RATIONALITY & COMPETITION
CRC TRR 190

Fairness in Matching Markets: Experimental Evidence

Tobias König (Linnaeus University)
Dorothea Kübler (WZB Berlin, TU Berlin)
Lydia Mechtenberg (University of Hamburg)
Renke Schmacker (WZB Berlin, DIW Berlin)

Discussion Paper No. 442

November 01, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Fairness in Matching Markets: Experimental Evidence*

Tobias König Dorothea Kübler
Lydia Mechtenberg Renke Schmacker

July 7, 2023

Abstract

We investigate fairness preferences in matching mechanisms using a spectator design. Participants choose between the Boston mechanism or the serial dictatorship mechanism (SD) played by others. In our setup, the Boston mechanism generates justified envy, while the strategy-proof SD ensures envy-freeness. When priorities are merit-based, many spectators prefer the Boston mechanism, and this preference increases when priorities are determined by luck. At the same time, there is support for SD, but mainly when priorities are merit-based. Stated voting motives indicate that choosing SD is driven by concerns for envy-freeness rather than strategy-proofness, while support for the Boston mechanism stems from the belief that strategic choices create entitlements.

JEL-codes: D47; C92; I24; D72
Keywords: Matching markets; school choice; voting; Boston mechanism; sincere agents; justified envy

*Financial support by the Deutsche Forschungsgemeinschaft through CRC TRR 190 ('Rationality and Competition') and Cluster of Excellence 'Contestations of the Liberal Script' (EXC 2055) is gratefully acknowledged. The paper has benefited from discussions with participants of the WZB MC seminar, the Berlin-Munich Seminar, the Matching in Practice Workshop, and the Lausanne Matching and Market Design Workshop 2020. We thank the Vienna Center of Experimental Economics (VCEE), University of Vienna, for allowing us to run part of our experiment in their laboratory.

Tobias König: Linnaeus University, tobias.koenig[at]lnu[dot]se
Dorothea Kübler: WZB Berlin Social Science Center & Technical University Berlin, dorothea.kuebler[at]wzb[dot]eu
Lydia Mechtenberg: University Hamburg, lydia.mechtenberg[at]uni-hamburg[dot]de
Renke Schmacker: University of Lausanne, renke.schmacker[at]unil[dot]ch

1 Introduction

Matching procedures are employed to allocate goods and services in many important domains of life. Examples include the allocation of seats at schools and universities, public housing, and appointments at public offices. Often, individuals are required to state their preferences for the available options. By taking into account the applicants' priorities, which can be based on test scores, lottery draws, or waiting time, the clearinghouse creates a matching outcome.

Fairness is a paramount criterion for assessing matching mechanisms. A long-standing theoretical literature criticizes the widely-used Boston mechanism for violating principles of fairness. The mechanism leads to justified envy and is not strategy-proof. Moreover, sophisticated applicants who act strategically can benefit at the expense of individuals who submit their true preferences (Pathak and Sönmez, 2008). In the school-choice context, a matching is considered free from justified envy if there is no individual who prefers another school over her school and is entitled to it according to that school's priorities (Abdulkadiroğlu and Sönmez, 2003; Balinski and Sönmez, 1999). In a strategy-proof mechanism, students cannot benefit from misreporting their preferences, regardless of other students' choices, since truth-telling is a dominant strategy. Policy-makers and economists have argued that individuals' ability to make smart strategic choices should not impact the allocation of goods and services they are otherwise entitled to (Abdulkadiroğlu *et al.*, 2006). This argument played a major role in the decision to abandon the Boston mechanism in the city of Boston.[1]

[1]In a memo to the Boston School Committee on May 25th 2005, former Boston Public Schools superintendent Thomas Payzant described the rationale for switching from a manipulable school-choice mechanism to a strategy-proof mechanism as follows: "A strategy-proof algorithm *levels the playing field* by diminishing the harm done to parents who do not strategize or do not strategize well." (Abdulkadiroğlu *et al.*, 2006)

