A. J., Neilson, C. A., Zimmerman, S. D. 2020. Heterogeneous beliefs and school choice mechanisms. American Economic Review, 110, 1274–1315.

Kapor, A., Karnani, M., Neilson, C. 2022. Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms.Technical report, National Bureau of Economic Research.

Kirkeboen, L. J., Leuven, E., Mogstad, M. 2016. Field of study, earnings, and self-selection. The Quarterly Journal of Economics, 131, 1057–1111.

Kline, P., Walters, C. R. 2016. Evaluating public programs with close substitutes: The case of head start. The Quarterly Journal of Economics, 131, 1795–1848.

Larroucau, T., Rios, I. 2022. Dynamic college admissions.Technical report.

Manning, A., Petrongolo, B. 2017. How local are labor markets?  evidence from a spatial job search model. American Economic Review, 107, 2877–2907.

Morris, C. N. 1983. Parametric empirical bayes inference: theory and applications. Journal of the American statistical Association, 78, 47–55.

Mountjoy, J., Hickman, B. 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.

Pathak, P. A., Shi, P. 2021. How well do structural demand models work?  counterfactual predictions in school choice. Journal of Econometrics, 222, 161–195.

Robbins, H. E. 1992. An empirical bayes approach to statistics. In Breakthroughs in statistics, Springer, 388–394.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies.. Journal of educational Psychology, 66, p. 688.

Tanaka, M., Narita, Y., Moriguchi, C. 2020. Meritocracy and its discontent: Long-run effects of repeated school admission reforms.Technical report, Research Institute of Economy, Trade and Industry (RIETI).

UFM. 2018. Frafald og studieskift på de videregående uddannelser.Technical report, Ministeriet for Forskning, Innovation og Videregående Uddannelser.

Zimmerman, S. D. 2014. The returns to college admission for academically marginal students. Journal of Labor Economics, 32, 711–754.
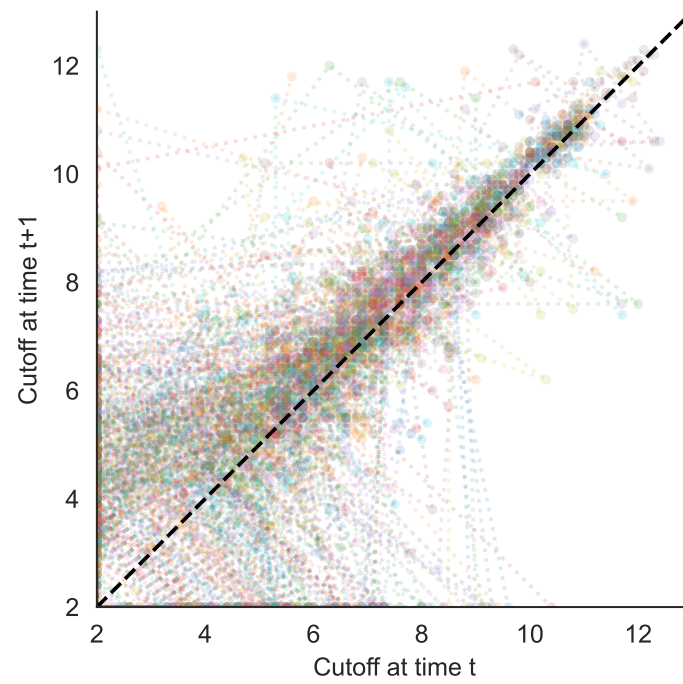
# A   Additional figures and tables



Figure A.1: Cutoff change from year to year

**Note:** The figure plots Quota 1 cutoffs as a function of cutoff from last year. Mass around the 45 degree line indicates the stability of cutoffs from year to year.
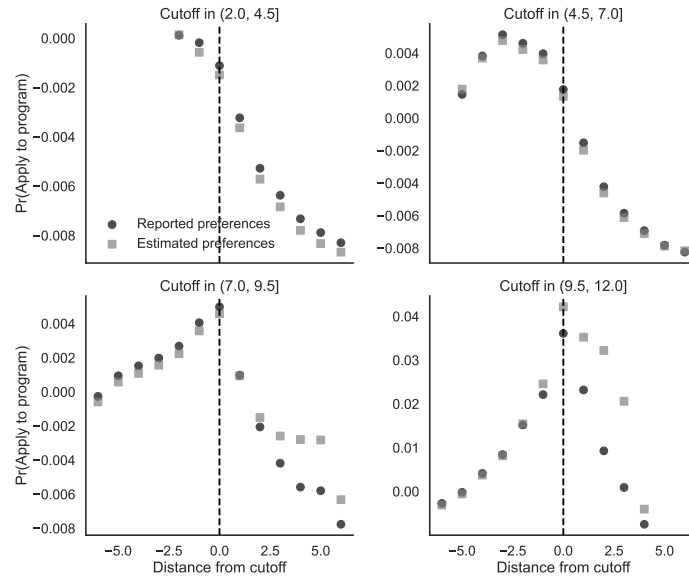
Figure A.2: Probability of applying as function of distance to cutoff

