(For Online Publication)

Appendix to

# Preference Discovery in University Admissions: The Case for Dynamic Multioffer Mechanisms

Julien Grenet    YingHua He    Dorothea Kübler

*The Journal of Political Economy*

April 2022

## List of Appendices

**A Data** **2**

**B Supplementary Figures and Tables** **7**

**C Early-Offer Effect: Regression Discontinuity Estimates** **20**

**D Additional Definitions, Proofs, and Results** **25**

**E Simulation-Based Comparison of Mechanisms** **29**

**Appendix References** **39**

# A Data

This appendix provides additional information about the data sets used in the empirical analysis.

## A.1 DoSV Data

The *Dialogorientierten Serviceverfahren* (DoSV) data for the winter term of 2015–16 is managed by the *Stiftung für Hochschulzulassung.* It consists of several files, all of which can be linked using encrypted identifiers for students and programs.

### A.1.1 Data Files

**Applicants.** A specific file provides information on applicants' basic sociodemographic characteristics (gender, year of birth, postal code), their *Abitur* grade, and their final admission outcome, i.e., the reason for exit, the date and time of exit, and (when relevant) the accepted program. The *Abitur* grade is available for only approximately 50% of the applicants but, as explained below (Section A.1.2), it can be inferred for a large fraction of those for whom the information is missing. Possible reasons for exit include (i) the active acceptance of an early offer; (ii) the automatic acceptance of the best offer during Phase 2; (iii) the cancellation of applications; and (iv) rejection due to application errors or rejection in the final stage for students who participated in Phase 2 but received no offer.

**Programs.** For each of the 465 programs that participated in the DoSV procedure in 2015–16, information is provided on the program's field of study and the university where it is located.

**Applicants' rank-order lists of programs.** Applicants' ROLs of programs are recorded on a daily basis throughout the duration of the DoSV procedure, i.e., between April 15 and October 5, 2015. During the Application Phase, students can apply to at most 12 university programs. By default, applications are ranked by their arrival time at the clearinghouse but students may actively change the ordering at any time before Phase 2—with the information recorded in the data.

**Programs' rankings of applicants.** In general, the ranking of applicants by the programs follows a quota system. The size, number, and nature of the quotas are determined by state laws and regulations, and by the universities themselves. For each quota, applicants are ranked according to quota-specific criteria. We make use of the complete rankings of applicants by the programs, including all quotas. So-called preselection quotas are filled before other quotas and are typically applied to 10–20% of a program's seats. They are open to, e.g., foreign students, applicants with professional qualifications, cases of special hardship, and minors. One of the main quotas is the *Abitur* quota (*Abiturbestenquote*) where the ranking is based on a student's average *Abitur* grade and typically applies to 20% of the seats. The Waiting Time Quota (*Wartezeitquote*) is devoted to applicants who have waited for the greatest number of semesters since obtaining the *Abitur*, and typically applies to 20% of the seats as well. Finally, the University Selection Quota (*Auswahlverfahren der Hochschulen*) tends to apply to around 60% of seats and employs criteria that are determined by the programs themselves. However, the ranking under the University Selection Quota is almost entirely determined by the students' *Abitur* grade, with an average correlation coefficient between the rankings submitted by programs and the *Abitur* grade of 0.86 across programs. The order in which the quotas are processed is specific to each university.

**Program offers.** The exact date and time at which offers are made by programs to applicants are recorded in a separate file.

### A.1.2 Additional Information

Based on the data from the DoSV procedure, we computed a number of auxiliary variables.

***Abitur* grades.** In the data, the *Abitur* grade is available for only 49.6% of applicants. However, this information can be inferred for a large fraction of the other applicants based on how they are ranked under programs' *Abitur* quota, because these rankings are strictly determined by an applicant's *Abitur* grade. The grade is given on a 6-point scale to one place after the decimal and ranges between 1.0 (highest grade) and 6.0 (lowest grade). Since the lowest passing grade is 4.0, all applicants in the data have *Abitur* grades between 1.0 and 4.0. Due to the discreteness of the *Abitur*, missing grades can be imputed without error in the following cases: (i) an applicant is ranked above any applicant with a grade of 1.0 (in which case the assigned grade is 1.0); (ii) an applicant is ranked below any applicant with a grade equal to $s$ and above any applicant with the same grade $s$ (in which case the assigned grade is $s$); and (iii) an applicant is ranked below any applicant with a grade of 4.0 (in which case the assigned grade is 4.0). Using this procedure, we were able to impute the *Abitur* grade for approximately two thirds of applicants with a missing grade in the data, bringing the overall proportion of students with a nonmissing *Abitur* grade to 83%.

**Distance to university.** To measure the distance between a student's home and the university of each of the programs she applied to, we geocoded students' postal codes and university addresses, and computed the cartesian distance between the centroid of the student's postal code and the geographic coordinates of each university.

**Feasible programs.** A program is defined as being ex post feasible to a student if the student was ranked above the last applicant to have received an offer from the program under any of the quota-specific rankings in which the student appears. The date the program became feasible to the student $i$ is determined as the first day when $i$, or any student ranked below $i$, received an offer from the program under any of the quota-specific rankings in which $i$ appears.

### A.1.3 Sample Restrictions

The DoSV data contain 183,028 students applying to university programs for the winter term of 2015–16. We exclude 31,066 students for whom the *Abitur* grade is missing and cannot be inferred using the procedure described above, as well as 2,252 students with missing sociodemographic or postal code information. We further remove from the sample 4,097 students who registered to the clearinghouse after the start of Phase 1. Finally, we exclude 34,832 students who applied to specific programs with complex ranking rules, these students being mostly those wanting to become teachers and who have to choose multiple subjects (e.g., math and English). This leaves us with a sample of 110,781 students.

Table 1 in the main text provides summary statistics for this sample, as well as for the subsample of students who applied to at least two programs (64,876 students). To estimate the impact of early offers on the acceptance of offers, we consider only students who applied to at least two feasible programs and either actively accepted an early offer in Phase 1 or were assigned to their best offer by the computerized algorithm in Phase 2. In total, there are 21,711 such students in the sample.

## A.2 Survey

We conducted an online survey between July 27 and October 10, 2015, among students who participated in the DoSV procedure for the winter term of 2015–16. All visitors of the application website were invited to participate in the survey. We collected around 9,000 responses. Of all respondents, 52% completed the survey in July and August while 48% completed it in September and October. The survey formed part of an official survey conducted by the *Stiftung für Hochschulzulassung*, which was aimed at collecting feedback on the DoSV procedure and its website.

Our survey questions focus on the general understanding of the procedure as well as the process of preference formation, including the effect of early offers and the acquisition of information. Since students were able to participate in the survey over a long period of time, we also asked questions regarding the status of their applications, including offers received, rejected, etc. For every question, we included the option "I do not want to answer this question." In the following, we document the complete list of questions (translated from German).

1. How many programs did you apply for through the DoSV? Please provide the number.
2. How many programs did you apply for outside the DoSV? Please provide the number.
3. Which subjects did you apply for through the DoSV? [The list of all subjects grouped in clusters was shown.]
4. Did you apply to some universities in the hope of going there with your friends? [Yes/no]
5. How many offers have you already received? Please consider both offers inside the DoSV and outside of it. Please provide the number.
6. If you have already received an offer, please answer questions 7, 8, 9, and 10. If not, please proceed with question 11.
7. Regarding the offers that you have received up to now [Rate on a Likert scale]
   - Did you talk to your parents about these universities?
   - Did you talk to your friends about these universities?
   - Did you talk to your friends about the possibility of accepting offers at the same university or at universities that are located close to each other?
8. When comparing universities that have made you an offer with universities that have not, can it then be said that [Choose one option]

- On average, I spend more time collecting information on the universities that have made me an offer.
- On average, I spend the same amount of time collecting information on the universities that have made me an offer.
- On average, I spend less time collecting information on the universities that have made me an offer.

9. Regarding the universities that have already made you an offer, which of the following statements best describes your situation? [Choose one option]
    - On average, I find these universities better than before receiving their offers.
    - I find some of these universities better and some worse than before receiving their offers.
    - On average, I find these universities worse than before receiving their offers.
    - The offers did not influence my evaluation of the universities.
10. What is your opinion regarding the acceptance of one of the offers that you have already received? [Rate on a Likert scale]
    - I will accept (or have already accepted) one of the offers since it is from my most preferred university.
    - I will accept (or have already accepted) one of the offers in order to be able to start planning future activities as soon as possible.
    - I will take my time since I want to find out more about the universities.
    - I will take my time since I want to find out where my friends are going to study.
    - I will take my time since I have not received an offer from my preferred university yet.
11. Have any of your friends already received an offer? [Yes/no]
12. If yes, did any of your friends... [Rate on a Likert scale]
    - ... talk to you about the advantages and disadvantages of these universities?
    - ... talk to you about accepting one of these offers?
    - ... consider the possibility of accepting one of the offers from the same or a nearby university together with you or some other friends?
13. Please remember the situation when you submitted your applications to the universities in the DoSV. We would like to know how well you knew at this point how to rank your applications, that is, which application was your most preferred, your second preferred, etc. How accurate are the following statements regarding your situation back then with respect to your preference ranking over the programs? [Rate on a Likert scale]
    - I had a clear preference ranking over the programs.
    - I did not have a clear ranking because I still needed to collect information in order to rank my applications according to my preferences.
    - I did not have a clear ranking because I did not know where my friends were going.
    - Getting to a ranking was very difficult, and I wanted to postpone this decision for as long as possible.
14. Did you actively change your ranking in the DoSV (that is, submitted a new ranking or actively prioritized the applications)? [Yes/no]
15. If no, please provide us with the reasons. [Rate on a Likert scale]

- I did not know that it was possible to change the ranking.
- I was happy with the initial ranking of the DoSV.
- I missed the deadline before which it was possible to change the ranking.
- I did not have a clear ranking of my applications.
- I assume that the ranking has no effect on the likelihood of being admitted.

16. Has your ranking changed between the beginning of the procedure on July 15 and now? [Yes/no]
17. If yes, what were the reasons for changing your ranking? [Rate on a Likert scale]
    - I did not have a ranking at the beginning of the procedure when I submitted my applications.
    - I have received new information during this time period.
    - Now I know where my friends are going.
    - I have received some early offers that have changed my perception of the universities.
18. Have you tried to collect information about the universities during the procedure, in particular... [Rate on a Likert scale]
    - ... via the internet?
    - ... from students of these universities?
    - ... from your school teachers?
    - ... from your parents or other members of your family?
    - ... from your friends?
19. Which of the following reasons have played a role for your selection of programs and universities and for your ranking of them? [Rate on a Likert scale]
    - The fit between the program offered by the university and my own interests.
    - The geographical proximity to my parents.
    - The geographical proximity to my friends.
    - Job market considerations.
    - Whether my application has a chance of being successful at this university.
    - Other reasons.
20. Please tell us your gender. [Female/male]

# B Supplementary Figures and Tables

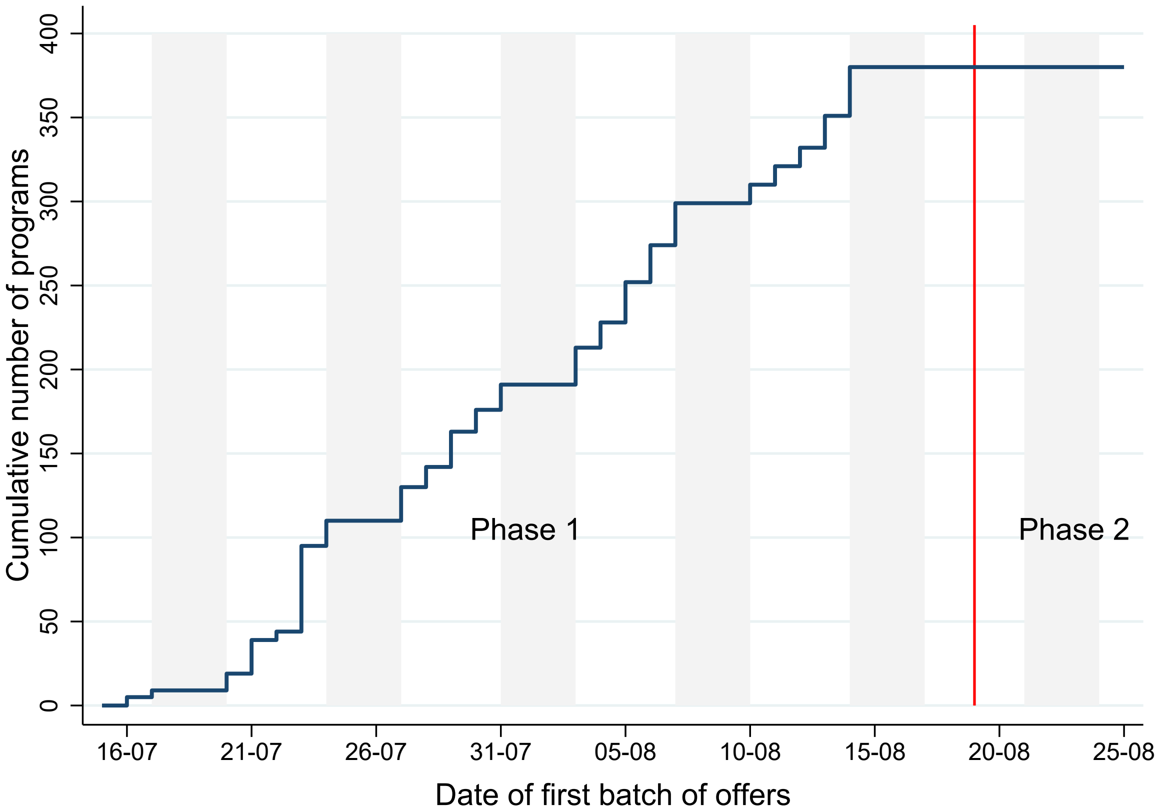



**Figure B1** – First Batch of Offers Sent Out by Programs

*Notes:* This figure shows the cumulative number of programs that have made their first round of offers throughout Phase 1 of the DoSV procedure, i.e., between July 16 and August 18, 2105, based on data from the winter term of 2015–16. Weekends—during which no first round of offers are sent by university programs—are denoted by gray shaded areas.