Despite extensive discussions in the theoretical matching literature, there remains a scarcity of empirical evidence regarding people's actual preferences for the fairness of matching mechanisms. This paper aims to address this gap by conducting laboratory experiments to examine individuals' fairness views on matching mechanisms. To isolate fairness preferences, we enable subjects who are not directly engaged in the matching problem to vote on the preferred mechanism to be applied for another group of subjects. The spectator design allows us to observe their choices regarding a mechanism in which they have no personal stakes. This approach provides valuable insights into the normative principles they adhere to, independent of immediate personal involvement and self-interest (Konow, 2000; Cappelen *et al.*, 2013). Furthermore, we introduce variations in whether the priorities are determined by earned entitlements or through a lottery. This enables us to investigate potential interactions between preferences for matching mechanisms and the perceived legitimacy of the matching priorities.

In a laboratory experiment, we let subjects choose between the Boston mechanism where misreporting preferences can be individually optimal and outcomes are characterized by justified envy, and the strategy-proof and envy-free serial dictatorship mechanism (SD). The Boston mechanism is frequently used for school choice, as is the deferred-acceptance algorithm that reduces to SD schools or universities have common priorities. Both theoretical fairness criteria point toward SD, but actual preferences may be different. For example, individuals may have different attitudes toward the frequently criticized aspect of the Boston mechanism, whereby sophisticated individuals who act strategically can gain advantages over those who honestly state their preferences (Pathak and Sönmez, 2008). While criticized by parents, policymakers, and economists for violating principles of equity and fairness, some individuals could perceive a mechanism that honors clever behavior as meritocratic.

Our experiment aims to ascertain preferences for matching mechanisms that differ in terms of two fairness criteria: strategy-proofness and justified envy. In order to elicit these preferences, we have designed an experiment with three key features: (i) We consider a simple school-choice matching problem where efficiency and distributional concerns are absent. By design, there is always full efficiency, and the distribution of payoffs is fixed. (ii) We exogenously impose heterogeneity in the individuals' ability to make strategic choices by mandating some participants to be sincere. Thus, we use a setup akin to Pathak and Sönmez (2008). This leads to justified envy in our setup. (iii) We shut down all elements of self-interest and personal involvement by eliciting the preferences of impartial spectators (Konow, 2000; Cappelen *et al.*, 2013).

Our results reveal that the vast majority of spectators are willing to sacrifice a portion of their payout to influence which matching mechanism will be used for the other subjects. Among these spectators, the majority prefer the Boston mechanism, regardless of whether priorities are based on lotteries or earned entitlements. Moreover, we observe that around one-third of spectators prefer the serial dictatorship mechanism when matching priorities are based on earned entitlements. However, preferences for the serial dictatorship interact with the way priorities are determined, and it is selected less frequently when priorities are based on a lottery. The voting results indicate that while concerns for the absence of justified envy and strategy-proofness exist, they are not the only factors individuals consider when assessing the fairness of matching mechanisms. In fact, many spectators perceive the Boston mechanism as fair, suggesting that they value a mechanism that rewards strategic behavior.

Combining voting choices with data from our post-experimental surveys supports this interpretation. We asked spectators to indicate their agreement with a set of fairness-related statements, and we utilized the least absolute shrinkage and selection operator (lasso) to identify the statements that have the strongest impact on voting behavior. Respondents' agreement with the statement that a mechanism should reward clever behavior significantly increases support for the Boston mechanism. Conversely, agreement with the statement "it is fair for students to receive a school place based on their rank," is significantly associated with voting for SD. Interestingly, preferences for strategy-proofness do not strengthen support for SD when controlling for envy-freeness. This suggests that people do not inherently value strategy-proofness. Instead, voting for the serial dictatorship mechanism is primarily driven by the desire for envy-freeness. Notably, although people widely express the opinion that matching mechanisms should be simple and transparent, these value statements were not selected by the lasso, and have no predictive power for voting choices in bivariate regressions.