**Note:** The figure shows the probability of applying to a program as a 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.
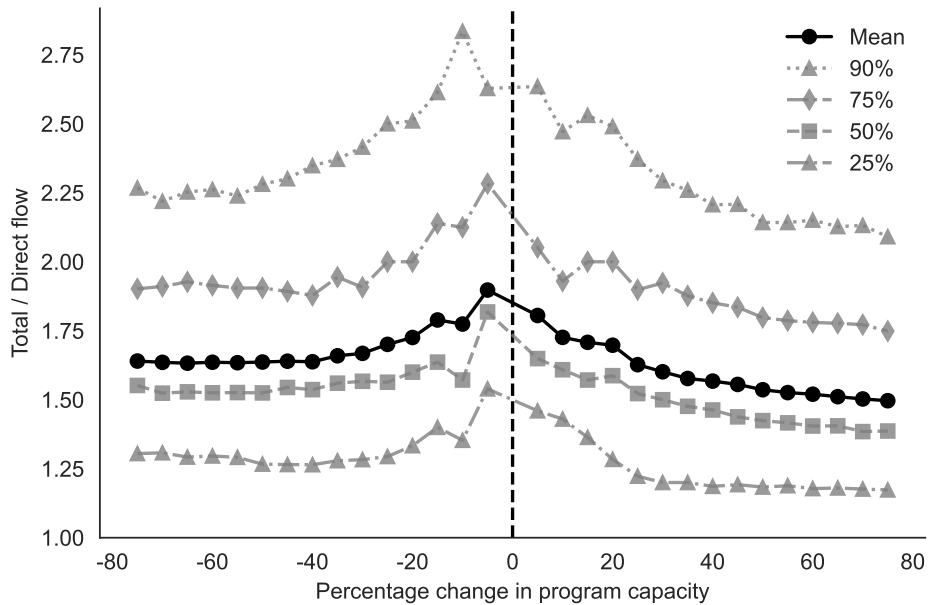


Figure A.3: Total flow relative to direct flow, reported preferences

**Note:** 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.
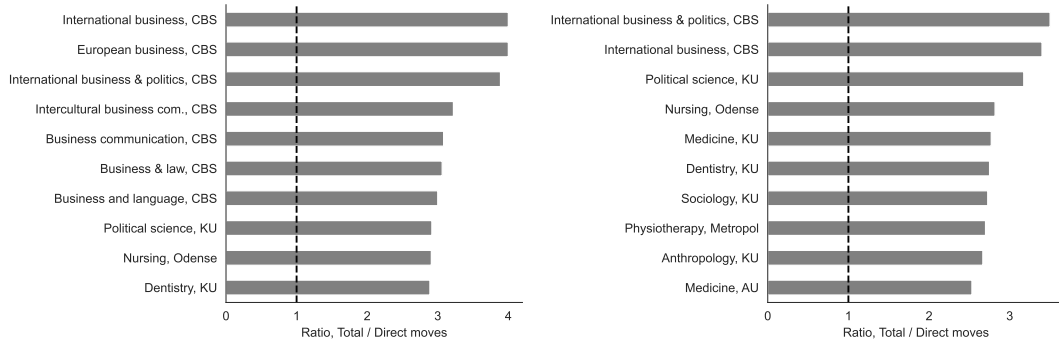
(a) Reported pref: 10% reduction in capacity    (b) Reported pref: 10% increase in capacity

Figure A.4: Programs with the largest ripple effects measured by flows of applicants, reported preferences

**Note:** Figures 5a)-b) 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 are sufficiently large to fulfill the 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 A.4 contains equivalent results using estimated preferences, rather than estimated preferences.