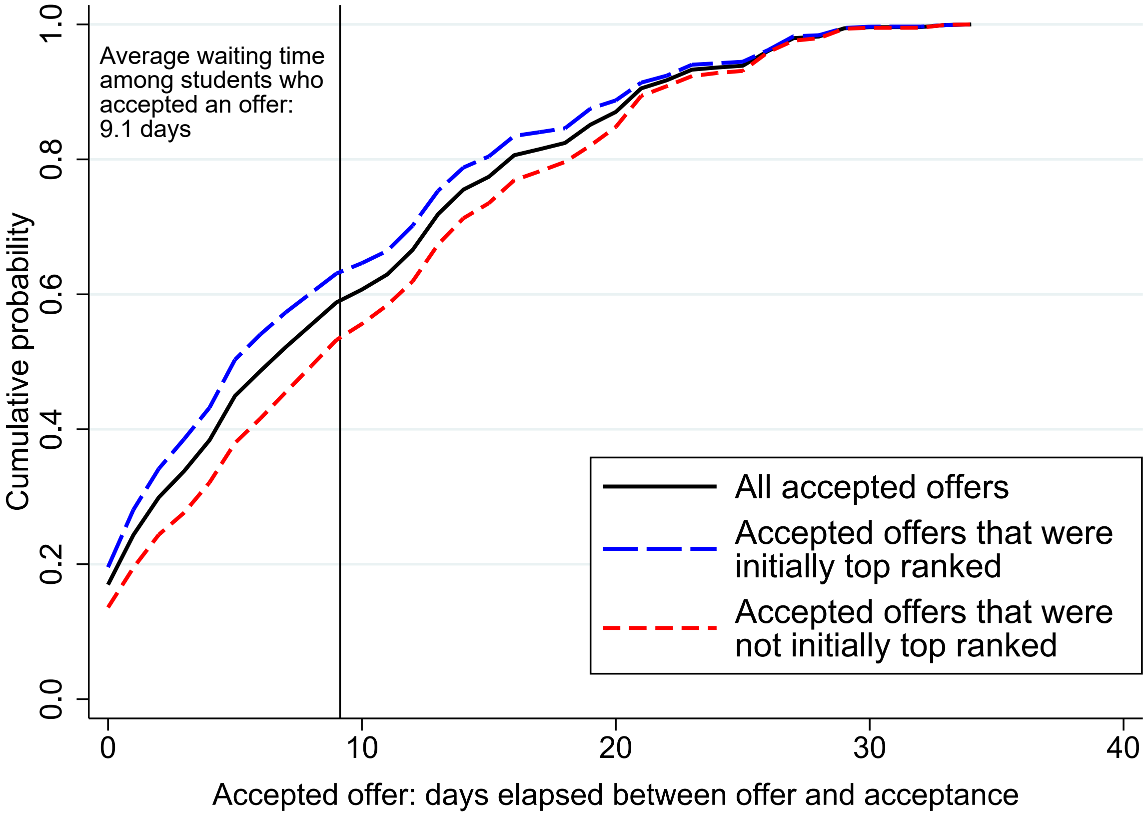



**Figure B2** – Accepted Offer: Cumulative Distribution of Number of Days Elapsed between Offer and Acceptance

*Notes:* This figure shows the cumulative empirical distribution of the number of days elapsed between the date an offer is received by a student and the date it is accepted. The sample is restricted to students who applied to at least two feasible programs and either actively accepted an early offer during Phase 1 or were assigned to their best offer by the computerized algorithm in Phase 2. The different lines correspond to different subsets of accepted offers: (i) all accepted offers (solid line); (ii) accepted offers that were initially top ranked by students (long-dashed line); and (iii) accepted offers that were not initially top ranked by students (short-dashed line).

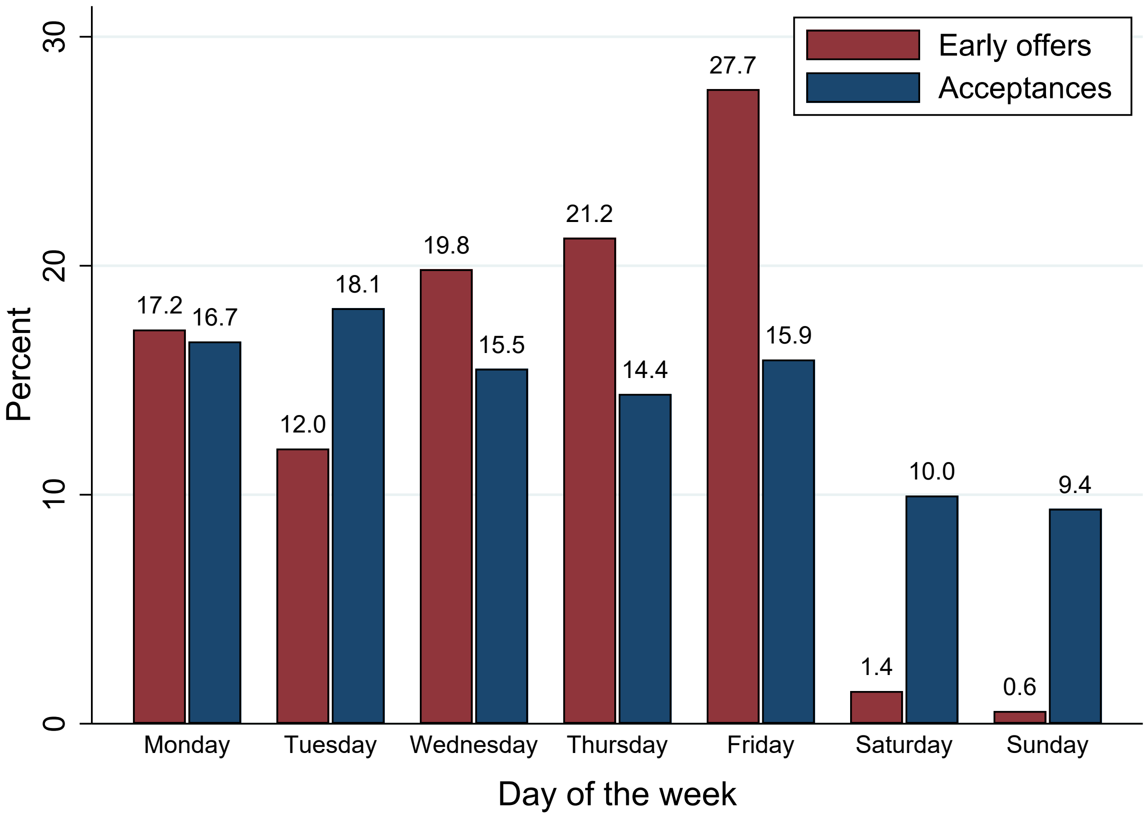



**Figure B3** – Distribution of Early Offers and Acceptances across the Days of the Week

*Notes:* This figure shows the distribution of early offers and acceptances during Phase 1 of the DoSV procedure (i.e., between Thursday, July 16 and Tuesday, August 18, 2015), across the days of the week. The proportions are adjusted to account for the fact that the distribution of days of the week is not balanced during the period (all days but Wednesday have 5 occurrences each whereas Wednesday has 4 occurrences).

**Table B1** – Early Offer and Acceptance among Feasible Programs: Heterogeneity Analysis

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.424*** | 0.568*** | 0.529*** | 0.432*** | 0.592*** |
| | (0.108) | (0.122) | (0.162) | (0.115) | (0.164) |
| × female student | | −0.202** | | | −0.196** |
| | | (0.081) | | | (0.084) |
| × *Abitur* percentile (between zero and one) | | | −0.109 | | −0.022 |
| | | | (0.155) | | (0.165) |
| × number of feasible programs (in excess of two) | | | | −0.005 | 0.004 |
| | | | | (0.024) | (0.025) |
| *FirstEarlyOffer*: First offer in Phase 1 | 0.147*** | 0.176*** | 0.326*** | 0.152*** | 0.339*** |
| | (0.023) | (0.031) | (0.051) | (0.026) | (0.054) |
| × female student | | −0.051 | | | −0.029 |
| | | (0.036) | | | (0.036) |
| × *Abitur* percentile (between zero and one) | | | −0.268*** | | −0.258*** |
| | | | (0.068) | | (0.069) |
| × number of feasible programs (in excess of two) | | | | −0.007 | −0.002 |
| | | | | (0.013) | (0.013) |
| *Controls* | | | | | |
| Distance to university (quadratic) | Yes | Yes | Yes | Yes | Yes |
| Program in student's region (*Land*) | Yes | Yes | Yes | Yes | Yes |
| Program's ranking of student (between zero and one) | Yes | Yes | Yes | Yes | Yes |
| Chances of not receiving an offer from program in Phase 2 | Yes | Yes | Yes | Yes | Yes |
| Program fixed effects (376 programs) | Yes | Yes | Yes | Yes | Yes |
| Number of students | 21,711 | 21,711 | 21,711 | 21,711 | 21,711 |
| Total number of feasible programs | 66,263 | 66,263 | 66,263 | 66,263 | 66,263 |

*Notes:* This table reports estimates from a conditional logit model for the probability of accepting a program among feasible programs. The sample and variables are the same as in Table 3 in the main text. Standard errors are shown in parentheses. **: $p < 0.05$; ***: $p < 0.01$.

**Table B2** – Early Offer and Acceptance among Feasible Programs: Robustness to Contracting Students' Feasible Sets

| | Contracted feasible sets: a program is considered as feasible if the student's rank $\leqslant \overline{r} \times$ admission cutoff rank (with $\overline{r} \leqslant 1$) | | | | | |
|---|---|---|---|---|---|---|
| | $\overline{r}$ = 1.0 | $\overline{r}$ = 0.9 | $\overline{r}$ = 0.8 | $\overline{r}$ = 0.7 | $\overline{r}$ = 0.6 | $\overline{r}$ = 0.5 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***A. Estimates*** | | | | | | |
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.404*** | 0.459*** | 0.465*** | 0.460*** | 0.478*** | 0.681*** |
| | (0.044) | (0.053) | (0.069) | (0.097) | (0.154) | (0.239) |
| *FirstEarlyOffer*: First offer in Phase 1 | 0.147*** | 0.137*** | 0.134*** | 0.131*** | 0.127*** | 0.125*** |
| | (0.023) | (0.023) | (0.024) | (0.024) | (0.024) | (0.024) |
| Distance to university (thousands of km) | −9.37*** | −9.35*** | −9.31*** | −9.39*** | −9.42*** | −9.40*** |
| | (0.33) | (0.34) | (0.34) | (0.34) | (0.35) | (0.35) |
| Distance to university squared | 12.54*** | 12.53*** | 12.41*** | 12.42*** | 12.49*** | 12.45*** |
| | (0.55) | (0.56) | (0.56) | (0.57) | (0.57) | (0.57) |
| Program in student's region (*Land*) | −0.006 | −0.004 | 0.011 | −0.008 | −0.017 | −0.021 |
| | (0.039) | (0.040) | (0.040) | (0.041) | (0.041) | (0.041) |
| Program's ranking of student (between zero and one) | 0.439* | 0.440* | 0.492** | 0.361 | 0.402* | 0.407* |
| | (0.227) | (0.231) | (0.236) | (0.239) | (0.240) | (0.241) |
| Program fixed effects (376) | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of students | 21,711 | 20,911 | 20,300 | 19,925 | 19,713 | 19,627 |
| Total number of feasible programs | 66,263 | 63,523 | 61,358 | 59,986 | 59,218 | 58,925 |
| ***B. Marginal effects on acceptance probability of feasible programs*** | | | | | | |
| Baseline (no early offer) acceptance probability | 0.385 | 0.386 | 0.387 | 0.388 | 0.388 | 0.389 |
| Nonfirst early offer (percentage points) | 8.3 | 9.4 | 9.5 | 9.4 | 9.8 | 14.1 |
| | (1.5) | (1.7) | (1.7) | (1.7) | (1.8) | (2.4) |
| First early offer (percentage points) | 11.3 | 12.3 | 12.3 | 12.2 | 12.5 | 16.6 |
| | (2.0) | (2.1) | (2.1) | (2.1) | (2.2) | (2.7) |

*Notes:* The sample and variables are the same as in Table 3 in the main text. This table shows the results based on the specification in column 4 of Table 3 when we artificially contract students' sets of feasible programs. We proceed by relabelling as "infeasible" from the student's perspective any program that a student included in her initial ROL and whose cutoff was barely cleared by the student in Phase 2. Starting from the main analysis sample, we modify students' feasible sets and acceptance decisions as follows: (i) we relabel as "infeasible" any program that became feasible to the student in Phase 2 and for which the ratio $r$ between the student's rank under the most favorable quota and the rank of the last student who received an offer from the program under that quota is between $\overline{r}$ and 1, with $\overline{r} < 1$ (the most favorable quota is the quota under which the program first became feasible to the student); (ii) we restrict the sample to students who applied to at least two feasible programs under the new definition of program feasibility; (iii) if the student accepted an offer from a program that became feasible in Phase 2 but is no longer feasible under the new definition, we modify the student's acceptance decision by considering that the student accepted the highest ranked offer among the programs that she ranked upon entering Phase 2 and that remain feasible under the new definition of feasibility. The baseline estimates obtained using the observed (ex post) feasible sets of programs are reported in column 1. The results using the contracted feasible sets are shown in columns 2–6 for various choices of the upper limit $\overline{r}$ between 0.5 and 0.9. Standard errors are shown in parentheses. *: $p$ <0.1; **: $p$ <0.05; ***: $p$ <0.01.