To gauge the robustness of our findings, we conduct additional experiments. First, we examine the possibility that spectators do not trust SD to produce an envy-free matching. Therefore, we run a control treatment where subjects choose between the Boston mechanism and a direct implementation of the envy-free matching.[2] The results of this experiment align with the findings in our main experiment. Second, we investigate whether our results are influenced by the fact that spectators are selecting mechanisms for *other* subjects. Therefore, we implement a "veil of ignorance" treatment where subjects choose a mechanism for their own group without knowledge of their own capability to manipulate their rank-order list. In this robustness treatment, which is only conducted for randomly assigned priorities, we once again observe a substantial majority of participants selecting the Boston mechanism.

Our experiment contributes to a growing literature on the behavioral aspects of matching markets.[3] Pathak and Sönmez (2008) show that sophisticated players can benefit from

[2]The experimental literature shows that even in strategy-proof mechanisms, individuals often depart from truthful reporting (see Hakimov and Kübler, 2020), which can lead to matching outcomes exhibiting justified envy.

[3]For a survey of matching experiments see Hakimov and Kübler (2020).

a non-strategy-proof matching algorithm (such as the Boston mechanism), as it allows them to exploit sincere players who do not act strategically. We examine whether these implications of the Boston mechanism are relevant for the preferences of impartial spectators. Surprisingly, we find that a considerable portion of individuals vote for the Boston mechanism because they believe that an outcome based on strategic choices is fair. This finding provides a novel explanation for the widespread use of the Boston mechanism.

The heterogeneity of people's ability to act strategically has received some attention in the empirical matching literature. For instance, Dur *et al.* (2018) document that there is variation in students' ability to understand the incentives of the Boston mechanism to avoid over-demanded schools, using data from the US. Moreover, a number of studies find that individuals hold incorrect beliefs about their admission chances, which prevents them from choosing optimally in the Boston mechanism (Kapor *et al.*, 2020; de Haan *et al.*, 2023; He, 2017). Basteck and Mantovani (2018) experimentally show that differences in cognitive skills translate into differences in matching outcomes in the Boston mechanism, with this link being less pronounced in a strategy-proof mechanism. In contrast, we exogenously vary the ability to manipulate the rank-order list and let subjects vote over matching mechanisms when some agents are known to be sincere.

In practice, the absence of justified envy has proven to be an important criterion for school-choice mechanisms. For example, after the Boston School Committee abolished the manipulable Boston mechanism, a decision had to be made between the Pareto inefficient deferred-acceptance mechanism (DA) and the efficient top-trading cycles mechanism (TTC) (Abdulkadiroğlu and Sönmez, 2003). While TTC was initially preferred on grounds of efficiency, concerns about justified envy arising from trading priorities in TTC led to the adoption of DA (Abdulkadiroğlu *et al.*, 2006). It was also discussed whether some priorities—e.g., lottery priorities—are more legitimate to be traded than others, such as sibling priority. Our paper contributes to this discussion by showing that justified envy is perceived as less problematic in environments where priorities are based on a lottery than in environments with merit-based priorities.

Lastly, our paper is connected to a literature that employs the spectator design to investigate distributive preferences. Previous studies have found that individuals are more likely to accept unequal payouts when the inequality is based on merit (e.g., due to choices or effort) as opposed to luck (e.g. Konow, 2000; Cappelen *et al.*, 2013; Durante *et al.*, 2014; Mollerstrom *et al.*, 2015; Deffains *et al.*, 2016; Cettolin and Riedl, 2017; Almås *et al.*, 2020; Andre, 2022).[4] Our study contributes to this literature by contrasting the merit of hard work with the merit of clever behavior, connecting to the ongoing debate in philosophy and the social sciences about what constitutes merit in modern societies (e.g. Sandel, 2020). Interestingly, we find that many individuals consider the Boston mechanism fair, even though strategic agents gain at the expense of subjects who are sincere at no fault of their own. This finding contributes to recent literature showing that many individuals are willing to accept "factual merit" arguments to justify inequality, without questioning its origins (Andre, 2022; Bartling *et al.*, 2018).