(a) Centrality with rep. pref.    (b) Cutoffs with rep. pref.

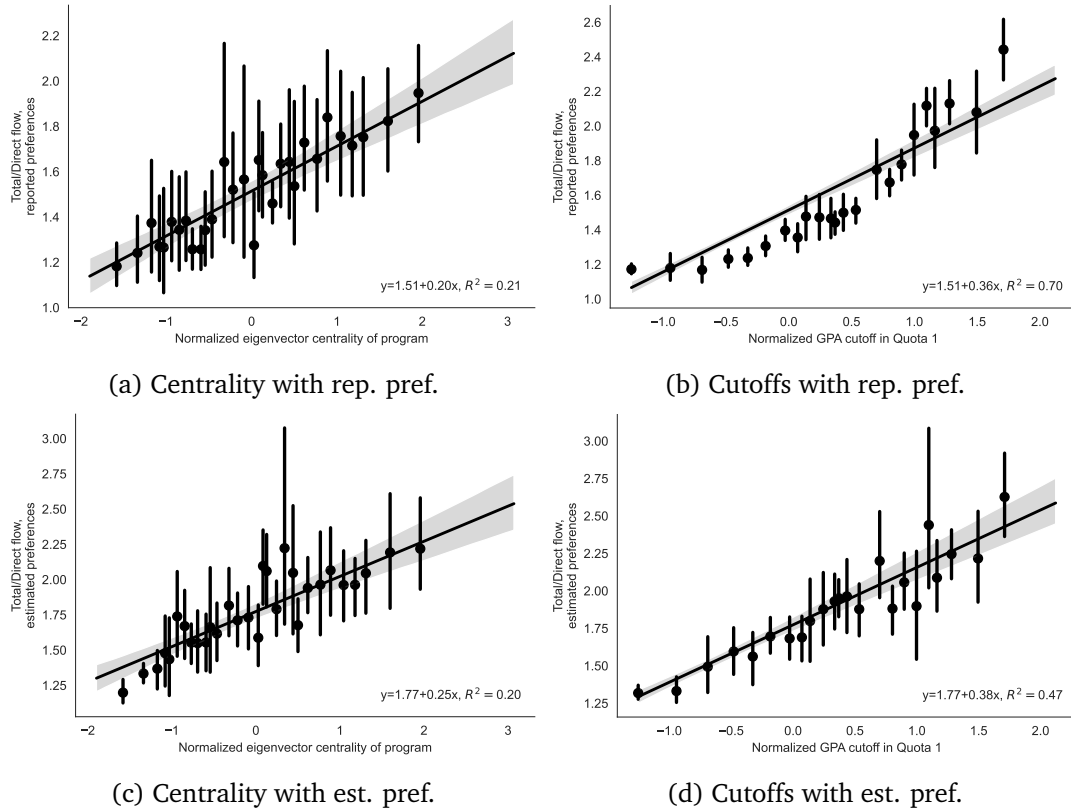(c) Centrality with est. pref.    (d) Cutoffs with est. pref.

Figure A.5: Predicting ripple effects in flows

**Note:** 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) show the results based on simulation with reported preferences. Panel c) and d) show the results based on simulation with estimated preferences. Centrality is measured as the Eigen-vector centrality of a weighted graph constructed from the applications for the round of 2016. An edge weight 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.
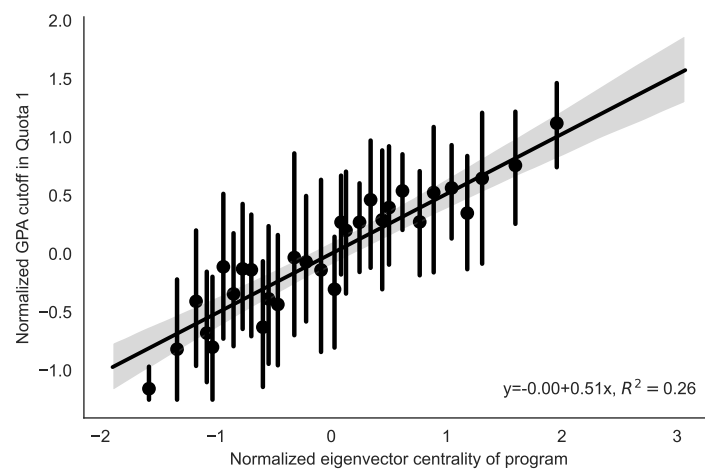
43

