Critical Literature Review:
On the Estimation of Causal Effects in Matching Markets

Term Paper in SM Matching and Market Design

University of Cologne
Faculty of Management, Economics and Social Sciences
Chair of Design and Behavior

Submitted by Lennart Bolwin
Supervised by Prof. Dr. Alexander Westkamp

February 17, 2020

Contents

1 Introduction **1**

2 Summary of Selected Papers **2**

2.1 Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation [1] 2

2.2 Breaking Ties: Regression Discontinuity Design Meets Market Design [2] . . . 5

2.3 Elite Colleges and Upward Mobility to Top Jobs and Top Incomes [3] 9

3 Critical Assessment and Further Research **10**

References **11**

1 Introduction

Many economic markets fail to allocate resources via market-clearing prices. In particular, for some goods and services, not marginal utilities and costs, but rather preferences and priorities determine the distribution of resources. In such scenarios, the theory of matching markets has proven to be a helpful tool in order to allocate scarce capabilities in an as efficient and fair manner as possible. A prominent example of employing the matching approach is the allocation of school or university seats. Here, some schools are more demanded than others, though it is apparent that prices are inapplicable to address the quest of allocating seats in a way that satisfies the interests of all market participants.

Sophisticated matching mechanisms like *Deferred Acceptance* (*DA*) are used nowadays in a growing number of schools and are faced with the challenge of distinguishing between applicants with same preferences and priorities. One method of resolution is random tie-breaking: In case of an indifference, each student in the tie draws a lottery number, and the student with the smallest random number is granted the highest priority. Hereafter, it is assumend that such random numbers are drawn from a continuous, uniform distribution. In other circumstances, non-random tie-breaking (e.g. entry tests) has more appealing features: Besides of breaking the tie, applicants are rewarded for good performance in the test. Independent of the specific tie-breaker nature, tie-breaking generates valuable data for the estimation of school treatment effects in matching markets, where the school treatment effect is defined as the causal effect a school has on certain outcome variables. Under random tie-breaking, the tie-breaker virtually generates an experimental setting, as school treatment is assigned randomly. In case a non-random tie-breaker is employed, the econometric technique of regression-discontinuity design is frequently in use to mimic quasi-experimental environments. The issue of self-selection is closely related to the question of random treatment assignment. In controlled laboratory experiments, the researcher can ensure that treatment allocation is truly random. Thus, it is reasonable not to assume any systematic differences between treatment and control group. When estimating causal effects in matching markets however, self-selection is a severe problem which has to be controlled for. If this is done adequately, causal effects are not compromised and can be estimated unbiasedly. These school treatment estimates are of pivotal importance as they explain partly why some schools are heavily over-demanded while other schools are trying in vain to fill their capacities. If there were no systematic differences across schools, applicant's preferences should be considerably easier to deal with. For this reason, it is important to quantify causal school effects empirically, as they help to explain behavioral patterns of matching market participants.

The following term paper summarizes three research articles that elaborate on the estimation of causal effects in matching markets. While all articles deal with school assignment problems, Abdulkadiroglu et al. (2017) [1] explain how the notion of the propensity score (PS) can be applied to measure causal effects of charter school attendance in Denver. Their model uses a single random tie-breaker which entails random treatment assignment by construction. In Abdulkadiroglu et al. (2019) [2], the authors extend the model from 2017 in a way that allows schools to incorporate random and non-random tie-breakers. Thus, they are able to utilize school data from NYC in order to estimate the causal effects highly demanded schools have on its students. Zimmermann (2019) [3] tries to find evidence for the question why elite college students in Chile reach top positions in the economy. To answer this research task, he makes use of an admission test (non-random tie-breaker) as a source of quasi-experimental variation in school treatment. The last part of this term paper summarises and gives a brief outlook into further research directions.