2 Matching markets: basic concepts

In this section, we introduce some concepts of matching theory that will be used to characterize the experimental school-choice problem. Each student requires one seat at a school while each school can admit many students, up to its capacity. The students have strict preferences over schools, and schools have priorities according to which they rank students. We consider a setup where the number of students is equal to the number of seats.

[4]Other papers use veil-of-ignorance designs, in which subjects decide on a redistribution scheme before they know whether they are net payers or receivers (e.g. Krawczyk, 2010; Schildberg-Hörisch, 2010; Durante *et al.*, 2014; Deffains *et al.*, 2016; Gerber *et al.*, 2019). We also include a veil-of-ignorance treatment as a robustness check, with similar results.

A matching is considered *free of justified envy* if the following condition is satisfied: There is no student who prefers another school to her match and has a higher priority at this school than an admitted student. This concept of justified envy-freeness is commonly regarded as fairness (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003).

In our setup, all students rank schools in the same way, that is, they agree on the best school, the second-best school, etc. At the same time, all schools rank students based either on their performance rank or based on a single lottery. Thus, there exists an *assortative* matching outcome. A matching outcome is assortative if the highest-ranked students are matched to the most-preferred school, the next-highest-ranked students are matched to the second-most-preferred school, and so on. Thus, the assortative matching outcome is free of justified envy.

In the experiment, subjects can choose between the *Boston mechanism* and the *serial dictatorship (SD) mechanism*, which work as follows:

Boston mechanism All students submit a rank-order list of schools. The student assignment mechanism iterates the following steps:

- Step 1: Each student applies to the school that she ranks first. Each school admits students up to its capacity, in the order of their priorities. These assignments are final. The remaining students are rejected.
- Step k, k$\geq$2: Each student who was rejected in the previous step applies to the next school in the rank-order list. Each school admits students up to its remaining capacity, according to their priority. These assignments are final. The remaining students are rejected.
- The algorithm terminates when no student is rejected, or all schools have filled their capacity.

The Boston mechanism is not strategy-proof: truthful preference revelation is not a (weakly) dominant strategy for agents. For example, it is straightforward to see that it can be optimal for a student to skip her most preferred school if she has no chance of being admitted. In Nash equilibrium, the outcome of the Boston mechanism is stable (Ergin and Sönmez, 2006), which means that no blocking pair of a student and a school can profitably deviate. In the context of school choice, this implies that the matching is free of justified envy (Abdulkadiroğlu and Sönmez, 2003). However, the equilibrium requires strategic play by the students. If some students do not act strategically, outcomes characterized by justified envy can occur in which a student prefers a seat at a school where she has a higher priority than a student who has been admitted (Pathak and Sönmez, 2008).

Serial dictatorship mechanism (SD) All students submit a rank-order list of schools. The student assignment mechanism iterates the following steps:

- Step 1: The student with the highest priority is assigned to the school at the top of her rank-order list and the assignment is final. The capacity of the respective school is reduced by one.
- Step k, k≥2: The student with the next highest priority is assigned to the highest school of her rank-order list that has remaining capacity. These assignments are final.

SD is strategy-proof and eliminates justified envy. In contrast to the Boston mechanism, students cannot be better off when deviating from truthful preference revelation.

3 Experimental Design

The purpose of the experiment is to investigate fairness preferences regarding matching mechanisms. Subjects can decide between the Boston mechanism and SD. In our experiment, we allow subjects to make a choice, incurring a personal cost, regarding which mechanism should be implemented for a separate group of subjects. Given that these individuals have no personal stake in the outcome, we can assume that their choices reflect their views on fairness principles rather than being influenced by their beliefs about personal benefits arising from the selected mechanism (see e.g., Cappelen *et al.*, 2007; Konow, 2000, 2009).

Matching market We choose a simple matching market with identical student preferences and school priorities. The experimental matching problem consists of eight students who need to be allocated to five schools. Schools A and B have one seat each, whereas schools C, D, and E have two seats each. Preferences for schools are induced as follows: students earn 12 Euros if they receive a seat at school A, 10 Euros at school B, 8 Euros at school C, 6 Euros at school D, and 4 Euros at school E.