Figure A.6: Predicting ripple effects in ratio of absolute gains

**Note:** The figure plots the linear relationship between GPA cutoffs and weighted eigenvector centrality in the applications for the 2016 cohort.
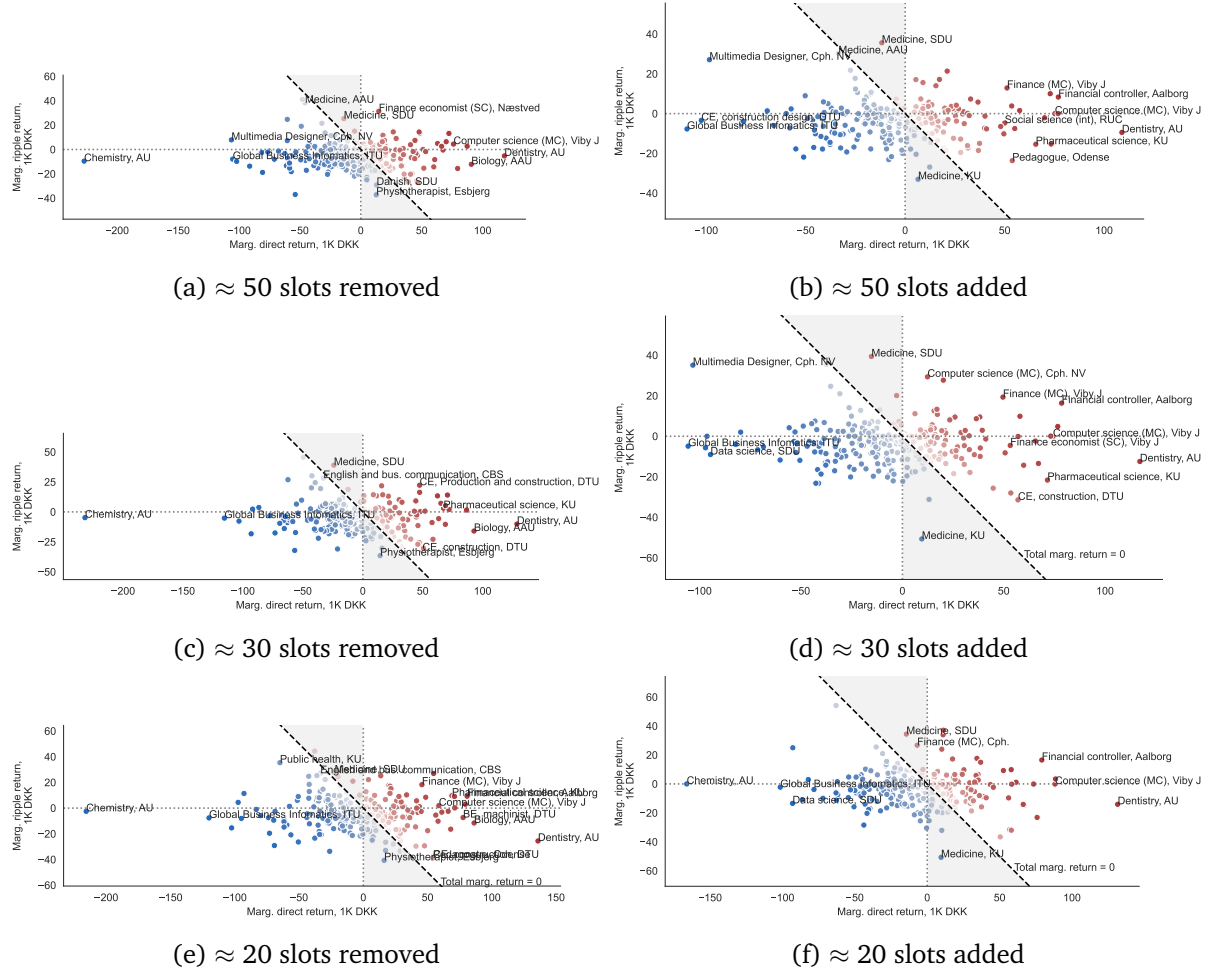
(a) ≈ 50 slots removed