2 Summary of Selected Papers

2.1 Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation [1]

Many matching mechanisms take applicants' preferences and school priorities as input to assign school seats. Hence, under random tie-breaking, the only variable that could possibly bias causal inference is the applicant's type θ, defined as the combination of preferences and priorities. Suppose for instance, students prefer neighborhood schools. If schools grant higher priorities to students who live nearby, students from certain neighborhoods are likely to self-select into schools close to their place of residence. A straight forward solution to this issue is to condition on applicants of the same type. In many settings however, type is too finely distributed, making full stratification on preferences and priorities infeasible. Abdulkadiroglu et al. exploit the fact that conditioning on a scalar function of type, the propensity score, has been proven to be sufficient to remove any selection bias induced by type. Henceforth, the propensity score is defined as the conditional probability of school assignment given type and provides a much coarser distribution. This feature makes propensity score conditioning in comparison to type conditioning feasible. Once an analytical form of the propensity score is derived, it is easily estimated by repeatedly drawing random student-orderings and computing the empirical probability of assignment to schools for students in the match. This propensity score is then used to measure the impact valuation of charter schools in Denver. The analyzed schools in Denver employ a widely-used version of DA, single tie-breaking DA, that differs from standard DA only by the fact that each student is assigned an uniform iid-random variable and that schools rank applicants first by priority, then by random variable. The next step in the research agenda is to derive an analytical form of the propensity score under single tie-breaking DA.

The probability of assignment to school s depends on two types of uncertainty: First of all, to be offered a seat at s, an applicant has to beat other students in the match that compete for the same seat. Second, an applicant is only seated at s if she fails to be seated at a school that is preferred to s. This structure becomes more comprehensible by considering an example: Suppose there are three applicants applying to two schools with one seat each. Independent of the exact preferences, with three applicants there are $3 \cdot 2 \cdot 1 = 6$ distinct possibilities of ordering the applicants, which is done via the iid-random variable. Then, for each applicant at each school, the propensity score can be calculated by counting the occasions in which the applicant is seated at a given school. Under single tie-breaking DA, the PS-calculation is more involved: As a first step, for each applicant at each school the priority status and the random variable are merged into a single number, named applicant rank:

$$\pi_{is} = \rho_{is} + r_i,$$

where i indexes applicants and s indexes schools. $\rho_{is} \in \{1, 2, ..., K, \infty\}$ encodes applicants' priority, with 1 indicating the highest and K the lowest priority for eligible applicants. ∞ illustrates the fact that i is ineligible at s. r_i denotes i's iid-random variable that is common to all schools. Priorities are assumed to be integers and $r_i \sim \mathcal{U}(0,1)\ \forall\ i$. To further analyse the probability of assignment, school cutoffs need to be incorporated. Therefore, consider the last step of single tie-breaking DA and define the cutoff of s in case of undersubscription as $\xi_s = K + 1$. Since applicants are offered a seat at s if $\pi_{is} \leq \xi_s$, each undersubscribed school ensures that it offers a seat to all eligible applicants. Otherwise, ξ_s is set equal to the last applicant's rank who is offered a seat at s.

Additionally, the concept of a marginal priority group at school s needs to be established in order to fully characterize the DA propensity score. The marginal priority (MP) is defined as the integer part of ξ_s ($= \rho_s$) and enters the stage when type-θ applicants are separated into groups that indicate the likelihood of assignment at s. Let Θ_s be the set of all type-θ applicants who list s and split this set into three groups: Θ_s^n includes applicants at s that fail to clear ξ_s because $\rho_{\theta s} < \rho_s$ (never seated). Likewise, Θ_s^a contains applicants that are always seated at s because $\rho_{\theta s} > \rho_s$[1]. Finally, define Θ_s^c as the group that is conditionally seated at s since $\rho_{\theta s} = \rho_s$. These applicants are seated at s if they clear ξ_s and are not assigned to another school listed above s. The final component before the PS can be calculated mirrors the fact that the probability of assignment at s for applicants in Θ_s^a and Θ_s^c is affected by

[1]As will be seen, this group could possibly improve on s by clearing the cutoff of a school listed above s.

cutoffs of schools preferred to s. The quantity that describes the degree of truncation for the PS of applicants at s is called most informative qualification ($MID_{\theta s}$) and equals zero, if the applicant qualifies at no school listed above s (no truncation). Reversely, complete truncation implies that $MID_{\theta s} = 1$. This occurs if the applicant belongs to the always-seated group of another school listed above s. $MID_{\theta s}$ is equal to the highest cutoff among the preferred schools, if there is at least one school at which the applicant belongs to the conditionally-seated group. The maximum cutoff among preferred schools is chosen, as applicants who fail to clear this cutoff, will also not qualify at schools with more restrictive cutoffs. On the other hand, applicants can qualify at s' (the school among the preferred schools with the highest cutoff) but fail to clear more restrictive cutoffs at preferred schools.