The experimental school choice game is implemented with groups of eight subjects each (referred to as *players* in what follows). Prior to conducting the player sessions, *impartial spectators* decide on the matching mechanism that will be applied to the matching problem.

Student priorities We employed a between-subjects design to vary whether student priorities were randomly assigned or based on performance in a real-effort task. This enables us to investigate whether individuals' preferences for the mechanisms are influenced by whether matching priorities are based on earned entitlements.

- **Treatment EFFORT:** In this treatment, student priorities at schools are determined by the performance in a real-effort task. The task involves counting zeros in matrices composed of zeros and ones for 10 minutes. The task is based on Abeler *et al.* (2011) (a screenshot of the task can be found in Figure B.4 in the appendix). During this time, subjects are permitted to access the internet by clicking an on-screen button. When performing the real-effort task, subjects do not know that they will participate in a school-choice game later on.[5]

- **Treatment RANDOM:** In this treatment, students are randomly assigned their rank and, thereby, their priority at schools.

Based on their priority, students are denoted as student 1, student 2, etc. (where lower numbers mean higher student priorities).

Sincere students The Boston mechanism violates the condition of no justified envy with respect to the stated preferences. Moreover, a well-known feature of the Boston mechanism is that, in the presence of heterogeneity in the ability to act strategically, certain individuals may benefit at the expense of those who do not manipulate their preferences, (e.g., Pathak and Sönmez, 2008). To capture both properties in the laboratory environment, we exogenously impose sincerity on a subset of agents: for these students, the induced preferences are automatically submitted to the clearinghouse. The remaining students are free to submit their preferences as they choose. While the Nash equilibrium outcome of the Boston mechanism is envy-free (Ergin and Sönmez, 2006), our design with sincere agents guarantees justified envy and makes this transparent for the participants. By contrast, the strategy-proof SD mechanism protects sincere individuals from being assigned to worse matches and guarantees the absence of justified envy.

Specifically, in each student group of eight, we randomly select two students for whom schools in the rank-order list are ranked according to their potential payout. We are

[5] They are only told that a better performance will, on average, pay out in the experiment. We ensure that this is true by randomizing between groups which ranks are forced to be sincere. Hence, exerting higher effort provides a higher expected payoff.

focusing on interesting cases where we select students in ranks 2 and 3, 3 and 4, or 4 and 5.[6] We ensure that these combinations are balanced across treatments and sessions.

This design feature ensures common knowledge about the existence of individuals who are sincere (as opposed to a design in which all students are unconstrained in reporting their preferences). Furthermore, by requiring some agents to be sincere and selecting a simple market with fully correlated preferences, we guarantee that there is a unique (Nash) equilibrium outcome in the matching mechanisms we employ. Additionally, we have included several incentivized comprehension questions in the instructions. The aim of our design is to create a simple and comprehensible setup for the participants, without introducing strategic uncertainty that may arise from voting on matching procedures with multiple equilibrium outcomes. This will allow us to focus on the fairness aspects of matching mechanisms. In particular, our research focus is *not* on whether agents in practice play sincerely in the Boston mechanism and are taken advantage of by strategic agents—this has been established in previous work (see Hakimov and Kübler, 2020). Instead, we aim to gain insight into the fairness preferences regarding matching mechanisms, given that it is common knowledge that some individuals cannot act strategically.

Nash equilibrium The assignment of students to schools is either based on the Boston or the serial dictatorship mechanism. Due to the presence of sincere students, the mechanisms have different Nash equilibria in strategies, but a unique equilibrium matching outcome.

In the **Boston mechanism** with sincere students on ranks 2 and 3, student 4 can get into school B by listing it first, students 5 and 6 can get into school C by listing it first, and students 7 and 8 get into school D as long as they list it first, second, or third. Students 2 and 3 end up at the least attractive school E. This is an equilibrium since no profitable

[6] These are the interesting cases since the equilibrium outcome is unchanged when the students in rank 1, 7, or 8 report their preferences truthfully.

deviation exists.[7] The resulting matching is

$$\text{(Boston Nash Equil.)} \qquad ((1,A),(2,E),(3,E),(4,B),(5,C),(6,C),(7,D),(8,D)).$$

This matching outcome violates justified envy-freeness: the sincere students 2 and 3 envy, for instance, student 4 as the latter is matched to school B. The envy is justified since students 2 and 3 have a higher priority at school B than student 4.