(b) ≈ 50 slots added

(c) ≈ 30 slots removed

(d) ≈ 30 slots added

(e) ≈ 20 slots removed

(f) ≈ 20 slots added

Figure A.7: Direct returns to program changes and ripple effects

**Note:** The figures show the direct returns and ripple effects of adding one slot to a program in the GPA-based quota (Quota 1). On the horizontal axis is the direct effect, meaning the changes in predicted incomes for applicants moving into or out of the program. On the vertical axis is the higher-order effect; the aggregate changes in predicted outcomes for applicants who change admission status but do not move in or out of the predicted program. The aggregate effect, the sum of direct and ripple effects, is illustrated by the colors. Red means positive aggregate effect, while blue means negative. The color coding is centered at zero, which coincides with the dashed line, which indicates where direct and ripple effects exactly offset each other. The shaded triangles show the programs where the direct effect has a different sign than the total effect. The predicted program× applicant-specific outcomes are based on the estimates of equation (10) which have been shrunk using the Empirical Bayes procedure described in section 7. For each program, I pick the simulation closest to an expansion of 30 slots and divide the return by the number of slots changed.
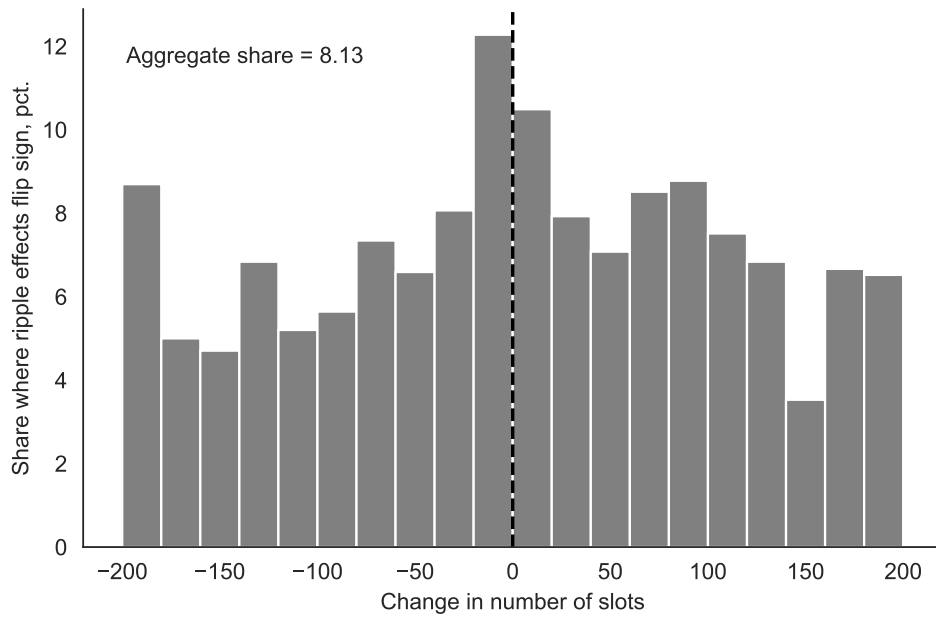
45

Figure A.8: Share of simulations where sign of total return $\neq$ direct return

**Note:** The figure shows the share of simulations, where the total return is of a different sign than the direct return. Simulations are binned by the number of changes in slots.
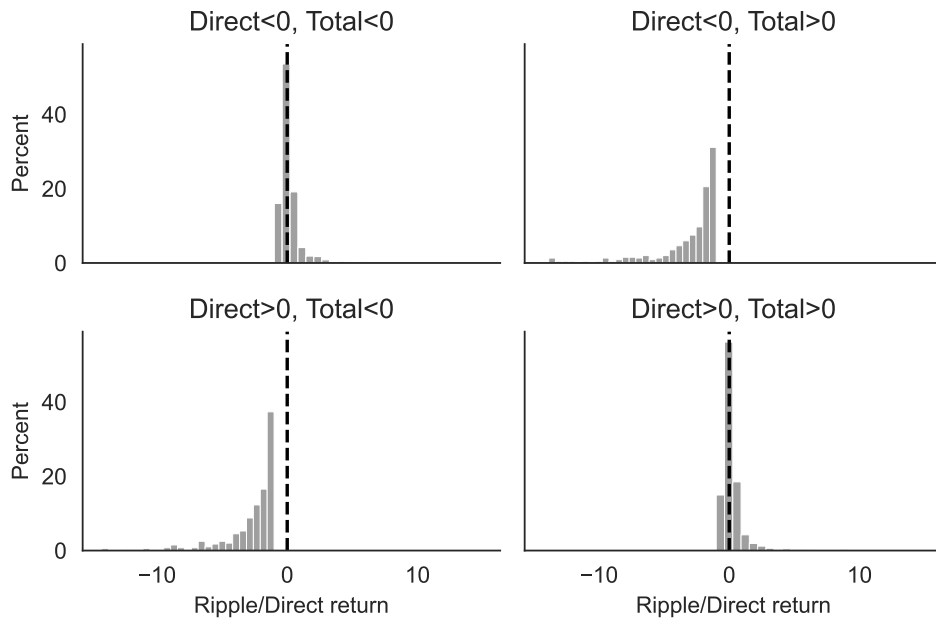