**Table B3** – Early Offer and Acceptance among Feasible Programs: Robustness to Expanding Students' Feasible Sets

| | Expanded feasible sets: a program is considered as feasible if the student's rank $\leqslant \bar{r} \times$ admission cutoff rank (with $\bar{r} \geqslant 1$) | | | | | |
|---|---|---|---|---|---|---|
| | $\bar{r} = 1.0$ | $\bar{r} = 1.1$ | $\bar{r} = 1.2$ | $\bar{r} = 1.3$ | $\bar{r} = 1.4$ | $\bar{r} = 1.5$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***A. Estimates*** | | | | | | |
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.404*** | 0.564*** | 0.654*** | 0.711*** | 0.752*** | 0.768*** |
| | (0.044) | (0.034) | (0.030) | (0.029) | (0.027) | (0.027) |
| *FirstEarlyOffer*: First offer in Phase 1 | 0.147*** | 0.152*** | 0.145*** | 0.130*** | 0.117*** | 0.116*** |
| | (0.023) | (0.022) | (0.021) | (0.021) | (0.021) | (0.020) |
| Distance to university (thousands of km) | −9.37*** | −9.36*** | −9.09*** | −9.05*** | −8.90*** | −8.83*** |
| | (0.33) | (0.31) | (0.30) | (0.29) | (0.28) | (0.28) |
| Distance to university squared | 12.54*** | 12.52*** | 12.05*** | 12.00*** | 11.83*** | 11.70*** |
| | (0.55) | (0.52) | (0.50) | (0.49) | (0.47) | (0.46) |
| Program in student's region (*Land*) | −0.006 | 0.008 | 0.021 | 0.026 | 0.021 | 0.033 |
| | (0.039) | (0.037) | (0.036) | (0.034) | (0.033) | (0.033) |
| Program's ranking of student (between zero and one) | 0.439* | 0.463** | 0.490** | 0.568*** | 0.545*** | 0.531*** |
| | (0.227) | (0.217) | (0.209) | (0.203) | (0.197) | (0.192) |
| Program fixed effects (376) | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of students | 21,711 | 23,513 | 25,101 | 26,590 | 27,915 | 29,159 |
| Total number of feasible programs | 66,263 | 72,521 | 78,508 | 84,154 | 89,610 | 94,677 |
| ***B. Marginal effects on acceptance probability of feasible programs*** | | | | | | |
| Baseline (no early offer) acceptance probability | 0.385 | 0.382 | 0.379 | 0.376 | 0.373 | 0.370 |
| Nonfirst early offer (percentage points) | 8.3 | 11.7 | 13.5 | 14.8 | 15.6 | 15.9 |
| | (1.5) | (2.0) | (2.3) | (2.4) | (2.5) | (2.5) |
| First early offer (percentage points) | 11.3 | 14.8 | 16.6 | 17.5 | 18.1 | 18.4 |
| | (2.0) | (2.4) | (2.6) | (2.7) | (2.7) | (2.8) |

*Notes:* The sample and variables are the same as in Table 3 in the main text. This table shows the results based on the specification in column 4 of Table 3 when we artificially expand students' sets of feasible programs. We proceed by relabelling as "feasible" from the student's perspective any program that a student included in her initial ROL and whose cutoff was barely missed by the student in Phase 2. Starting from the sample of students who applied to at least two programs (not necessarily feasible) and did not cancel their application at some point during the procedure, we modify students' feasible sets and acceptance decisions as follows: (i) we relabel as "feasible" any program for which the ratio $r$ between the student's rank under the most favorable quota and the rank of the last student who received an offer from the program under that quota is at most $\bar{r}$, with $\bar{r} > 1$ (the most favorable quota is approximated as the quota under which the ratio $r$ is he smallest for the student); (ii) we restrict the sample to students who applied to at least two feasible programs under the new definition of program feasibility; (iii) if the student participated in Phase 2 and would have been assigned to program $j$ under the new definition of feasibility, i.e., if program $j$ is the highest-ranked feasible program in the student's final ROL, the student is considered as having accepted an offer from that program. The baseline estimates using the observed (ex post) feasible sets of programs are reported in column 1. The results using the expanded feasible sets are shown in columns 2–6 for various choices of the upper limit $\bar{r}$ between 1.1 and 1.5. Standard errors are shown in parentheses. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

**Table B4** – Early Offer and Acceptance among Feasible Programs: By Week in which Program Became Feasible

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *EarlyOffer*: Potential offer from program in Phase 1 | | | | | |
| × Weeks 1, 2 | 0.790*** | 0.838*** | 0.817*** | 0.810*** | 0.800*** |
| | (0.060) | (0.075) | (0.076) | (0.076) | (0.123) |
| × Weeks 3–5 | 0.434*** | 0.372*** | 0.375*** | 0.367*** | 0.356*** |
| | (0.042) | (0.044) | (0.044) | (0.044) | (0.109) |
| *FirstEarlyOffer*: First offer in Phase 1 | | | | | |
| × Weeks 1, 2 | | −0.111** | −0.090* | −0.090* | −0.090* |
| | | (0.047) | (0.048) | (0.048) | (0.048) |
| × Weeks 3–5 | | 0.152*** | 0.169*** | 0.168*** | 0.168*** |
| | | (0.028) | (0.029) | (0.029) | (0.029) |
| Distance to university (thousands of km) | | | −9.35*** | −9.36*** | −9.36*** |
| | | | (0.33) | (0.33) | (0.33) |
| Distance to university squared | | | 12.51*** | 12.53*** | 12.53*** |
| | | | (0.55) | (0.55) | (0.55) |
| Program in student's region (*Land*) | | | −0.007 | −0.008 | −0.008 |
| | | | (0.039) | (0.039) | (0.039) |
| Program's ranking of student (between zero and one) | | | | 0.445* | 0.444* |
| | | | | (0.227) | (0.227) |
| Chances of not receiving an offer from program in Phase 2 | | | | | −0.009 |
| | | | | | (0.076) |
| Program fixed effects (376 programs) | Yes | Yes | Yes | Yes | Yes |
| Number of students | 21,711 | 21,711 | 21,711 | 21,711 | 21,711 |
| Total number of feasible programs | 66,263 | 66,263 | 66,263 | 66,263 | 66,263 |

*Notes:* This table reports estimates from a conditional logit model for the probability of accepting an offer from a feasible program. The sample and variables are the same as in Table 3 in the main text. Standard errors are shown in parentheses. *: $p$ <0.1; **: $p$ <0.05; ***: $p$ <0.01.

**Table B5** – Acceptance among Feasible Programs and Final ROLs: Controlling for How Students Initially Rank Programs

| | Acceptance among feasible (conditional logit) (1) | Final ROL (rank-ordered logit) (2) |
|---|---|---|
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.707*** | 0.653*** |
| | (0.134) | (0.130) |
| *FirstEarlyOffer*: First offer in Phase 1 | 0.189*** | 0.169*** |
| | (0.028) | (0.027) |
| Distance to university (thousands of km) | −6.54*** | −6.35*** |
| | (0.39) | (0.38) |
| Distance to university squared | 8.55*** | 8.23*** |
| | (0.66) | (0.64) |
| Program in student's region (*Land*) | −0.032 | −0.021 |
| | (0.047) | (0.046) |
| Program's ranking of student (between zero and one) | 0.534* | 0.549** |
| | (0.274) | (0.271) |
| Chances of not receiving an offer from program in Phase 2 | 0.050 | 0.052 |
| | (0.095) | (0.092) |
| Student's initial ranking of program (ref.: rank=1) | | |
| rank=2 | −1.422*** | −1.410*** |
| | (0.022) | (0.022) |
| rank=3 | −2.043*** | −2.039*** |
| | (0.032) | (0.031) |
| rank=4 | −2.358*** | −2.362*** |
| | (0.041) | (0.040) |
| rank=5 or above | −3.051*** | −3.068*** |
| | (0.042) | (0.041) |
| Program fixed effects (376 programs) | Yes | Yes |
| Number of students | 21,711 | 21,711 |
| Total number of feasible programs | 66,263 | 66,263 |

*Notes:* The sample and variables are the same as in Table 3 in the main text. Column 1 reports estimates from a conditional logit model for the probability of accepting an offer from a feasible program. Column 2 reports estimates from a rank-ordered logit model for the probability of observing a student's final rank-order list (ROL), using only the payoff-relevant information in each ROL. Standard errors are shown in parentheses. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

**Table B6** – Initial vs. Final Ranking of Feasible Programs: Students who Submitted an Initial ROL that they Reranked in the Application Phase

| | Rank-order list | | | | | |
|---|---|---|---|---|---|---|
| | Initial ROL (at start of Phase 1) | | | Final ROL (at end of Phase 1) | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *EarlyOffer*: Potential offer from program in Phase 1 | −0.069 | −0.051 | −0.069 | 0.490*** | 0.450*** | 0.604*** |
| | (0.043) | (0.044) | (0.125) | (0.067) | (0.070) | (0.186) |
| *FirstEarlyOffer*: First offer in Phase 1 | | −0.050* | −0.035 | | 0.080* | 0.108*** |
| | | (0.027) | (0.027) | | (0.041) | (0.042) |
| Distance to university (thousands of km) | | | −5.19*** | | | −9.85*** |
| | | | (0.33) | | | (0.55) |
| Distance to university squared | | | 6.53*** | | | 12.33*** |
| | | | (0.55) | | | (0.90) |
| Program is in student's region (*Land*) | | | −0.007 | | | −0.001 |
| | | | (0.042) | | | (0.065) |
| Program's ranking of student (between zero and one) | | | −0.001 | | | −0.032 |
| | | | (0.278) | | | (0.453) |
| Chances of not receiving an offer from program in Phase 2 | | | 0.011 | | | 0.131 |
| | | | (0.091) | | | (0.133) |
| Program fixed effects (364) | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of students | 6,473 | 6,473 | 6,473 | 6,473 | 6,473 | 6,473 |
| Total number of feasible programs | 23,116 | 23,116 | 23,116 | 23,116 | 23,116 | 23,116 |

*Notes:* This table repeats the same analysis as in Table 4 in the main text on a restricted sample. The sample here includes only students who applied to at least two feasible programs, who submitted an initial ROL that they reranked at some point in the Application Phase (i.e., before Phase 1), and actively accepted an early offer during Phase 1 or were assigned to their best offer by the computerized algorithm in Phase 2. Standard errors are shown in parentheses. *: $p < 0.1$; ***: $p < 0.01$.

**Table B7** – How Long do Students Wait before Accepting an Offer?

| | Dependent variable: number of days between offer arrival and acceptance | |
|---|---|---|
| | Sample 1: Students with a least two feasible programs who actively accepted an offer in Phase 1 | Sample 2: Sample 1 + students who were automatically assigned to a program in Phase 2 |
| | (1) | (2) |
| Intercept[a] | 11.17*** | 18.17*** |
| | (0.17) | (0.15) |
| Female | −0.228* | 0.004 |
| | (0.100) | (0.088) |
| *Abitur* percentile (between zero and one) | 0.270 | −0.369* |
| | (0.182) | (0.162) |
| Distance to university (thousands of km) | 4.99*** | 15.91*** |
| | (1.31) | (1.10) |
| Distance to university squared | −8.02*** | −24.40*** |
| | (2.41) | (1.98) |
| Program is not in student's region (*Land*) | 0.032 | 0.365** |
| | (0.138) | (0.121) |
| Student's initial ranking of program (ref.: rank=1) | | |
| rank = 2 | 2.637*** | 1.150*** |
| | (0.125) | (0.113) |
| rank = 3 | 2.855*** | 1.615*** |
| | (0.176) | (0.155) |
| rank = 4 | 3.590*** | 1.841*** |
| | (0.229) | (0.202) |
| rank=5 or above | 3.566*** | 2.166*** |
| | (0.212) | (0.183) |
| Number of days between start of Phase 1 and date of offer arrival | −0.419*** | −0.597*** |
| | (0.006) | (0.005) |
| Number of programs in initial ROL (in excess of 2) | 0.086*** | 0.046* |
| | (0.024) | (0.021) |
| Number of other offers held when accepting offer | 0.659*** | 0.579*** |
| | (0.039) | (0.036) |
| Number of observations | 12,025 | 21,711 |
| Adjusted $R$-squared | 0.343 | 0.435 |
| Mean waiting time before accepting offer (in days) | 6.67 | 9.11 |
| | (6.50) | (8.30) |

*Notes:* This table reports estimates from a regression where the dependent variable is the number of days between the date an offer was received by a student and the date it was accepted. The sample in column 1 includes students who applied to at least two feasible programs and actively accepted an early offer in Phase 1. The sample in column 2 further includes students who were assigned to their best offer by the computerized algorithm in Phase 2 (with an acceptance date set to the first day of Phase 2, i.e., August 19, 2015). Standard errors are shown in parentheses. *: $p$ <0.1; **: $p$ <0.05; ***: $p$ <0.01.

[a] The regression intercept can be interpreted as the mean waiting time before accepting an offer that was received by a male student at the lowest percentile of the *Abitur* grade distribution, from a program located in the student's region, that was initially ranked in first position in a two-choice rank-order list, when the offer arrives on the first day of Phase 1 and no other offers are held.

**Table B8** – Early Offer and Acceptance among Feasible Programs: Using Flexible Controls for a Program's Ranking of the Student

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| ***A. Estimates*** | | | | | |
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.424*** | 0.455*** | 0.454*** | 0.436*** | 0.443*** |
| | (0.108) | (0.109) | (0.109) | (0.108) | (0.109) |
| *FirstEarlyOffer*: First offer in Phase 1 | 0.147*** | 0.153*** | 0.154*** | 0.148*** | 0.151*** |
| | (0.023) | (0.023) | (0.023) | (0.023) | (0.023) |
| Distance to university (thousands of km) | −9.37*** | −9.41*** | −9.41*** | −9.39*** | −9.39*** |
| | (0.33) | (0.33) | (0.33) | (0.33) | (0.33) |
| Distance to university squared | 12.54*** | 12.60*** | 12.60*** | 12.57*** | 12.58*** |
| | (0.55) | (0.55) | (0.55) | (0.55) | (0.55) |
| Program in student's region (*Land*) | −0.006 | −0.006 | −0.005 | −0.005 | −0.005 |
| | (0.039) | (0.039) | (0.039) | (0.039) | (0.039) |
| Chances of not receiving an offer from program in Phase 2 | 0.016 | 0.024 | 0.022 | 0.018 | 0.021 |
| | (0.076) | (0.076) | (0.076) | (0.076) | (0.076) |
| Program's ranking of student (between zero and one) | Linear | Quadratic | Quartic | Quartiles | Deciles |
| Program fixed effects (376) | Yes | Yes | Yes | Yes | Yes |
| Number of students | 21,711 | 21,711 | 21,711 | 21,711 | 21,711 |
| Total number of feasible programs | 66,263 | 66,263 | 66,263 | 66,263 | 66,263 |
| ***B. Marginal effects on acceptance probability of feasible programs*** | | | | | |
| Baseline (no early offer) acceptance probability: 38.5% | | | | | |
| Nonfirst early offer (percentage points) | 8.7 | 9.3 | 9.3 | 8.9 | 9.1 |
| | (1.6) | (1.7) | (1.7) | (1.7) | (1.7) |
| First early offer (percentage points) | 11.8 | 12.5 | 12.5 | 12.0 | 12.2 |
| | (2.1) | (2.2) | (2.2) | (2.1) | (2.1) |

*Notes:* The sample and variables are the same as in Table 3 in the main text. This table shows the results obtained when using alternative ways of controlling for a program's ranking of the student, which is measured as the student's percentile (between zero and one, with a higher value indicating a better rank) among all applicants under the *Abitur* quota: a linear control (column 1, which replicates column 5 of Table 3); a quartic (column 2) or quadratic (column 3) function; dummies for quartiles of this variable (column 4); and dummies for deciles (column 5). Standard errors are shown in parentheses. ***: $p < 0.01$.