Analogous equilibrium outcomes occur if students 3 and 4, or 4 and 5 are sincere. All students in ranks below the sincere students move up to a better school while the sincere students are admitted to school E.

In the **serial dictator mechanism**, the (weakly) dominant strategy for agents is to submit their preferences truthfully. Thus, sincere students, who are forced to submit their preferences truthfully, are not made worse off; they are protected by the strategy-proofness of the matching mechanism. The unique Nash equilibrium is the stable assortative matching where the student ranked first is admitted by school A, the student ranked second is admitted by school B, and so on:

$$\text{(Assortative)} \qquad ((1,A),(2,B),(3,C),(4,C),(5,D),(6,D),(7,E),(8,E)).$$

This matching outcome is free of justified envy; any envy that exists can be reconciled by the fact that students matched to higher-ranked schools had higher priorities.

Experimental procedures At the start of a session, spectators are informed that their task is to select the mechanism that will be used for a future group of participants. They

[7] Note that there are multiple equilibrium strategies, since the ranking of schools after the first does not matter. Moreover, students 7 and 8 can submit any ranking as long as school D is not ranked last or second to last. Despite multiple equilibrium strategies, the equilibrium outcome is unique since the congested schools (A, B, C, and D) are filled in the first round of the Boston mechanism based on students' first ranked school.

are then presented with the instructions given to the participants in the player sessions. In the EFFORT treatment, spectators have the opportunity to try the counting-zeros task for ten minutes (without receiving feedback on their performance). In the RANDOM treatment, this initial stage is omitted. Afterward, spectators read the instructions for the players and are asked to answer comprehension questions regarding their understanding of the mechanism. The order in which matching procedures are explained is balanced on the session level. Incentives are provided for correctly answering the comprehension questions to encourage careful reading of the instructions.[8]

After reading the instructions for the player sessions, spectators can vote for one of the matching procedures. For every group of eight players, there is a corresponding group of eight spectators. Each group of spectators is asked to vote for the implementation of the matching procedure for one group of players, after being informed about which players in that group are sincere (either players in ranks 2 and 3, 3 and 4, or 4 and 5). The voting outcome is determined using the random dictator rule. The randomly chosen subject whose voting decision is implemented pays 0.05 Euros. If this subject has decided to abstain from voting, she pays nothing and a random draw determines which mechanism is applied. Since each group of spectators decides for one group of players, spectators are pivotal with a probability of $\frac{1}{8}$. We label the sessions in which subjects play the mechanism determined by the spectators as PLAYER_EFFORT and PLAYER_RANDOM for the EFFORT and RANDOM treatments, respectively.

The sample consists of 96 subjects in the EFFORT treatment and 96 subjects in the RANDOM treatment. Correspondingly, there are 96 subjects in the PLAYER_EFFORT and PLAYER_RANDOM sessions each, for whom the selected mechanism is implemented.

[8]In addition to the show-up fee, spectators receive 5 Euros for answering comprehension questions about the matching procedures and 2 Euros for answering comprehension questions about the voting stage. If subjects do not give the correct answer on the first attempt, they can try another time or ask the experimenter for clarification.

Descriptive statistics for the spectator sessions are given in Table A.1 and for the player sessions in Table A.3. Spectators received a show-up fee of 10.05 Euros, and players 8.00 Euros. The experimental sessions of both treatments lasted on average 60 minutes. Average earnings (including the show-up fee) were 17.05 Euros in the spectator sessions and 16.88 Euros in the player sessions. The experimental software was programmed using z-Tree (Fischbacher, 2007) and subjects were recruited using ORSEE (Greiner, 2015). The spectator sessions were carried out at the WZB-TU Berlin lab; the player sessions were conducted in the VCEE lab in Vienna.