Having collected all necessary parameters, the PS for type-θ applicants at s can be computed:

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_s^n \\ (1 - MID_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ (1 - MID_{\theta s}) \cdot max\left(0, \frac{\xi_s - MID_{\theta s}}{1 - MID_{\theta s}}\right) & \text{if } \theta \in \Theta_s^c \end{cases}$$

The first case is trivial: The probability of assignment at s is zero if an applicant belongs to the never-seated group of s. For applicants in Θ_s^a, the PS is truncated by the fact that applicants could improve on s by clearing $MID_{\theta s}$. With probability $1 - MID_{\theta s}$ applicants fail to qualify at s' and are seated at s because they clear ξ_s with certainty. For applicants in Θ_s^c, one again has to ensure that nobody improves on s by clearing the cutoff of s'. However, this group is only conditionally seated at s, therefore one also has to condition on the probability of clearing ξ_s but not $MID_{\theta s}$. If $MID_{\theta s} > \xi_s$, the max-function yields zero, which generates a PS of zero, as applicants are then assigned to s' instead of s. Consequently, the PS under single tie-breaking DA is a function of few, easily observed parameters and can be estimated for each applicant at each school. Moreover, Abdulkadiroglu et al. prove that their PS-estimate is a strongly consistent estimator of the theoretical PS.

The ultimate goal of PS conditioning is to eliminate a selection bias induced by variables that are related to school treatment assignment. Therefore, after PS conditioning, there should be no systematic difference in applicant characteristics between applicants with the same conditional probability of treatment. The authors provide such evidence in the empirical application, which shows that their PS-estimate was able to remove a possible selection bias from applicant characteristics like sex, skin color and baseline scores in different fields. Subsequently, the PS can be used to measure the causal effect of charter schools in Denver. Loosely speaking, the treatment effect is derived by estimating the effect of a charter school on different variables of interest, while controlling for the PS. The estimation results suggest that charter schools boost scores in math, reading, and writing, having the largest gain in math. It is also worth noting that estimates without score controls are smaller for all three variables, indicating a negative selection bias towards charter schools in Denver.

Abdulkadiroglu et al. (2017) have shown how centralized assignment mechanisms with random tie-breaking can be used to measure causal effects of school attendance. As a random tie-breaker is not assumed to compromise causal inference, the only covariate that has to be controlled for is type. This is done via the empirical counterpart of the PS which proved to be a sufficient statistic to eliminate possible self-selection. The next challenge to tackle is to answer how causal effects can be estimated when both type and tie-breaker are likely to be related to the school treatment.

2.2 Breaking Ties: Regression Discontinuity Design Meets Market Design [2]

In 2017, Abdulkadiroglu et al. have shown how causal effects can be computed under random tie-breaking. However, many schools in matching markets break ties by admission tests, which symbolise non-random tie-breaking with unknown probability distribution on the unit level. Again, only those applicants who successfully clear the school cutoff are offered a seat. Furthermore, many schools apply school-specific tie-breakers (e.g. TM Wiso). Consequently, the two main challenges in the upcoming analysis are to work out how randomization can be ensured under non-random tie-breaking, and how the model from 2017 has to be extended in order to allow schools to only use a subset of all available random and non-random tie-breakers. Apart from these differences, the authors are again seeking for a notion of the PS that can be consistently estimated using school data to assess the question of school quality. It is already known that PS-conditioning was able to remove the selection bias from type. If non-random tie-breakers are employed, there might also be a relation between school offers and the tie-breaker, which has to be controlled for. For example, it is reasonable to argue that more talented applicants perform, on average, better in admission tests than less talented test-takers. They are therefore more likely to be offered seats at highly demanded schools with admission tests, a fact that has to be taken into consideration when calculating the PS. One possibility to warrant random treatment assignment under non-random covariates is given by non-parametric regression-discontinuity (RD) design. In general, RD is used when treatment status is a discontinuous and deterministic function of some covariate x_0. In this application, the covariate is illustrated by the tie-breaker, which satisfies both requirements for the RD design to be valid. As students with vastly different tie-breaker values are likely to have distinct characteristics, the non-parametric RD design is employed subsequently. Here, only applicants within an δ-interval around the randomization cutoff $(x_0 - \delta, x_0 + \delta)$ for $\delta \to 0$ are considered.