**Table B9** – Early Offer and Acceptance among Feasible Programs: Students who Applied only to Programs Located in their Municipality of Residence

| | (1) | (2) | (3) |
|---|---|---|---|
| ***A. Estimates*** | | | |
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.661*** | 0.466*** | 0.461*** |
| | (0.121) | (0.131) | (0.131) |
| *FirstEarlyOffer*: First offer in Phase 1 | | 0.312*** | 0.307*** |
| | | (0.079) | (0.079) |
| Program's ranking of student (between zero and one) | | | 0.289 |
| | | | (0.649) |
| Program fixed effects (273 programs) | Yes | Yes | Yes |
| Number of students | 2,459 | 2,459 | 2,459 |
| Total number of feasible programs | 6,612 | 6,612 | 6,612 |
| ***B. Marginal effects on acceptance probability of feasible programs*** | | | |
| Baseline (no early offer) acceptance probability: 41.5% | | | |
| Nonfirst early offer (percentage points) | 13.7 | 9.6 | 9.5 |
| | (3.3) | (2.4) | (2.4) |
| First early offer (percentage points) | | 16.0 | 15.7 |
| | | (3.9) | (3.9) |

*Notes:* This table repeats the same analysis as in Table 3 on a restricted sample. The sample is restricted to students who applied only to programs located in their municipality of residence. Standard errors are shown in parentheses. ***: $p < 0.01$.

**Table B10** – Early Offer and Acceptance among Feasible Programs: Students who Did not Accept an Early Offer until at least Halfway Through Phase 1

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| ***A. Estimates*** | | | | | |
| *EarlyOffer*: Potential offer from program in Phase 1 | 0.446*** | 0.411*** | 0.410*** | 0.405*** | 0.412*** |
| | (0.042) | (0.044) | (0.045) | (0.045) | (0.109) |
| *FirstEarlyOffer*: First offer in Phase 1 | | 0.063*** | 0.078*** | 0.077*** | 0.077*** |
| | | (0.023) | (0.024) | (0.024) | (0.024) |
| Distance to university (thousands of km) | | | −8.77*** | −8.78*** | −8.78*** |
| | | | (0.34) | (0.34) | (0.34) |
| Distance to university squared | | | 11.62*** | 11.63*** | 11.63*** |
| | | | (0.56) | (0.56) | (0.56) |
| Program in student's region (*Land*) | | | −0.002 | −0.003 | −0.003 |
| | | | (0.040) | (0.040) | (0.040) |
| Program's ranking of student (between zero and one) | | | | 0.346 | 0.347 |
| | | | | (0.237) | (0.237) |
| Chances of not receiving an offer from program in Phase 2 | | | | | 0.006 |
| | | | | | (0.076) |
| Program fixed effects (273 programs) | Yes | Yes | Yes | Yes | Yes |
| Number of students | 19,693 | 19,693 | 19,693 | 19,693 | 19,693 |
| Total number of feasible programs | 60,394 | 60,394 | 60,394 | 60,394 | 60,394 |
| ***B. Marginal effects on acceptance probability of feasible programs*** | | | | | |
| Baseline (no early offer) acceptance probability: 38.4% | | | | | |
| Nonfirst early offer (percentage points) | 9.5 | 8.7 | 8.4 | 8.3 | 8.4 |
| | (1.5) | (1.4) | (1.6) | (1.5) | (1.6) |
| First early offer (percentage points) | | 10.1 | 10.0 | 9.9 | 10.1 |
| | | (1.5) | (1.8) | (1.8) | (1.8) |

*Notes:* This table repeats the same analysis as in Table 3 on a restricted sample. The sample is restricted to students who did not accept an early offer until at least halfway through Phase 1, i.e., until August 2 (Phase 1 lasted 34 days, from July 16 to August 18). Standard errors are shown in parentheses. ***: $p < 0.01$.

# C Early-Offer Effect: Regression Discontinuity Estimates

This appendix uses a regression discontinuity (RD) design to provide supplementary evidence that early offers are accepted more often than later ones. This design exploits the fact that a student's receipt of an early offer during Phase 1 of the DoSV procedure is determined by the student's position in the program's ranking of its applicants. The effect of early offers on the acceptance probability can therefore be estimated by comparing the acceptance behavior of students ranked just above versus just below a program's Phase 1 cutoff rank, i.e., the rank of the last student who received an early offer from the program in Phase 1.

**Limitations of the RD design in the DoSV setting.** The reason why we do not adopt an RD design as our main empirical strategy is that it has a number of limitations in our setting.

A first limitation comes from the fact that in the DoSV procedure, each program has several quotas (the average program has six) and an applicant appears on multiple rankings of the same program, e.g., the one for the *Abitur* quota (*Abiturbestenquote*) and the one for the Waiting time quota (*Wartezeitquote*). An undesirable consequence of these multiple rankings is that a student who missed a program's Phase 1 cutoff under quota $q$ can receive an early offer from the program under a different quota $q'$ provided that she clears this other quota's cutoff by the end of Phase 1. As a result, the RD design is fuzzy rather than sharp—the observed discontinuity in the probability of receiving a potential early offer at the Phase 1 cutoff of a program's quota is less than one.

A second limitation of the RD design is that the comparison of students' acceptance decisions around the Phase 1 cutoffs allows us to estimate only the early-offer effect on the probability of accepting a program and not to compare the effects of the first versus subsequent early offers or to analyze students' reranking behavior.

A third limitation is that the RD design identifies the early-offer effect only for the subgroup of students who barely cleared or barely missed the cutoff to receive an early offer, whereas we are interested in estimating this effect for a broader population of applicants.

**Sample restrictions.** Bearing in mind these limitations, we implement a fuzzy RD design by restricting the sample to the subset of students and programs for which this approach can be applied. We start by considering all programs that made offers in both Phase 1 and Phase 2 and keep only the relevant quotas, i.e., those under which offers were made in both phases. Let $\overline{R}^1_{j,q}$ denote the rank of the last student who received an offer from program $j$ under quota $q$ by the end of Phase 1, and let $\overline{R}^2_{j,q}$ denote the rank of the last applicant who received an offer from program $j$ under quota $q$ by the end of Phase 2. By construction, $\overline{R}^2_{j,q} \geqslant \overline{R}^1_{j,q}$. We restrict the set of program quotas $(j, q)$ to those having at least 10 applicants ranked above the Phase 1 cutoff ($\overline{R}^1_{j,q}$) and at least 10 applicants ranked between the Phase 1 and Phase 2 cutoffs (i.e., between $\overline{R}^1_{j,q}$ and $\overline{R}^2_{j,q}$). To ensure consistency with our main empirical results, we consider only students who applied to at least two feasible programs and accepted an offer.

**Empirical specification.** Our RD estimates of the early-offer effect are based on the following empirical specification:

$$Accept_{i,j,q} = \delta\, EarlyOffer_{i,j,q} + f(\tilde{R}^1_{i,j,q}) + \epsilon_{i,j,q}, \tag{A.1}$$
$$EarlyOffer_{i,j,q} = \pi \mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\} + g(\tilde{R}^1_{i,j,q}) + \nu_{i,j,q}, \tag{A.2}$$

where $Accept_{i,j,q}$ is an indicator that equals one if student $i$, ranked under program quota $(j, q)$, accepted an offer (in Phase 1 or Phase 2) from program $j$ under any quota; $EarlyOffer_{i,j,q}$ is an indicator that takes the value one if the student had a potential early offer (i.e., before the end of Phase 1) from program $j$ under any quota; the forcing variable $\tilde{R}^1_{i,j,q} \equiv R_{i,j,q} - \overline{R}^1_{j,q}$ is the distance between the student's rank under the program quota $(j, q)$ and the rank of the last student who received an offer in Phase 1 under that quota; $\mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\}$ is an indicator that is equal to one if the student cleared the Phase 1 cutoff for program $j$ under quota $q$, and hence was eligible to receive an early offer from the program under that quota; $f(\cdot)$ and $g(\cdot)$ are polynomial functions of the forcing variable $\tilde{R}^1_{i,j,q}$.

The RD instrumental variable estimator using $\mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\}$ as an instrument for $EarlyOffer_{i,j,q}$ identifies the local average treatment effect of early offers on the acceptance probability under two main assumptions: (i) $\mathbb{E}(\epsilon_{i,j,q}|\tilde{R}^1_{i,j,q})$ and $\mathbb{E}(\nu_{i,j,q}|\tilde{R}^1_{i,j,q})$ are continuous, i.e., the assignment of students on either side of Phase 1 cutoff is as good as random; (ii) crossing the cutoff affects students' acceptance probability only through the increased probability of receiving an early offer.

We implemented this fuzzy RD design using the statistical package `rdrobust` described in Calonico et al. (2017). We present estimates from data-driven bandwidths that are mean square error (MSE)-optimal as proposed by Imbens and Kalyanaraman (2012) as well as the bias-corrected estimates and robust standard errors proposed by Calonico, Cattaneo and Titiunik (2014). Since a student can be ranked under multiple program quotas, standard errors are clustered at the individual level using the nearest neighbor variance estimation method described in the same study.

The linear reduced-form specifications for observations within a distance $h$ of the Phase 1 cutoff are as follows:

$$EarlyOffer_{i,j,q} = \pi \mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\} + \rho_0 + \rho_1 \tilde{R}^1_{i,j,q} + \rho_2 \tilde{R}^1_{i,j,q} \times \mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\} + \eta_{i,j,q}, \tag{A.3}$$
$$Accept_{i,j,q} = \gamma \mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\} + \lambda_0 + \lambda_1 \tilde{R}^1_{i,j,q} + \lambda_2 \tilde{R}^1_{i,j,q} \times \mathbb{1}\{\tilde{R}^1_{i,j,q} \leqslant 0\} + \xi_{i,j,q}, \tag{A.4}$$

where $\gamma = \delta \times \pi$.

Note that, by construction, the probability of receiving an early offer is equal to one for the last student who received an offer in Phase 1. To mitigate concerns arising from this endogenous stopping rule, we follow the recommendation in de Chaisemartin and Behaghel (2020) of dropping the last applicant who received an offer in Phase 1.[1]

**Graphical evidence.** Figure C1 plots the density of applicants on either side of the Phase 1 cutoffs, after pooling all program quotas and centering their cutoffs at 0. The manipulation testing procedure is implemented using the local polynomial density estimators proposed by Cattaneo, Jansson and Ma (2020). The results show no statistical evidence

[1]The authors consider the closely related problem of estimating treatment effects allocated by randomized waiting lists.

against the null hypothesis that the density is smooth around the Phase 1 cutoff.[2]

Panel A of Figure C2 provides graphical evidence of the first stage, i.e., the discontinuity in the probability of receiving a potential early offer when crossing the Phase 1 cutoff of a program's quota. The x-axis represents the distance between a student's rank under the program's quota and the Phase 1 cutoff rank for this quota. The graphical evidence shows that the early-offer probability increases discontinuously for students who barely cleared the Phase 1 cutoff. The induced discontinuity is smaller than one because some students who barely missed the cutoff for the considered program quota were ranked above the Phase 1 cutoff for a different quota of the same program, and hence could receive an early offer from that program.

Panel B of Figure C2 presents graphical evidence of the discontinuity in the probability of accepting program $j$ at the Phase 1 cutoff for a quota of the program. There is clear evidence that the acceptance probability increases discontinuously for students who barely cleared the cutoff.

**RD estimates of the early-offer effect.** Table C1 presents the results. The specifications in columns 1 and 2 do not include covariates whereas those in columns 3 and 4 include program-quota fixed effects. Columns 1 and 3 uses the mean-squared-error-optimal bandwidth following Imbens and Kalyanaraman (2012). Columns 2 and 4 in addition report bias-corrected estimates and robust standard errors following Calonico, Cattaneo and Titiunik (2014). Standard errors are clustered at the student level using the nearest neighbor variance estimation method described in the same study.

Panel A reports first-stage RD estimates of the discontinuity in the probability of receiving an early offer from a program when crossing the Phase 1 cutoff of a program's quota. The results indicate that the early-offer probability increases significantly at the cutoff, by 76.2 to 76.4 percentage points from a baseline of 26.8% for students ranked below the Phase 1 cutoff.

Panel B reports reduced-form RD estimates of the corresponding discontinuity in the probability of accepting a program's offer. Consistent with an early-offer effect, the results indicate that the probability of accepting a program's offer increases significantly at the cutoff, by 6.2 to 7.0 percentage points from a baseline of 25.8% for students ranked below the Phase 1 cutoff.

Panel C reports RD IV estimates of the impact of receiving an early offer from a program on the probability of ultimately accepting that program, where the estimand of interest is the ratio between the estimand from the reduced-form equation (discontinuity in acceptance probability) and the estimand from the first-stage equation (discontinuity in the early-offer probability). The results indicate that receiving an early offer from a program significantly increases the probability of accepting that program's offer by 8.2 to 9.1 percentage points. These RD estimates are remarkably similar to those we obtain using the conditional logit model: in our preferred specification (Table 3, column 5), an early offer is estimated to increase the acceptance probability by 8.7 percentage points.

[2]The reason why the density exhibits a spike at the Phase 1 cutoff is that we pool together all program quotas and center their cutoffs. Some program quotas have many students ranked above and below the Phase 1 cutoff whereas others have few.