Figure A.9: Distributions of ripple effects, separated by sign of total return and direct return

**Note:** The figure shows the distribution of ripple effects divided by the direct effect conditional on the signs of the direct effect and the total return. See text below table 5 for further details.
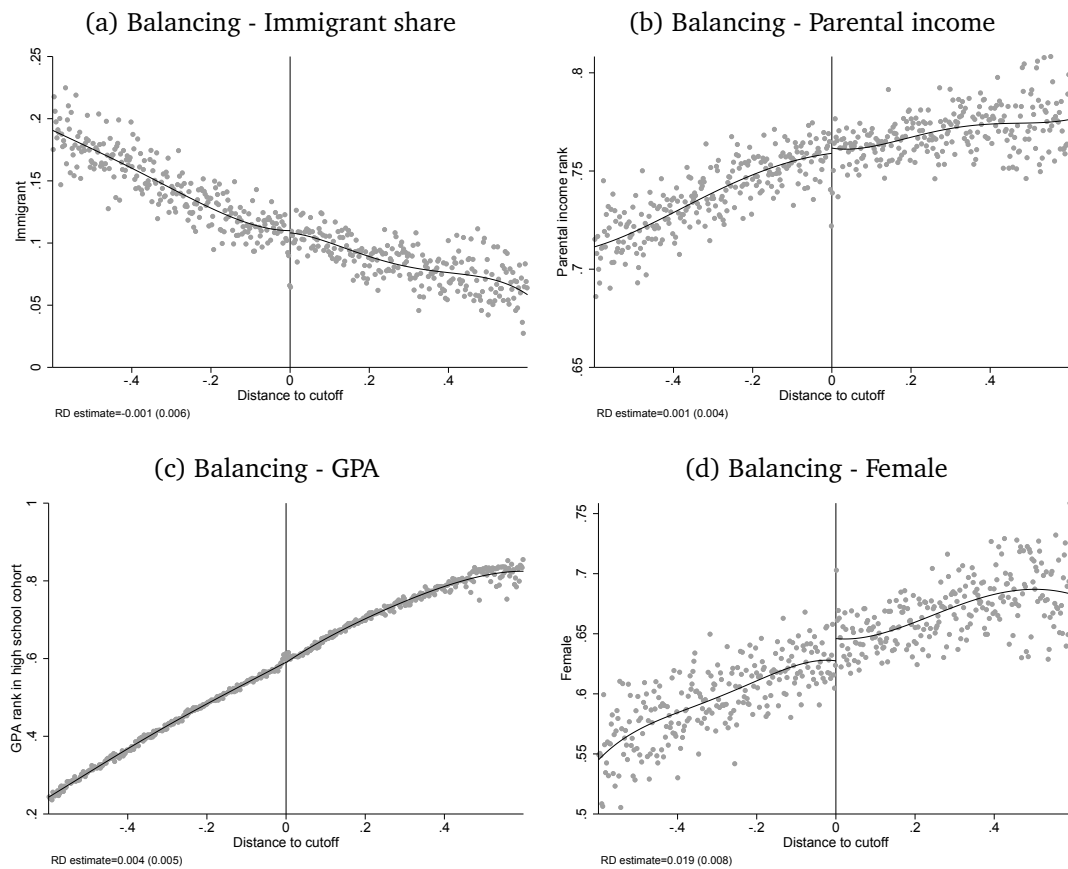
46