Since the probability distribution within the δ-interval is for $\delta \to 0$ approximately uniform, the probability of drawing a value below x_0 is equal to the probability of drawing a value above x_0. Consequently, even under non-random tie-breaking, treatment assignment is locally random if applicants within a δ-interval around the treatment cutoff are considered.

School cutoffs and applicant rank are again described through the variables ξ and π, by which the applicant group is partitioned into always-, never-, and conditionally-seated subgroups. Additionally, all possible tie-breakers are indexed by $v \in \{0, 1, ...V\}$, where $v = 0$ represents the random tie-breaker. The MID-quantity now exists for every tie-breaker in use, and the truncation for tie-breaker v at school s depends on cutoffs of preferred schools that also use v. Apart from that, the truncation follows the same argumentation as in the case with only one tie-breaker: There is no truncation ($MID^v_{\theta s} = 0$) if no other preferred school is using v or if type θ-applicants fail to clear all v-cutoffs among preferred schools. Maximum truncation ($MID^v_{\theta s} = 1$) occurs if the student belongs to the always-seated group of a preferred school using v. $MID^v_{\theta s}$ equals the most forgiving cutoff among preferred schools at which θ-applicants are conditionally seated. The PS for type θ-applicants at s, which embeds the formulation of 2017, has thus the following form:

$$p_s(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta^n_s \\ \prod_v (1 - F_v(MID^v_{\theta s}|\theta)) & \text{if } \theta \in \Theta^a_s \\ \prod_{v \neq v(s)} (1 - F_v(MID^v_{\theta s}|\theta)) \cdot max\{0, max\{0, F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s})\} & \text{if } \theta \in \Theta^c_s \end{cases}$$

This proposition summarises the PS at s for three types of θ-applicants. First, applicants who fail to clear marginal priority at s have no chance of being seated there. Second, applicants that are within the always-seated subgroup of s are only seated there, if they fail to improve on s by clearing the cutoff of a preferred school. As tie-breakers are assumed to be independently distributed, the probability of not improving on s is given by the product of the probabilities of failing to clear $MID^v_{\theta s}$. This probability simplifies to zero respectively one, if there is full, respectively no truncation via $MID^v_{\theta s}$. The last line of the theorem considers again conditionally-seated applicants at s: The first part describes the probability of not clearing the cutoff of a preferred school that is using a tie-breaker different from v. The second argument demonstrates again that applicants are seated at s if and only if their tie-breaker lies between $MID^v_{\theta s}$ and τ_s. Contrary to the model of 2017, these probabilities have no closed form, as the density of a non-random tie-breaker $F(\cdot)$ is unknown. A supplementary difficulty arises, because school assignment is still related to tie-breaking. Therefore, the non-parametric RD design is applied to locally randomize school treatment assignment for applicants within the δ-interval around each schools' cutoff.

To do so, the classification of applicants into subgroups, indicating their probability of assignment, has to be modified. Hereafter, the groups of always- and never-seated applicants at s are extended by those applicants who were originally seated at Θ_s^c but drew a tie-breaker outside the δ-interval around the randomization cutoff τ_s[2]. After this adjustment Θ_s^c contains only applicants within the δ-interval around τ_s. Consequently, treatment assignment is locally random for this subgroup. Before the local PS can be computed, a further variable has to be introduced. Since the probability of assignment at s is truncated by cutoffs of preferred schools, at which θ-applicants belong to the conditionally-seated group, we need to know how many such cutoffs exist. The variable that counts these cases is denoted by $m_s(\theta)$. It is sufficient to focus on preferred schools where the applicant is conditionally seated at because in the remaining two cases, there is either complete or no truncation. In the former case, the probability of assignment at s simplifies to zero, in the latter scenario, preferred schools do not impact the PS at s because the applicant fails to clear their cutoffs anyway. The PS also equals zero if the applicant belongs to the never-seated group of s. In all other cases, the local PS, denoted $\psi_s(\theta)$ with general tie-breaking takes on the following form:

$$\psi_s(\theta) = \begin{cases} 0.5^{m_s(\theta)}(1 - MID^0_{\theta s}) & \text{if } \theta \in \Theta_s^a \\ 0.5^{m_s(\theta)} \cdot max\{0, \tau_s - MID^0_{\theta s}\} & \text{if } \theta \in \Theta_s^c \text{ and } v = 0 \\ 0.5^{1+m_s(\theta)}(1 - MID^0_{\theta s}) & \text{if } \theta \in \Theta_s^c \text{ and } v > 0 \end{cases}$$