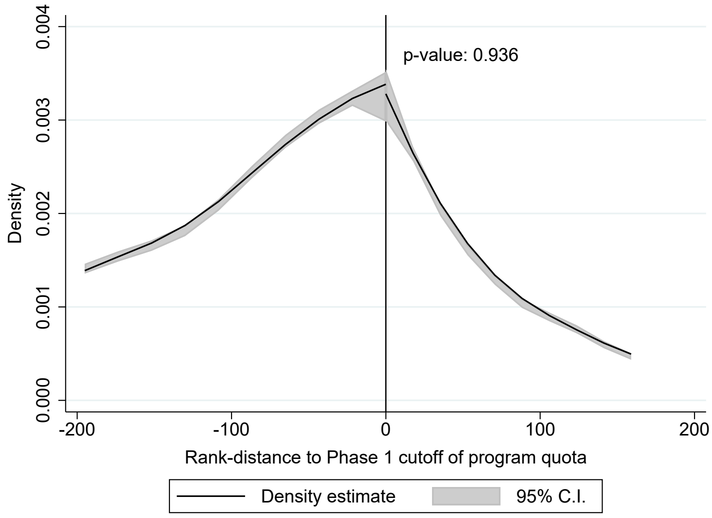
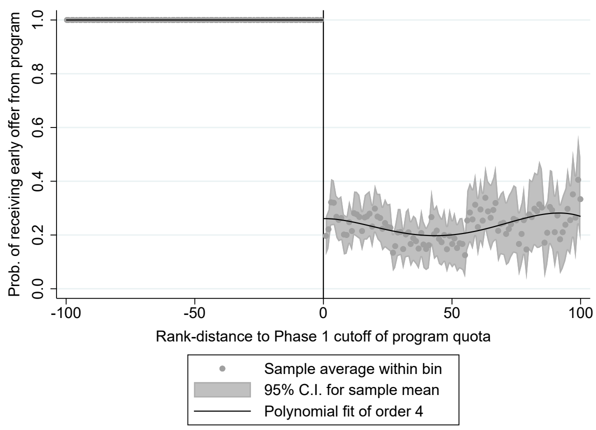



**Figure C1** – Phase 1 Cutoff: Density Test

*Notes:* This figure implements a manipulation testing procedure using the local polynomial density estimators described in Cattaneo, Jansson and Ma (2020) (Stata command `rddensity`). The solid line indicates the density estimate and the shaded area shows the 95% confidence interval. The program quotas considered are those under which offers were made in both Phase 1 and Phase 2, with at least 10 students ranked above the Phase 1 cutoff rank and at least 10 students ranked below the Phase 1 cutoff and above the Phase 2 cutoff. The sample consists of all students who applied to at least two feasible programs and whose rank under one of the selected program quotas was above the quota's Phase 2 cutoff.

**A.** Probability of a (potential) early offer from program




**B.** Probability of accepting an offer from program

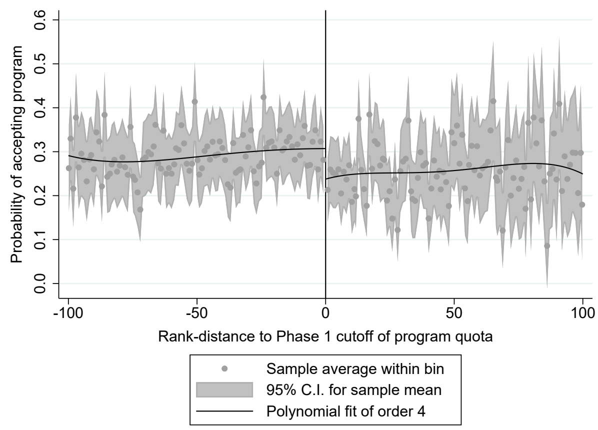



**Figure C2** – Probability of Receiving and Accepting an Early Offer at a Program's Phase 1 Cutoff

*Notes:* This figure shows the probability of receiving a (potential) early offer (Panel A) and the probability of accepting an offer (Panel B) from a program as a function of the student's rank-distance to the Phase 1 cutoff of one of the program's quotas. The Phase 1 cutoff is the rank of the last student who received an early offer from the program under the considered quota in Phase 1. We restrict program quotas to those under which offers were made in both Phase 1 and Phase 2, with at least 10 students ranked above the Phase 1 cutoff rank and at least 10 students ranked below the Phase 1 cutoff and above the Phase 2 cutoff. The sample pools together all students who applied to at least two feasible programs and whose rank under one of the considered program quotas was above the quota's Phase 2 cutoff. Following de Chaisemartin and Behaghel (2020), the last student to receive an early offer under any program quota is removed. The plots are obtained using the Stata command `rdplot` described in Calonico et al. (2017). The dots represent early-offer probabilities (Panel A) and acceptance probabilities (Panel B) averaged within bins. The bins are selected to balance squared-bias and variance so that the integrated mean squared error is approximately minimized. The solid lines represent kernel-weighted fourth-order polynomial fits using a uniform kernel. The shaded area represents the 95% confidence intervals for local means within each bin.

**Table C1** – Impact of (Potential) Early Offers on Acceptance Probability: RD Estimates

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| ***A. First Stage: Discontinuity in probability of (potential) early offer from program*** | | | | |
| $\mathbb{1}\{\text{Rank} \leqslant \text{Phase 1 cutoff}\}$ | 0.764*** | 0.762*** | 0.762*** | 0.763*** |
| | (0.009) | (0.010) | (0.008) | (0.009) |
| Baseline mean (Rank > Phase 1 cutoff) | 0.268 | 0.268 | 0.268 | 0.268 |
| ***B. Reduced Form: Discontinuity in probability of accepting a program's offer*** | | | | |
| $\mathbb{1}\{\text{Rank} \leqslant \text{Phase 1 cutoff}\}$ | 0.066*** | 0.070*** | 0.062*** | 0.066*** |
| | (0.012) | (0.014) | (0.012) | (0.013) |
| Baseline mean (Rank > Phase 1 cutoff) | 0.258 | 0.258 | 0.258 | 0.258 |
| ***C. RD IV: Impact of early offer on acceptance probability*** | | | | |
| *EarlyOffer* | 0.086*** | 0.091*** | 0.082*** | 0.087*** |
| | (0.016) | (0.018) | (0.016) | (0.017) |
| Number of observations | 40,103 | 40,103 | 40,103 | 40,103 |
| Number of students | 19,659 | 19,659 | 19,659 | 19,659 |
| RD estimation method | Conventional | Bias-corrected | Conventional | Bias-corrected |
| Bandwidth (rank-distance to Phase 1 cutoff) | ±124.5 | ±124.5 | ±125.7 | ±125.7 |
| Spline | Linear | Linear | Linear | Linear |
| Program-quota fixed effects (208) | No | No | Yes | Yes |

*Notes:* Panel A reports the first-stage RD estimates of the discontinuity in the probability of a (potential) offer from a program when crossing the Phase 1 cutoff of a program's quota. The Phase 1 cutoff of a program quota is the rank of the last student who received an early offer from the program under that quota in Phase 1. Panel B reports the reduced-form RD estimates of the corresponding discontinuity in the probability of accepting a program's offer. Panel C reports the RD IV estimates of the impact of an early offer from a program on the probability of accepting the program's offer. The program quotas considered in the analysis are those under which offers were made in both Phase 1 and Phase 2, with at least 10 students ranked above the Phase 1 cutoff rank and at least 10 students ranked below the Phase 1 cutoff and above the Phase 2 cutoff. The sample pools together all students who applied to at least two feasible programs and whose rank under one of the considered program quotas was above the quota's Phase 2 cutoff. Following de Chaisemartin and Behaghel (2020), the last student to receive an early offer under any program quota is removed. The estimates are obtained using the `rdrobust` package described in Calonico et al. (2017). The specifications in columns 1 and 2 do not include covariates whereas those in columns 3 and 4 include program-quota fixed effects. Columns 1 and 3 use the mean-squared-error-optimal bandwidth following Imbens and Kalyanaraman (2012). Columns 2 and 4 in addition report bias-corrected estimates and robust standard errors following Calonico, Cattaneo and Titiunik (2014). Standard errors are clustered at the student level using the nearest neighbor variance estimation method described in the same study. ***: $p$ <0.01.

# D Additional Definitions, Proofs, and Results

This appendix first defines the Gale-Shapley algorithm (Section D.1). Then, it provides additional results and proofs for Section IV in the main text. Specifically, Section D.2 considers a special case in which the student knows her own ordinal preferences, an assumption that is often imposed in the matching literature. Under this assumption, we show that there are no early-offer effects. Section D.3 provides an example in which the early-offer effect on offer acceptance is negative.

Section D.4 proves that the sequence of optimal strategies under the M-DA can be summarized into one strategy. Thus, it provides a theoretical foundation for Section IV.B.2 in the main text. Section D.5 proves Proposition 1.

## D.1 The Gale-Shapley Algorithm

We distinguish between a "mechanism" and an "algorithm," although the literature often uses them interchangeably. Let us consider the case that students apply to university programs and that each program admits students. In our setting, a university-admissions mechanism, such as the M-DA, is a complete procedure that specifies how students and programs exchange information with the mechanism and how a matching outcome is determined. In contrast, an algorithm is a computer code that takes as inputs information from applicants and programs and delivers a matching outcome; importantly, it is silent on how relevant information is collected. Therefore, an algorithm is always one of the multiple components of a mechanism.

The Gale-Shapley algorithm can be either student-proposing or program-proposing. We focus on the latter. Specifically, taking as inputs program capacities, the programs' ranking over applicants, and students' rank-order lists of programs (ROLs), it proceeds as follows:

*Round* 1. Every program with capacity $q_j$ and $n_j^1$ applicants extends an admission offer to its top-min$\{q_j, n_j^1\}$ applicants in its ranking over applicants. Each student who receives multiple offers keeps the highest-ranked offer according to her ROL and rejects the rest.

Generally, in:

*Round* $(r > 1)$. Every program with $m_j^{(r-1)}$ of its previous offers rejected in round $r-1$ and $n_j^r$ applicants who have never received its offer extends offers to the top-min$\{m_j^{(r-1)}, n_j^r\}$ applicants among those who have not received its offer. Each student who has multiple offers keeps the highest ranked offer according to her ROL and rejects the rest.

The process terminates after any round $r$ in which no offers are rejected by students. Each program is then matched with the students that are currently holding its offer.

## D.2 Ordinally Informed Student

We now investigate if the early-offer effect in Section II can emerge when we impose the following assumption that is common in the matching literature: agents know their own ordinal preferences (Roth and Sotomayor, 1990; Bogomolnaia and Moulin, 2001).

We call the student *ordinally informed* if the distribution of $X_j$, $G_j$, is such that with probability one, (i) either $u(X_j) > u(X_{j'})$ or $u(X_j) < u(X_{j'})$ for any $j \neq j'$, and (ii) either $u(X_j) > 0$ or $u(X_j) < 0$ for any $j$. In other words, the student knows her ordinal preferences and whether a program is (ex post) acceptable without any further learning. Importantly, this assumption includes the following as a special case: there is no uncertainty about program quality and there are no ties in the student's cardinal preferences.

**Proposition D1.** *If the student is ordinally informed, she does not learn about program quality and submits the same ROL for all $O$, $p^0$, and $c$ under the DA, the M-DA, or the BM-DA. Her expected utility is constant across the mechanisms.*

This proposition is straightforward, and we therefore only sketch the proof here. Since the optimal strategy for the student is to submit an ROL respecting her true ordinal preferences, which are known to her without additional learning, the marginal benefit of learning more about the programs is zero under any mechanism. Since her submitted ROL is constant across mechanisms, her expected utility is constant.

Proposition D1 also leads to the following corollary.

**Corollary D1.** *If the student is ordinally informed, for all $O$, $p^0$, and $c$, there is no early-offer effect on offer acceptance under the M-DA.*

In contrast to Corollary D1, our empirical study documents a sizable early-offer effect on offer acceptance; it therefore implies that students being ordinally informed may not be plausible in our empirical setting.

## D.3 Ambiguous Effects of Early Offers on Offer Acceptance

Our empirical analysis finds that an early offer increases the probability that the offer is accepted. This result does not hold in some cases. Below, we provide an example in which an early offer from a program reduces the likelihood that the offer is accepted on average; moreover, there is no first-early-offer effect. One of the key features of this example is that the student never has any incentives to learn about one of the two programs.

Suppose that the student is risk neutral and applies to two programs, $j = 1, 2$. The offer probabilities (as seen in period 0) are $p_1^0$ and $p_2^0$, while $0 < p_1^0 < p_2^0 < 1$. The distribution of $X_1$ is *Uniform*$(\mu_1 - \delta, \mu_1 + \delta)$ with $\mu_1 \in (1/2, 1)$; importantly, $0 < \delta < c/2$ and thus it never pays to learn about $X_1$. Moreover, $X_2 \sim$ *Uniform*$(0, 1)$, and there is an outside option whose value is zero.

When there is no learning, the student submits ROL 1–2 (the first choice is program 1 and the second choice is program 2) with an expected payoff $V_0(p_1, p_2) = p_1\mu_1 + (1-p_1)p_2/2$, for $p_1 \in \{p_1^0, 1\}$ and $p_2 \in \{p_2^0, 1\}$. If the student learns $X_2$, the expected payoff is:

$$\begin{aligned} V_1(p_1, p_2) = &\int_{\mu_1}^{1} (p_2 x_2 + (1-p_2)p_1\mu_1)dx_2 + \int_0^{\mu_1} (p_1\mu_1 + (1-p_1)p_2 x_2)dx_2 - c \\ =&V_0(p_1, p_2) + \frac{p_1 p_2}{2}(1-\mu_1)^2 - c. \end{aligned}$$

Suppose that $c$ is such that $\frac{p_1^0}{2}(1-\mu_1)^2 < c < \frac{p_2^0}{2}(1-\mu_1)^2$. Therefore, when there is no early offer or only one early offer from program 2, the student will not learn about $X_2$ and thus will always submit ROL 1–2.

When the student receives only one early offer from program 1, she will learn about $X_2$ and ex ante (before learning) submit ROL 1–2 with probability $\mu_1$. When there are two early offers, regardless of the arrival order, the student's learning and ranking behaviors are the same as when she receives an early offer from program 1.

We calculate the early-offer effects as follows (similar to Result 2):

$$\Pr(\text{top rank univ. 1} \mid \text{early offer from 1}) - \Pr(\text{top rank univ. 1} \mid \text{no early offer}) = \mu_1 - 1,$$
$$\Pr(\text{top rank univ. 1} \mid \text{two early offers}) - \Pr(\text{top rank univ. 1} \mid \text{early offer from 2}) = \mu_1 - 1,$$
$$\Pr(\text{top rank univ. 2} \mid \text{early offer from 2}) - \Pr(\text{top rank univ. 2} \mid \text{no early offer}) = 0,$$
$$\Pr(\text{top rank univ. 2} \mid \text{two early offers}) - \Pr(\text{top rank univ. 2} \mid \text{early offer from 1}) = 0.$$

If early offers arrive randomly, the average early-offer effect on offer acceptance is a weighted average of the above four effects, while each weight is strictly positive. Thus, it leads to a negative average effect.