(a) Balancing - Immigrant share

(b) Balancing - Parental income

(c) Balancing - GPA

(d) Balancing - Female

Figure A.10: Regression discontinuity plots - balancing

**Note:** 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 across programs and years.
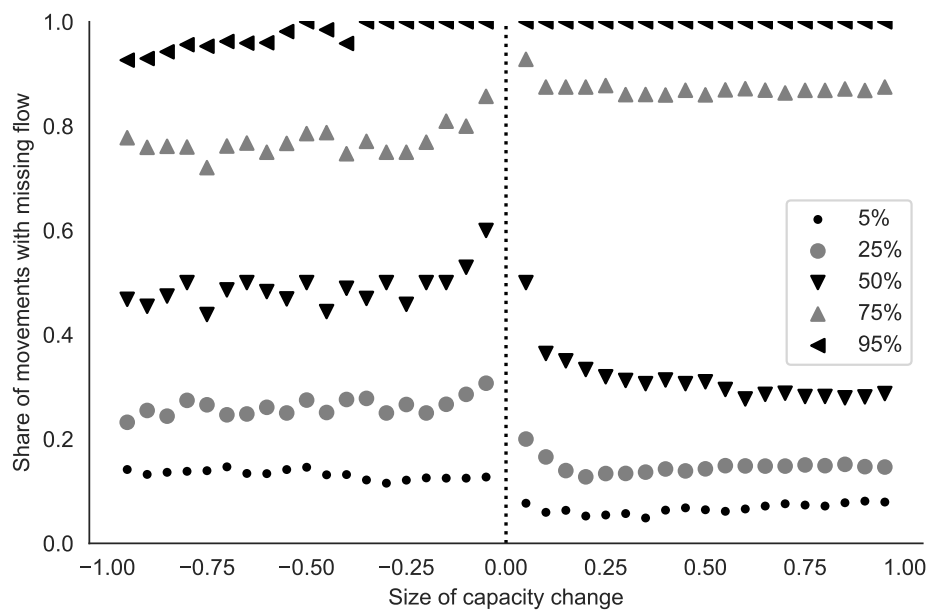
Figure A.11: Share of affected applicants with a missing LATE

**Note:** The figure shows the share of applicants with a missing LATE. For expansions, I use the standard fuzzy RDD design, where the running variable is the ranking in the more preferred program. For contractions, I use the reverse design, where the running variable is the ranking in the less preferred program. The size of the capacity changes is measured in percentage of initial capacity in Quota 1 (GPA-based quota)

# B  Preference estimation

As described in Section 5 I employ the stability-based estimator introduced by Fack et al. (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 $\mu$ 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 a personalized feasible 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 $V_{ip}$:

$$V_{ip} = \mu_{r(p)} + \eta_{l(p)} + \lambda_{f(p)} + \sum_{l'} \mathbf{1}(l(p) = l') \times \gamma_{gpa,l} GPA_i$$

$$+ \mathbf{1}(CV_i = 1) \sum_{f'} \mathbf{1}(f(p) = f') \times (\gamma_{gpa,f} GPA_i + \alpha_{1,l} STEM_i + \alpha_{2,l} SOC_i + \alpha_{3,l} HUM_i + \alpha_{4,l} BUS_i)$$

$$+ \mathbf{1}(CV_i = 0) \sum_{f'} \delta_f \mathbf{1}(f(p) = f'),$$

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.

Standard errors are estimated using the score-bootstrap as described by Abdulkadiroğlu et al. (2020) with a 100 replications.

## C  Empirical Bayes shrinkage

In this section I provide details on the Empirical Bayes shrinkage procedure outlined in Section 7. As mentioned in the main text, this procedure follows Abdulkadiroğlu et al. (2020) closely. The estimates based on conditional independence and the control function approach each return a set of program-specific estimates, $\left\{\hat{\beta}_p\right\}_{p=1}^{P}$.

Let $K$ be the length of the vector of parameters. Under the hierarchical model outlined in Equations (12) and (13), the likelihood of the estimates for program $p$ conditional on the unobserved parameters, $\beta_p$, and the associated covariance matrix, $\Omega_p$, is

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

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

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

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

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

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 $\beta_p$ and $\Omega_p$ are obtained as follows:

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

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