In all three cases, the PS is affected by two types of truncation: The applicant can improve on s by clearing random and non-random cutoffs of preferred schools. If the applicant is always seated at s, she can improve on s with 50% at each preferred school with non-random tie-breaker where she falls into the δ-interval around the cutoff[3]. The first term of line one mirrors the probability that this does not occur. The second term depicts the likelihood of not clearing the random cutoff of a preferred school. If s employs a random tie-breaker and the applicant belongs to Θ_s^c, to be seated at s, the applicant has again to fail being seated at a preferred school with non-random tie-breaker, where she is conditionally seated and has to draw a random tie-breaker between τ_s and $MID^0_{\theta s}$. The last line differs from the first line only by the fact that a conditionally-seated applicant at s also only faces a 50% chance of assignment, reflected by the summation of one in the exponent. Examining the limiting behavior, the authors prove again that this theoretical PS can be consistently estimated.

In 2007, the NYC department of education introduced a school accountability system which

[2]Applicants in Θ_s^c have priority $\rho_{\theta s} = \rho_s$. Those who draw a tie-breaker $R_i > \tau_s + \delta$ are added to Θ_s^n, applicants with $R_i \leq \tau_s + \delta$ join the group of Θ_s^a.

[3]The probability of clearing a cutoff for applicants within the δ-interval equals 50%. This follows from the fact that any cumulative density function is approx. uniform within the $\delta-$interval for $\delta \to 0$.

ranked high schools from A to F. These grades where derived on the basis of historical data and reflect achievement levels as well as growth of the schools. Using school application data from 2011-2014, the authors aim to measure the effect of being assigned to a Grade A school. A first comparison between Grade A students and Grade B-F students yields the expected outcome: Students at Grade A schools have higher average SAT scores, higher graduation rates, and are more likely to be considered "college- and career-prepared", a variable that is constructed from multiple scores and grades. The question at hand is to answer to which extend these effects are related to Grade A school attendance rather than to the fact that students at Grade A schools are systematically different from those at Grade B-F schools. Therefore, the authors compare the Grade A effect for all applicants in the match with those that have the same conditional probability of being offered a seat at a Grade A school. More specifically, the estimation of the Grade A effect for all applicants does not take into account that some applicants have, due to different types, a higher probability of Grade A enrollment, whereas the estimation with PS-control conditions on applicants with the same chance of being offered a Grade A seat. Abdulkadiroglu et al. show that most of the gains of Grade A attendance reflect a selection bias. PS-control reduced the estimated gain in the SAT math score by almost 65%, the SAT reading score of Grade A students with similar PS is not significantly larger than the score of Grade B-F students. The effect on the variable "college- and career-preparation" also experienced a decline under PS-control. These results show that Grade A schools do have a causal effect on certain outcomes, though by less than a simple comparison of A and B-F students would suggest.

In the remaining analysis, the authors also estimate the different effects between Grade A schools that use random tie-breakers versus Grade A schools with non-random tie-breakers. Furthermore, they are able to prove that the causal effects of Grade A attendance, which are based on a small subset of students with similar PS, are representative for the entire population in the NYC-match. Finally, it is worth noting that the provided theoretical model embeds several other matching mechanisms: Besides school and student proposing DA, the PS for serial dictatorship (a variant of DA without priorities), as well as for the immediate acceptance mechanism (Boston), can be derived based on the authors' model. Remarkably, top trading cycles are not included as this mechanism fails to ensure that an applicant is only seated at a school if she clears the specific school cutoff and simultaneously does not clear the cutoff of any preferred school.