We also calculate the first-early-offer effects as follows (similar to Result 3):

$$\Pr(\text{top rank univ. 1} \mid \text{offers from 1 then 2}) - \Pr(\text{top rank univ. 1} \mid \text{offers from 2 then 1}) = 0,$$
$$\Pr(\text{top rank univ. 2} \mid \text{offers from 2 then 1}) - \Pr(\text{top rank univ. 1} \mid \text{offers from 1 then 2}) = 0.$$

Therefore, there is no first-early-offer effect.

## D.4 Optimal Learning Strategy under the M-DA

Recall that $\underline{t} = \min\{\min\{O\}, J\}$. In each period $t = \underline{t}, \ldots, J$, given $p^t(O)$, the student has a myopic strategy, $\psi^t(\cdot, \cdot \mid p^t(O))$, leading to (subjective) expected utility $V^t(\mathcal{U}^{(t-1)}, l^{(t-1)}, p^t(O) \mid \psi^t(\cdot, \cdot \mid p^t(O)))$, where $\mathcal{U}^{(t-1)}$ and $l^{(t-1)}$ are the learning outcomes at the end of period $(t-1)$ with $\mathcal{U}^{(\underline{t}-1)} = \mathcal{J}$ and $l^{(\underline{t}-1)} = \varnothing$. Hence, $\psi^t \in \arg\max_\psi V^t(\mathcal{U}^{(t-1)}, l^{(t-1)}, p^t(O) \mid \psi)$, for $t = \underline{t}, \ldots, J$. Below, we show that this sequence of strategies is equivalent to one single strategy.

Suppose that the student adopts the sequence of strategies $\{\psi^t\}_{t=\underline{t}}^J$. In a given period $t$, only $\psi^t$ is applied, and at the end of the period, the student must reach a state, $(\mathcal{U}, l)$, such that $\psi^t(\mathcal{U}, l \mid p^t(O)) = 0$ (i.e., the student stops learning).

We say that a state $(\mathcal{U}, l)$ that is not $(\mathcal{J}, \varnothing)$ is *never reached* in period $\underline{t}$ if there is no learning sequence, $\{j_1, ..., j_n\} = \overline{\mathcal{U}} = \mathcal{J} \backslash \mathcal{U}$ and $\{(j_1, x_{j_1}), \ldots, (j_n, x_{j_n})\} = l$, such that $\psi^{\underline{t}}(\mathcal{J}, \varnothing \mid p^t(O)) = j_1$ and $\psi^{\underline{t}}(\mathcal{J} \backslash \{j_1, \ldots, j_m\}, \{(j_1, x_{j_1}), \ldots, (j_m, x_{j_m})\}) = j_{m+1}$ for $m = 1, \ldots, n-1$. Let $\mathcal{NR}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$ be the collection of states that are never reached in period $\underline{t}$ given $\{\psi^t\}_{t=\underline{t}}^J$. We can sequentially define $\mathcal{NR}(t, \{\psi^t\}_{t=\underline{t}}^J)$ for $t = \underline{t}+1, \ldots, J$.

For a state that is never reached in a period, a strategy can prescribe an arbitrary action without affecting the student's payoff. We impose a no-learning assumption on such off-equilibrium paths:

**Assumption D1.** *The student's strategy does not require her to learn anything in a never reached state, i.e.,* $\psi^t(\mathcal{U}, l \mid p^t(O)) = 0$ *if* $(\mathcal{U}, l) \in \mathcal{NR}(t, \{\psi^t\}_{t=\underline{t}}^J)$.

Further, we let $\mathcal{T}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$ be the collection of terminal states that can be reached *at the end* of period $\underline{t}$, such that $\psi^{\underline{t}}(\mathcal{U}, l \mid p^{\underline{t}}(O)) = 0$, $\forall (\mathcal{U}, l) \in \mathcal{T}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$. It must be that $\mathcal{T}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J) \cap \mathcal{NR}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J) = \varnothing$.

We can sequentially define $\mathcal{T}(t, \{\psi^t\}_{t=\underline{t}}^J)$ for $t = \underline{t}+1, \ldots, J$ by noting that a possible initial state in period $t$ must be in $\mathcal{T}(t-1, \{\psi^t\}_{t=\underline{t}}^J)$.

**Lemma D1.** *Under Assumption D1, for any* $(\mathcal{U}, l)$*, if* $\psi^t(\mathcal{U}, l \mid p^t(O)) \neq 0$ *for some* $t \geqslant \underline{t}$*, then* $\psi^{t'}(\mathcal{U}, l \mid p^{t'}(O)) = 0$ *for all* $t' \in \{\underline{t}, \ldots, J\} \backslash \{t\}$.

*Proof.* Suppose that $\psi^t(\mathcal{U}, l \mid p^t(O)) \neq 0$ for some $t \geqslant \underline{t}$.

First, we consider $t = \underline{t}$. With the off-equilibrium restriction in Assumption D1 and $\psi^t(\mathcal{U}, l \mid p^t(O))$ being a pure strategy, $\psi^t(\mathcal{U}, l \mid p^t(O)) \neq 0$ implies that there exists a unique learning sequence, $\{j_1, ..., j_n\} = \overline{\mathcal{U}} = \mathcal{J} \backslash \mathcal{U}$ and $\{(j_1, x_{j_1}), \ldots, (j_n, x_{j_n})\} = l$, such that $\psi^{\underline{t}}(\mathcal{J}, \varnothing \mid p^t(O)) = j_1$ and $\psi^{\underline{t}}(\mathcal{J} \backslash \{j_1, \ldots, j_m\}, \{(j_1, x_{j_1}), \ldots, (j_m, x_{j_m})\}) = j_{m+1}$ for $m = 1, \ldots, n-1$. Uniqueness is because we consider pure strategies. Moreover, $(\mathcal{U}, l) \notin \mathcal{T}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$; that is, $(\mathcal{U}, l)$ is not a possible terminal state in period $\underline{t}$. This means that $(\mathcal{U}, l) \in \mathcal{NR}(t', \{\psi^t\}_{t=\underline{t}}^J)$ for $t' > \underline{t}$ and thus $\psi^{t'}(\mathcal{U}, l \mid p^{t'}(O)) = 0$ (by the off-equilibrium restriction).

Second, a similar argument can be made for $t = \underline{t}+1$. We note the following observation: $\psi^t(\mathcal{U}, l \mid p^t(O)) \neq 0$ implies that there exists a unique state $(\mathcal{U}^{(t-1)}, l^{(t-1)}) \in \mathcal{T}(t-1, \{\psi^t\}_{t=\underline{t}}^J)$ and a unique learning sequence $\{j_1, ..., j_n\} = \mathcal{U}^{(t-1)} \backslash \mathcal{U}$ and $\{(j_1, x_{j_1}), \ldots, (j_n, x_{j_n})\} = l \backslash l^{(t-1)}$ such that $\psi^t(\mathcal{U}^{(t-1)}, l^{(t-1)} \mid p^t(O)) = j_1$ and, for $m = 1, \ldots, n-1$,

$$\psi^t\left(\mathcal{U}^{(t-1)} \backslash \{j_1, \ldots, j_m\}, l^{(t-1)} \cup \{(j_1, x_{j_1}), \ldots, (j_m, x_{j_m})\}\right) = j_{m+1}.$$

Therefore, $(\mathcal{U}, l)$ is either in $\mathcal{NR}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$ or $\mathcal{T}(\underline{t}, \{\psi^t\}_{t=\underline{t}}^J)$, and thus $\psi^{\underline{t}}(\mathcal{U}, l \mid p^{\underline{t}}(O)) = 0$. Moreover, $(\mathcal{U}, l) \in \mathcal{NR}(t', \{\psi^t\}_{t=\underline{t}}^J)$ for $t' > t$ and thus $\psi^{t'}(\mathcal{U}, l \mid p^{t'}(O)) = 0$.

By continuing this argument, we can show that if $\psi^t(\mathcal{U}, l \mid p^t(O)) \neq 0$ for $t \geqslant \underline{t}$, then $\psi^{t'}(\mathcal{U}, l \mid p^{t'}(O)) = 0$ for all $t' \in \{\underline{t}, \ldots, J\} \backslash \{t\}$. ∎

By Lemma D1, we can define a single strategy that is equivalent to applying $\{\psi^t\}_{t=\underline{t}}^J$ sequentially, as we do in Section IV.B.2.

## D.5 Proof of Proposition 1

*Proof.* The weak inequalities in part (i) are implied by the fact that $\psi^B(\cdot, \cdot \mid p^J(O)) \in \arg\max_\psi V(\mathcal{J}, \varnothing, p^J(O) \mid \psi)$ under the BM-DA, while $\psi^{DA}(\cdot, \cdot \mid p^0)$ and $\psi^M(\cdot, \cdot \mid p^J(O))$ may not maximize $V(\mathcal{J}, \varnothing, p^J(O) \mid \psi)$.

Furthermore, Result 4 in Section IV.C gives an example of $(\mathcal{J}, p^0, O, F, c)$ such that the dominance of the BM-DA is strict.

For part (ii), we notice that the BM-DA and the M-DA are equivalent when there is only one early offer (i.e., there is a unique program $j$ such that $O_j < J+1$ and $O_{j'} = J+1$ if $j' \neq j$). In this case, the M-DA achieves a higher expected utility than the DA, as implied by part (i). The opposite can be true in cases where the M-DA is not equivalent to the BM-DA. Result 4 in Section IV.C gives a concrete example. ∎

# E Simulation-Based Comparison of Mechanisms

This appendix describes the simulations we carry out to compare the welfare properties of the three mechanisms studied in Section IV in the main text: the DA, the M-DA, and the BM-DA. In keeping with our theoretical model whenever possible, we construct a stylized, closed market in which student behavior can be simulated under each mechanism using our empirical estimates based on the DoSV data, which makes it possible to compare the welfare properties of these mechanisms.

## E.1 Setup

We use the same data set as for the main empirical analysis, namely the data from the DoSV procedure for 2015–16, to construct a market in which students are matched with university programs under the DA, M-DA, and BM-DA mechanisms.

### E.1.1 The Market

**Students.** As in the main empirical analysis, we use the set of 21,711 students who applied to at least two feasible programs and accepted an offer.

**Students' applications and programs.** Throughout the simulations, we keep fixed the set of programs that each student $i$ applies to, which we denote by $\mathcal{A}_i$. This set includes an outside option as well as all programs that are in the student's initial rank-order list (ROL) in the DoSV procedure for 2015–16. Introducing this outside option accounts for the possibility that participants may have applied to programs outside the platform and devoted some time learning about them. Since all students in our sample accepted an offer from the platform, we make the simplifying assumption that this outside option is never feasible. We denote by $A_i \equiv |\mathcal{A}_i|$ the number of programs in $\mathcal{A}_i$ including the outside option.

The union of $\mathcal{A}_i$ across all students in the simulation, $\bigcup_i \mathcal{A}_i$, is the set of programs in the simulated market. In total, there are 376 programs.

**Program capacities.** For each program $j$, the number of available seats, denoted by $q_j$, is set equal to the number of students in the simulation sample who accepted an offer from the program in reality.

**Programs' ranking over students.** To simplify the analysis, we depart from the German setting by imposing that each program ranks its applicants under a single ranking (instead of the multiple-quota system). We generate the programs' rankings on the basis of a student-program-specific priority score, denoted $score_{i,j}$ (higher is better), which is the average of the student's *Abitur* percentile rank and a program-specific random component:

$$score_{i,j} = \frac{Abitur_i + \nu_{i,j}}{2}, \quad \forall i, j, \tag{A.1}$$

where $Abitur_i$ represents student $i$'s *Abitur* percentile rank (between zero and one) and $\nu_{i,j} \sim Uniform(0,1)$.

To ensure that our analyses are performed on a subset of programs for which student preferences can be estimated, while allowing for some variation in feasible sets across simulations, we put some restrictions on the programs that can ever be feasible to a student. We define an *extended feasible set* for each student. It includes the programs in $\mathcal{A}_i$ that were ex post feasible to the student in reality and the infeasible program in $\mathcal{A}_i$ (if any) that was the closest to being feasible under the most favorable quota to the student. In the simulations, a student who applies to programs outside of this extended feasible set is considered unacceptable to the corresponding programs (i.e., she never receives offers from those programs).

### E.1.2 Timeline under the Three Mechanisms

We assume that the following components are constant across the three mechanisms: (i) each student always applies to the same subset of programs in $\mathcal{A}_i$; (ii) the programs' capacities; and (iii) the programs' rankings over students.

As described in Section IV, the three mechanisms differ in terms of the existence and timing of early offers as well as the timing for students to submit their ROL.

**DA.** Students submit their ROL without having received early offers. The matching is determined by the program-proposing Gale-Shapley (GS) algorithm, using as inputs the students' submitted ROLs, the programs' rankings of students, and the programs' capacities.

**M-DA.** Each program sends out a single batch of early offers to its highest-ranked applicants up to its capacity. We assume that early offers are sent out on different dates and that the order of offer arrival is random for every student. Students are required to submit an ROL of programs after all early offers have been sent out. The matching is then determined by the program-proposing GS algorithm.

**BM-DA.** The timing is the same as under the M-DA mechanism, except that all early offers are sent out to students on the same date.

## E.2 Learning, Rank-Order Lists, and Matching Outcome

Students' preferences, their learning behavior under the three mechanisms, and the determination of their submitted ROL and matching outcome are simulated using a model whose parameters are estimated based on the DoSV data.

### E.2.1 Utility under Full Information and No Information

As in Section IV, a student's preferences over programs are unknown and can only be learned at a cost. Student $i$'s true utility from program $j$ (i.e., conditional on having learned her preferences for this program) is denoted by $U_{i,j}^{\text{FullInfo}}$, while $U_{i,j}^{\text{NoInfo}}$ denotes her expected—or perceived—utility without learning.

**Utility under full information.** Student $i$'s utility from program $j$ under full information, $U_{i,j}^{\text{FullInfo}}$, takes the following form:

$$U_{i,j}^{\text{FullInfo}} = V_{i,j}^{\text{FullInfo}} + \epsilon_{i,j}^{\text{FullInfo}}, \quad \forall i, j \in \mathcal{A}_i, \tag{A.2}$$

where $V_{i,j}^{\text{FullInfo}}$ represents the deterministic component of the student's utility, which depends on observable student-program-specific characteristics (e.g., field of study and distance), and $\epsilon_{i,j}^{\text{FullInfo}}$ represents the (random) idiosyncratic component, which is unobserved and i.i.d. type I extreme value (Gumbel) distributed.

To quantify $V_{i,j}^{\text{FullInfo}}$, we rely on the same sample as in the main empirical analysis of the early-offer effect, i.e., those students who applied to at least two feasible programs and accepted an offer. Under the assumption that a student always learns her preference for the program from which she has received her first early offer, $V_{i,j}^{\text{FullInfo}}$ is calculated by assuming that $i$ receives her first early offer from program $j$.

We use students' final ROLs. After restricting each student's choice set to the ex post feasible programs that she included in her initial ROL, we estimate the following specification using a rank-ordered logit to extract information from students' final ROLs as described in Section II.C:

$$\begin{aligned} U_{i,j} &= V_{i,j}(\boldsymbol{X}_{i,j}, \boldsymbol{W}_{i,j}, EO_{i,j}, FEO_{i,j}) + \eta_{i,j} \\ &= \boldsymbol{X}_{i,j}\beta + \delta_1\, EO_{i,j} + \delta_2\, FEO_{i,j} + (EO_{i,j} \cdot \boldsymbol{W}_{i,j})\gamma_1 + (FEO_{i,j} \cdot \boldsymbol{W}_{i,j})\gamma_2 + \eta_{i,j}, \quad \forall i, j, \end{aligned} \tag{A.3}$$

where $\boldsymbol{X}_{i,j}$ and $\boldsymbol{W}_{i,j}$ are row vectors of student-program-specific characteristics; $\boldsymbol{X}_{i,j}$ includes program fixed effects, distance, distance squared, and a dummy for whether the program is in the student's region (*Land*); $\boldsymbol{W}_{i,j}$ includes university fixed effects, field-of-study fixed effects,[3] distance, distance squared, and a dummy for whether the program is in the student's region; $EO_{i,j}$ is an indicator for whether student $i$ has received an early offer from program $j$; $FEO_{i,j}$ is an indicator for whether the first early offer received by student $i$ was from program $j$; and $\eta_{i,j}$ is a type I extreme value.

In this specification, the coefficients $\gamma_1$ and $\gamma_2$ on the interaction terms between the indicators for early offer/first early offer and the student-program-specific characteristics $\boldsymbol{W}_{i,j}$ capture indirectly the learning effects induced by early offers. They measure how early offers modify the weights that students place on the observable characteristics of the programs from which they received such offers.

We then compute $V_{i,j}^{\text{FullInfo}}$ as follows:

$$\begin{aligned} V_{i,j}^{\text{FullInfo}} &= \widehat{V}_{i,j}(\boldsymbol{X}_{i,j}, \boldsymbol{W}_{i,j}, EO_{i,j} = 1, FEO_{i,j} = 1) \\ &= \boldsymbol{Z}_{i,j}\widehat{\beta} + \widehat{\delta}_1 + \widehat{\delta}_2 + \boldsymbol{W}_{i,j}(\widehat{\gamma}_1 + \widehat{\gamma}_2), \quad \forall i, j, \end{aligned} \tag{A.4}$$

where $(\widehat{\beta}, \widehat{\delta}_1, \widehat{\delta}_2, \widehat{\gamma}_1, \widehat{\gamma}_2)$ are the parameter estimates from Equation (A.3).

**Utility under no information.** Similar to $U_{i,j}^{\text{FullInfo}}$, we assume that student $i$'s utility from program $j$ without learning, $U_{i,j}^{\text{NoInfo}}$, takes the following form:

$$U_{i,j}^{\text{NoInfo}} = V_{i,j}^{\text{NoInfo}} + \epsilon_{i,j}^{\text{NoInfo}}, \quad \forall i, j \in \mathcal{A}_i, \tag{A.5}$$

where $V_{i,j}^{\text{NoInfo}}$ and $\epsilon_{i,j}^{\text{NoInfo}}$ are the deterministic and idiosyncratic components, respectively; $\epsilon_{i,j}^{\text{NoInfo}}$ is assumed to be i.i.d. type I extreme value distributed and $\epsilon_{i,j}^{\text{NoInfo}} \perp \epsilon_{i,j}^{\text{FullInfo}}$.

We further assume that $V_{i,j}^{\text{NoInfo}}$ is drawn from a normal distribution centered at

[3] The programs are grouped into 12 fields of study (architecture and design, business and economics, engineering, language and culture, law, mathematics and computer science, medicine, natural sciences, psychology, social sciences, social work, and teaching programs) and a residual group for other fields.

$V_{i,j}^{\text{FullInfo}}$:

$$V_{i,j}^{\text{NoInfo}} \sim N\big(V_{i,j}^{\text{FullInfo}}, \big(\text{s.e.}(V_{i,j}^{\text{FullInfo}})\big)^2\big), \quad \forall i, j \in \mathcal{A}_i, \tag{A.6}$$

where s.e.$(V_{i,j}^{\text{FullInfo}})$ is the standard error of the predicted value in Equation (A.4).

### E.2.2 Preference Discovery under the Three Mechanisms

As in Section IV, for each mechanism, we assume that the learning technology is such that a student either learns her true utility from program $j$, $U_{i,j}^{\text{FullInfo}}$, or learns nothing beyond $U_{i,j}^{\text{NoInfo}}$. We denote by $\lambda_{i,j}^m$ an indicator that takes the value of one if student $i$ learns her true utility from program $j$ under mechanism $m$, and zero otherwise.

**Learning costs.** We do not have an estimate of learning costs. Therefore, we impose the simplifying assumption that under any mechanism, a student learns her true preferences for half of the programs in $\mathcal{A}_i$ (which may include learning the outside option). While this assumption neglects any potential effects of a matching mechanism on the amount of learning, it allows us to ignore the learning costs when comparing welfare between mechanisms.

**Learning under the DA.** Under the DA mechanism, we assume that each student learns her true preferences for a random half (rounded up to the next lower integer) of the programs to which she has applied. Let $\omega_i^{\text{DA}} : \mathcal{A}_i \to \{1, 2, ..., A_i\}$ denote a function such that $\omega_i^{\text{DA}}(j)$ returns the order of program $j$ at which it might be learned by student $i$. In the simulations, $\omega_i^{\text{DA}}(j)$ is chosen randomly. Student $i$'s learning outcome for program $j$ under the DA is:

$$\lambda_{i,j}^{\text{DA}} = \begin{cases} 1 & \text{if } \omega_i^{\text{DA}}(j) \leqslant \left\lfloor \frac{A_i}{2} \right\rfloor \\ 0 & \text{if } \omega_i^{\text{DA}}(j) > \left\lfloor \frac{A_i}{2} \right\rfloor \end{cases}, \quad \forall i, j \in \mathcal{A}_i, \tag{A.7}$$

where $\lfloor \cdot \rfloor$ represents the floor function.

**Learning under the M-DA.** Under the M-DA mechanism, a student may receive early offers at different dates before submitting her ROL. As in our theoretical model, an early offer may change a student's learning behavior. Compared to the DA, students' learning under the M-DA is modified by taking into account early offers and the order in which they are received.

Specifically, we assume that a student always learns her first early offer and then alternates between (i) learning a randomly chosen program from the ones in $\mathcal{A}_i$ she has not learned yet (including those from which she may later receive an early offer) and (ii) learning her early offers (if any) in the order in which they arrive. Similar to Section IV, the underlying assumption is that under the M-DA, a student's learning decision is made "myopically" period by period: each time a student receives an early offer, she learns her utility from this program (unless she has already learned her true preferences for half of the programs in $\mathcal{A}_i$); during the time interval between two consecutive early offers (or if the students has already learned her preferences for all early offers), she learns at random one of the programs that have not yet been learned.

Define $e_i : \{1, 2, \ldots, A_i\} \to \mathcal{A}_i \bigcup \varnothing$ such that student $i$'s $j^{\text{th}}$ early offer is from program $e_i(j)$ if $e_i(j) \neq \varnothing$ and such that $i$ has no more than $j-1$ early offers if $e_i(j) = \varnothing$.

Further, we define $\omega_i^{\text{M-DA}} : \mathcal{A}_i \to \{1, 2, ..., A_i\}$ such that $\omega_i^{\text{M-DA}}(j)$ is the (potential) learning order of program $j$ under the M-DA. Specifically, if the student does not receive any early offers, $\omega_i^{\text{M-DA}} = \omega_i^{\text{DA}}$. If the student receives one or more early offers, $\omega_i^{\text{M-DA}}$ is constructed in $A_i$ steps as follows:

(1) We define $L$ as the latest early offer and set $L = 1$. Let $\omega_i^{\text{M-DA}}(e_i(L)) = 1$, i.e., the first early offer is learned first.

($l$) ($2 \leqslant l \leqslant A_i$) There are two different cases:

  (a) $\omega_i^{\text{M-DA}}(e_i(L)) = l-1$; i.e., the latest early offer, $e_i(L)$, was chosen to be learned in step $l-1$ because it was the latest early offer then. In this case, we let $\omega_i^{\text{M-DA}}(j) = l$ where $j = \arg\min_{j' \in \mathcal{A}_i:\ \omega_i^{\text{M-DA}}(j') \notin \{1,...,l-1\}} \omega_i^{\text{DA}}(j')$. That is, the student learns in step $l$ the earliest program, as determined by $\omega_i^{\text{DA}}$, among those that have not been learned.

  (b) $\omega_i^{\text{M-DA}}(e_i(L)) < l-1$; i.e., the latest early offer, $e_i(L)$, was chosen to be learned in a step earlier than $l-1$ (or, equivalently, the program learned in step $l-1$ was not an early offer then). Let $L = L+1$, i.e., the next early offer becomes the latest early offer. If $e_i(L) \neq \varnothing$ and $\omega_i^{\text{M-DA}}(e_i(L)) \notin \{1, \ldots, l-1\}$, we let $\omega_i^{\text{M-DA}}(e_i(L)) = l$; otherwise, $\omega_i^{\text{M-DA}}(j) = l$ where $j = \arg\min_{j' \in \mathcal{A}_i:\ \omega_i^{\text{M-DA}}(j') \notin \{1,...,l-1\}} \omega_i^{\text{DA}}(j')$. That is, the student learns either the latest early offer (if any and if it has not been learned) or the earliest program, as determined by $\omega_i^{\text{DA}}$, among those that have not been learned.

Student $i$'s learning outcome for program $j$ under the M-DA is then defined as

$$\lambda_{i,j}^{\text{M-DA}} = \begin{cases} 1 & \text{if } \omega_i^{\text{M-DA}}(j) \leqslant \left\lfloor \frac{A_i}{2} \right\rfloor \\ 0 & \text{if } \omega_i^{\text{M-DA}}(j) > \left\lfloor \frac{A_i}{2} \right\rfloor \end{cases}, \quad \forall i, j \in \mathcal{A}_i. \tag{A.8}$$

By construction, if a student does not receive early offers, her learning outcomes under the M-DA are the same as under the DA, i.e., $\lambda_{i,j}^{\text{M-DA}} = \lambda_{i,j}^{\text{DA}}$ for all $j \in \mathcal{A}_i$. We maintain such a correlation between $\lambda_{i,j}^{\text{M-DA}}$ and $\lambda_{i,j}^{\text{DA}}$ (or, equivalently, between $\omega_i^{\text{M-DA}}$ and $\omega_i^{\text{DA}}$) so that the differences between the two mechanisms are driven only by the arrival of early offers.

*Example:* Suppose that $\mathcal{A}_i = \{j_1, j_2, j_3, j_4\}$ and that student $i$'s potential learning sequence under the DA is $(j_1, j_2, j_3, j_4)$. If the student receives early offers from three of these programs in the order $(j_4, j_2, j_1)$, the learning sequence under the M-DA is $(j_4, j_1, j_2, j_3)$, implying that the student first learns $j_4$ and then $j_1$.[4] If instead the arrival order of the early offers is $(j_2, j_1, j_4)$, the learning sequence under the M-DA is $(j_2, j_1, j_3, j_4)$, so the student first learns $j_2$ and then $j_1$.[5]

[4]The learning sequence under the M-DA is determined as follows: (i) the first program in the sequence is the student's first early offer, i.e., $j_4$; (ii) the next program is the first one in the learning sequence under the DA that has not yet been learned, i.e., $j_1$; (iii) then comes the second early offer ($j_2$), as it has not been learned in the previous step; (iv) the last program is the one in the learning sequence under the DA that has not yet been learned, i.e., $j_3$.

[5]The learning sequence under the M-DA is determined as follows: (i) the first program in the sequence is the student's first early offer, i.e., $j_2$; (ii) the next program is the first one in the learning sequence under the DA that has not yet been learned, i.e., $j_1$; (iii) since the second early offer ($j_1$) has been learned in the previous step, the next program to be learned is the first program in the learning sequence under the DA that has not yet been learned, i.e., $j_3$; (iv) the last program is the next one in the learning sequence under the DA that has not yet been learned, i.e., $j_4$.

**Learning under the BM-DA.** Under the BM-DA mechanism, each student receives her early offers on a single date before submitting her ROL. In contrast to the M-DA, we assume that a student always learns her early offers before learning other programs, up to the point where she has learned half of the programs to which she has applied.

Define $\omega_i^{\text{BM-DA}} : \mathcal{A}_i \rightarrow \{1, 2, ..., A_i\}$ such that $\omega_i^{\text{BM-DA}}(j)$ returns the (potential) learning order of program $j$ under the BM-DA mechanism. If the student does not receive early offers, we assume that the learning order is the same as under the DA and the M-DA, i.e., $\omega_i^{\text{BM-DA}} = \omega_i^{\text{M-DA}} = \omega_i^{\text{DA}}$. If instead the student receives one or more early offers, we make the following assumptions: (i) programs that made an early offer to the student are learned before programs that did not; (ii) the relative learning order of early offers is given by $\omega_i^{\text{DA}}$; and (iii) programs that did not extend an early offer to the student are learned in the same relative order as under the DA (as given by $\omega_i^{\text{DA}}$).