2.3 Elite Colleges and Upward Mobility to Top Jobs and Top Incomes [3]

Zimmermann tries to explain how elite colleges in Chile impact upward mobility into top jobs and top incomes. For this purpose, he uses administrative data from 1980 until 2001 of the two most selective universities in Chile, stated elite degree programs. The research goal is to identify how the treatment of these universities and two measures of top attainment, namely holding an executive management position, or belonging to the top 0.1% of the income distribution, are related. In Chile, students take a standardized admission exam in their final year of high school and apply to up to eight university programs via a centralized clearinghouse, after observing the exam score. Universities then rank students according to this score and high school grades. Subsequently, university seats are allocated via a version of the deferred acceptance mechanism[4]. Similar to the school assignment mechanisms in Denver and NYC, students are admitted to their most preferred program at which they clear the cutoff. In addition, students who are sufficiently close to the admission cutoff are placed on a waiting list which is made publicly available by the clearinghouse. Students at the top of this waiting list and students at the bottom of the admission list constitute the ingredients for a feasible RD design. While Abdulkadiroglu et al. determined the bandwidth of the δ-interval empirically, Zimmermann compares, due to data restrictions, accepted and rejected applicants who are reasonably similar in terms of average grades and admission test results. Consequently, it is worth examining whether his interval is adequately narrow to ensure local treatment randomization. In order to validate the RD design, Zimmermann first looks for discontinuities in the score density around the admission cutoff. Such a discontinuity would compromise causal inference as students are then assumed to be able to manipulate their exam score. If so, the proportion of students with a score value merely-above the cutoff should be significantly larger than the density of other score values. By plotting the histogram of scores, Zimmermann shows that the score density has a small left skew, but that there is no evidence of individuals self-selecting into the area above the threshold. Additionally, the author performs a balancing test for certain applicant characteristics within the interval, that is similar to the approach by Abdulkadiroglu et al., described on page four of this term paper. Again, applicants within the interval around the treatment cutoff seem to be reasonably similar to each other.

Zimmermann's main results can be summarised as follows: Despite the fact that students admitted to elite programs account for only 1.8% of the entire test-taker population, they make up 41% of the leadership positions and 39% of the richest thousandth in the income

[4]Unfortunately, Zimmermann does not provide the reader with more details regarding the exact structure of the algorithm.

distribution. Stated as a causal effect, enrollment to one of the two elite programs in Chile increased the probability of holding a top leadership position by 44%, and the probability to achieving a top 0.1% income by 55%. Even though it is not straightforward to transfer these results to other countries of interest, they encourage the research on matching mechanisms and demonstrate their importance in order to achieve good welfare properties.

3 Critical Assessment and Further Research

This term paper aimed to summarise two key research contributions in the estimation of causal effects in matching markets. The model of Abdulkadiroglu et al. shows how treatment effects can be consistently estimated under arbitrary tie-breaking. For this purpose, the connection of the propensity score with matching algorithms as proposed by the authors constitutes an important cornerstone. Using the notion of the propensity score enables researchers to determine the probability of treatment in any desired matching market. Apart from the estimation of causal effects, this knowledge can be valuable when assessing the consistency of individual's behavior: For example, it would be irrational to list a school with zero probability of enrollment, given applicant type. However, the calculation of the propensity score is very complex and requires knowledge about cutoffs and preferences of all involved market participants. Therefore, to evaluate consistency in behavior, a simpler and more applicable alternative would be required. Moreover, Abdulkadiroglu et al. demonstrate how the technique of RD design can be incorporated into matching markets. A valid RD design ensures local treatment randomization in non-random settings. While Abdulkadiroglu et al. pursue a formal and theoretically sound approach to locally randomize the sample, Zimmermann follows the practitioners way. Here it is worth noting, that Zimmermanns approach is not less able to compute a valid RD estimate. This is particularly remarkable because there exists a bunch of econometric literature elaborating on the choice of the optimal interval bandwidth. RD could also be used to measure treatment effects in matching markets without tie-breaking. For instance, in any large market approximation, matchings are determined by cutoffs. By considering only individuals with scores sufficiently close to these cutoffs, one again would locally randomize the population.
The presented theoretical model can also be applied to the allocation of medical study places in Germany. Compared to other fields of study, medicine has the appealing feature that seats are allocated via a centralized clearinghouse, named *Hochschulstart*. Among all applicants, 20% of the seats are allocated via the graduation grade of the German Abitur. Since some universities are in such extreme demand, occasionally, even applicants with the best possible graduation grade have to be rejected. These applicants could represent the counterpart of accepted applicants when assessing the effect of highly popular universities.

References

[1] Atila Abdulkadiroglu, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak (2017). *Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation.* Econometrica, 85(5), 1373–1432.

[2] Atila Abdulkadiroglu, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak (2019). *Breaking Ties: Regression Discontinuity Design Meets Market Design.* Cowles Foundation Discussion Papers 2170, Cowles Foundation for Research in Economics, Yale University.

[3] Zimmerman, S. D. (2019). *Elite Colleges and Upward Mobility to Top Jobs and Top Incomes.* American Economic Review, 109 (1), 1–47.