Under the BM-DA mechanism, student $i$'s learning outcome for program $j$ is then defined as

$$\lambda_{i,j}^{\text{BM-DA}} = \begin{cases} 1 & \text{if } \omega_i^{\text{BM-DA}}(j) \leqslant \left\lfloor \frac{A_i}{2} \right\rfloor \\ 0 & \text{if } \omega_i^{\text{BM-DA}}(j) > \left\lfloor \frac{A_i}{2} \right\rfloor \end{cases}, \quad \forall i, j \in \mathcal{A}_i. \tag{A.9}$$

If a student does not receive early offers, her learning outcomes under the BM-DA mechanism are the same as under the DA and the M-DA, i.e., $\lambda_{i,j}^{\text{BM-DA}} = \lambda_{i,j}^{\text{DA}} = \lambda_{i,j}^{\text{M-DA}}$ for all $j \in \mathcal{A}_i$. Again, we maintain such correlations among $\lambda_{i,j}^{\text{BM-DA}}$, $\lambda_{i,j}^{\text{M-DA}}$, and $\lambda_{i,j}^{\text{DA}}$ (or, equivalently, among $\omega_i^{\text{BM-DA}}$, $\omega_i^{\text{M-DA}}$, and $\omega_i^{\text{DA}}$) so that the differences between any two mechanisms are driven only by the arrival of early offers.

### E.2.3 Determination of Submitted ROL and Matching Outcome

**Perceived utility.** At the time of submitting her final ROL (i.e., conditional on all her information at that time), student $i$'s perceived utility from program $j$ under mechanism $m$, $U_{i,j}^m$, depends on whether or not she has learned her preferences for that program:

$$U_{i,j}^m = \lambda_{i,j}^m \cdot U_{i,j}^{\text{FullInfo}} + (1 - \lambda_{i,j}^m) U_{i,j}^{\text{NoInfo}}, \quad \forall i, j \in \mathcal{A}_i. \tag{A.10}$$

**Submitted ROL.** Each student is assumed to submit a complete and truthful (w.r.t. $U_{i,j}^m$) ranking of the programs in $\mathcal{A}_i$.

**Matching outcome.** For each mechanism, the program-proposing GS algorithm is used to match the students and programs. We assume that after the matching takes place, students always experience their true preference for the program to which they have been matched. A student's utility of her matching outcome, $\mu(i)$, can therefore be evaluated at $U_{i,\mu(i)}^{\text{FullInfo}}$.

## E.3 Monte Carlo Simulations

The simulations are performed among the same $S$ Monte Carlo samples under each of the three mechanisms. We set $S$ =10,000.

### E.3.1 Components Fixed across Simulation Samples

Across the simulation samples, the following components are held fixed as specified in Section E.1.1:

**Market participants.** Student characteristics and program attributes are fixed. The set of students is $\mathcal{I} \equiv \{1, ..., I\}$ while the set of programs is $\mathcal{J} \equiv \{1, \ldots, J\}$. In the simulations, $I$ =21,711 and $J$=376.

**Student Applications.** Each student $i$ applies to all programs in $\mathcal{A}_i$. On average, students in the simulation sample apply to 5.7 programs (including the outside option).

**Program capacities.** The programs' capacities are $\{q_j\}_{j=1}^{J}$.

**Programs' rankings of students.** Each program ranks its applicants based on the student-program specific score defined by Equation (A.1). If the program does not belong to the student's extended feasible set as defined in Section E.1.1, the student is assumed to be unacceptable to the program.

**Early offers.** Under the M-DA and BM-DA mechanisms, early offers are made to each program's top-ranked applicants up to the program's capacity. The set of early offers is the same under both mechanisms.

**Utility under full information.** For each student $i$ and program $j \in \mathcal{A}_i$, the deterministic component of the student's utility from the program under full information, $V_{i,j}^{\text{FullInfo}}$, is calculated using Equation (A.4) and is stored together with the standard error of the prediction, s.e.$(V_{i,j}^{\text{FullInfo}})$.

### E.3.2 Simulation Steps

Other than those in Section E.3.1, a component in general is independently drawn in each simulation sample. This includes idiosyncratic utility shocks, utility without learning, early offer arrival orders, and potential learning orders. Note that in a given simulation sample, we have the same market for each of the three mechanisms.

The $S$ independent Monte Carlo samples are generated as follows:

**Step 1: Utility functions with and without learning.** Let $U_{i,j,s}^{\text{FullInfo}}$ denote student $i$'s utility from program $j$ in sample $s$ under full information and $U_{i,j,s}^{\text{NoInfo}}$ her utility without learning. For each student $i$ in sample $s$:

(i) Draw a set of i.i.d. type I extreme values $\epsilon_{i,j,s}^{\text{FullInfo}}$ and $\epsilon_{i,j,s}^{\text{NoInfo}}$ for all $j \in \mathcal{A}_i$.

(ii) Use Equation (A.2) to compute the student's true utility from program $j$ in sample $s$, $U_{i,j,s}^{\text{FullInfo}}$:

$$U_{i,j,s}^{\text{FullInfo}} = V_{i,j}^{\text{FullInfo}} + \epsilon_{i,j,s}^{\text{FullInfo}}, \quad \forall i, j \in \mathcal{A}_i.$$

Note that $V_{i,j}^{\text{FullInfo}}$ is constant across simulation samples.

(ii) Use Equation (A.5) to compute the student's utility from program $j$ in sample $s$ without learning, $U_{i,j,s}^{\text{NoInfo}}$:

$$U_{i,j,s}^{\text{NoInfo}} = V_{i,j,s}^{\text{NoInfo}} + \epsilon_{i,j,s}^{\text{NoInfo}}, \quad \forall i, j \in \mathcal{A}_i,$$

where $V_{i,j,s}^{\text{NoInfo}} \sim N\left(V_{i,j}^{\text{FullInfo}}, \left(\text{s.e.}(V_{i,j}^{\text{FullInfo}})\right)^2\right)$ as specified in Equation (A.6).

**Step 2: Early offer arrival and learning order.** For each student $i$ in sample $s$:

(i) Draw an arrival order $e_{i,s}$ of $i$'s early offers under the M-DA mechanism. Recall that the set of $i$'s early offers is fixed across simulation samples.

(ii) Generate the (potential) learning sequences $\omega_{i,s}^{\text{DA}}$, $\omega_{i,s}^{\text{M-DA}}$, and $\omega_{i,s}^{\text{BM-DA}}$ as specified in Section E.2.2.

**Step 3: Learning outcomes.** Let $\lambda_{i,j,s}^{m}$ be an indicator for whether student $i$ learns her true preferences for program $j$ in sample $s$ under mechanism $m \in \{\text{DA, M-DA, BM-DA}\}$. The learning outcomes $\lambda_{i,j,s}^{\text{DA}}$, $\lambda_{i,j,s}^{\text{M-DA}}$, and $\lambda_{i,j,s}^{\text{BM-DA}}$ are computed from the learning sequences $\omega_{i,s}^{\text{DA}}$, $\omega_{i,s}^{\text{M-DA}}$, and $\omega_{i,s}^{\text{BM-DA}}$ as specified in Equations (A.7), (A.8), and (A.9).

**Step 4: Submitted ROLs.** Let $U_{i,j,s}^{m}$ denote student $i$'s perceived utility from program $j$ in sample $s$ under mechanism $m \in \{\text{DA, M-DA, BM-DA}\}$. For each student $i$ in sample $s$ under mechanism $m$:

(i) Use Equation (A.10) to compute

$$U_{i,j,s}^{m} = \lambda_{i,j,s}^{m} \cdot U_{i,j,s}^{\text{FullInfo}} + (1 - \lambda_{i,j,s}^{m}) U_{i,j,s}^{\text{NoInfo}}, \quad \forall i, j \in \mathcal{A}_i.$$

(ii) Let each student submit a complete ranking of the programs in $\mathcal{A}_i$, truthful w.r.t. $U_{i,j,s}^{m}$.

**Step 5: Matching.** For each sample $s$ and mechanism $m \in \{\text{DA, M-DA, BM-DA}\}$, the program-proposing GS algorithm is used to match the students and programs based on (i) students' submitted ROLs, (ii) the programs' rankings of applicants, and (iii) the programs' capacities. Note that the last two components are constant across $s$. Let $\mu(i, s, m)$ denote student $i$'s match in sample $s$ under mechanism $m$.

As a benchmark, we also simulate the match that would be observed under full information. Specifically, we let each student rank the programs in her ROL by $U_{i,j,s}^{\text{FullInfo}}$, and then only Step 5 is needed to be run.

## E.4 Comparisons between the Mechanisms

We compare the DA, M-DA, and BM-DA mechanisms along two dimensions: students' submitted ROLs and the utility that students derive from the matching outcome.

**Submitted ROLs.** To contrast the different mechanisms in terms of how they affect students' preference discovery, we compare in each simulation the ROL that a student submits under each of the three mechanisms to the ROL that she would submit under full information (her "true" preferences).

Because it is payoff-irrelevant how an infeasible program is ranked, these comparisons are restricted to the programs that are ex post feasible to the student under the considered mechanism. Specifically, the ex post feasible programs are those to which the student applied that either did not fill their capacity or for which the student was ranked above the lowest-ranked student who was admitted to the program.

We compute the following statistic for each mechanism $m \in \{\text{DA, M-DA, BM-DA}\}$:

$$\theta^m \equiv \frac{1}{S \cdot I} \sum_{S=1}^{S} \sum_{i=1}^{I} \mathbb{1} \begin{pmatrix} \text{student } i\text{'s ex post feasible programs in sample } s \text{ under} \\ \text{mechanism } m \text{ are ranked in order of full info preferences} \end{pmatrix},$$

where $\mathbb{1}(\cdot)$ is an indicator function. In other words, $\theta^m$ is the fraction of students who, under mechanism $m$, rank ex post feasible programs in the order of their true preferences, averaged across the simulation samples.

**Expected utility of students.** To compare student welfare across mechanisms, we adopt an "ex ante" perspective by taking an average across the simulation samples. As detailed in Section E.3.2, sample-specific components include idiosyncratic utility shocks, utility without learning, early offer arrival orders, and potential learning sequences. As a result, a student's match may change across the samples.

Recall that each student learns her true preferences for the same number of programs under the DA, M-DA, and BM-DA mechanisms (see the discussion in Section E.2.2). Thus, learning costs can be ignored in the welfare comparison.

For each student, her expected utility is the average of the student's full-information utility of her matches across the simulation samples. We then perform pairwise comparisons of mechanisms, say between $m_1$ and $m_2$, based on the shares of students whose expected utility is (i) strictly higher under mechanism $m_1$ than under mechanism $m_2$, (ii) strictly lower, or (iii) equal.

Formally, let $U_{i,\mu(i,m,s),s}$ denote student $i$'s utility from $\mu(i,m,s)$, her match in sample $s$ under mechanism $m$. If the student is assigned to a program (i.e., $\mu(i,s,m) \neq \varnothing$), $U_{i,\mu(i,m,s),s}$ is evaluated as the student's utility from the program under full information, i.e., $U_{i,\mu(i,s,m),s} \equiv U^{\text{FullInfo}}_{i,\mu(i,s,m),s}$. If the student is unmatched (i.e., $\mu(i,s,m) = \varnothing$), we assume that her utility is below that of the least preferred program in her extended feasible set (as defined in Section E.1.1), which we denote by $\mathcal{F}_i$. Specifically, this utility is equal to the utility of the student's least preferred program in $\mathcal{F}_i$ minus the standard deviation of $\epsilon^{\text{FullInfo}}_{i,j}$:

$$U_{i,\varnothing,s} \equiv \left( \min_{j \in \mathcal{F}_i} U^{\text{FullInfo}}_{i,j,s} \right) - \frac{\pi}{\sqrt{6}}.$$

Therefore, student $i$'s expected utility under mechanism $m$, denoted by $\text{EU}^m_i$, is:

$$\text{EU}^m_i \equiv \frac{1}{S} \sum_{s=1}^{S} U_{i,\mu(i,s,m),s}.$$

To compare students' expected utility under two mechanisms $m_1$ and $m_2$, we compute the following statistics:

(i) $\pi_{(m_1 \succ m_2)}$: Share of students whose expected utility is strictly higher under mechanism $m_1$ than under mechanism $m_2$:

$$\pi_{(m_1 \succ m_2)} \equiv \frac{1}{I} \sum_{i=1}^{I} \mathbb{1}(\mathrm{EU}_i^{m_1} > \mathrm{EU}_i^{m_2}).$$

(ii) $\pi_{(m_2 \succ m_1)}$: Share of students whose expected utility is strictly lower under mechanism $m_1$ than under mechanism $m_2$:

$$\pi_{(m_2 \succ m_1)} \equiv \frac{1}{I} \sum_{i=1}^{I} \mathbb{1}(\mathrm{EU}_i^{m_1} < \mathrm{EU}_i^{m_2}).$$

(iii) $\pi_{(m_1 \sim m_2)}$: Share of students whose expected utility is the same under both mechanisms $m_1$ and $m_2$:

$$\pi_{(m_1 \sim m_2)} \equiv \frac{1}{I} \sum_{i=1}^{I} \mathbb{1}(\mathrm{EU}_i^{m_1} = \mathrm{EU}_i^{m_2}).$$

In Section IV.D of the main text, we use the above statistics to compare student welfare under (1) full information versus the DA, (2) the M-DA versus the DA, (3) the BM-DA versus the DA, and (4) the BM-DA versus the M-DA.

# Appendix References


**Bogomolnaia, Anna and Hervé Moulin**, "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 2001, *100* (2), 295–328.

**Calonico, Sebastian, Matias D. Cattaneo, and Rocco Titiunik**, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

**_ , _ , Max H. Farrell, and Rocco Titiunik**, "`rdrobust`: Software for Regression Discontinuity Designs," *Stata Journal*, 2017, *14* (4), 909–946.

**Cattaneo, Matias D., Michael Jansson, and Xinwei Ma**, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

**de Chaisemartin, Clément and Luc Behaghel**, "Estimating the Effect of Treatments Allocated by Randomized Waiting Lists," *Econometrica*, 2020, *88* (4), 1453–147.

**Imbens, Guido and Karthik Kalyanaraman**, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 2012, *79* (3), 933–959.

**Roth, Alvin E. and Marilda A. Oliveira Sotomayor**, *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph Series, Cambridge University Press, 1990.