4 Results

Before presenting the main result on the voting behavior of spectators, we show that behavior and outcomes in the matching markets are close to the theoretical predictions. Based on the spectator's voting behavior, the random dictator rule selected the Boston mechanism for eleven groups in the RANDOM session, in six groups in the EFFORT session, and in six groups in the ASSORT session. Table A.4 displays the share of players who choose the equilibrium strategy in the Boston mechanism. A large fraction of subjects exhibit equilibrium behavior: in RANDOM the share is 77.3%, in EFFORT it is 83.3%. Panel B illustrates that approximately 70% of the subjects are matched with the school to which they are assigned in Nash equilibrium. Thus, the instructions were effective in helping subjects to coordinate on equilibrium outcomes.

Moreover, we can examine the percentage of players who adhere to the equilibrium strategy in SD and reveal their preferences truthfully. On average, the rate of truth-telling is relatively low at 57.1%. This is consistent with prior research (Hakimov and Kübler, 2020). At the same time, as shown in Table A.5, deviations from truth-telling are concentrated among subjects with low priorities and have no impact on the matching outcome. Ultimately, we find that in all player sessions where SD is selected, the final outcome is the assortative matching.

4.1 Baseline vote shares

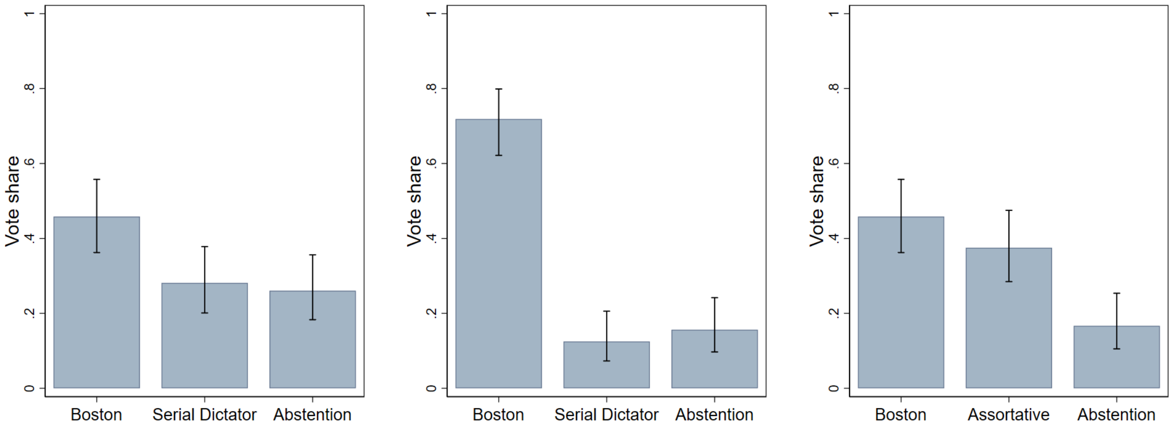
We start by presenting the baseline vote shares in the EFFORT treatment in which priorities are based on performance in the initial real-effort task. Figure 1a shows the voting results. First, we see that 73.9% of spectators are willing to incur the fee to participate in the voting process. This finding implies that a substantial share of spectators wishes to influence which matching mechanism is used for other subjects.

We observe that 45.8% of the spectators (62.0% of voters) vote for the Boston mechanism. Hence, the majority of voters express a preference for the mechanism that violates envy-freeness and strategy-proofness. At the same time, 28.1% of spectators (38.0% of the voters) opt for SD. The difference between the share favoring the Boston versus SD is statistically significant (binomial test, $p < 0.01$).

Result 1. *The majority of spectators (74%) incur a cost to participate in the voting mechanism. Among these voters, more than half (62%) prefer the Boston mechanism, while the remaining 38% opt for SD.*

The findings are in line with the idea that individuals have fairness preferences over matching mechanisms, although there is heterogeneity in their views. A larger number of respondents favor the Boston mechanism, even though it violates envy-freeness. Those individuals who express a preference for SD appear to value envy-freeness and strategyproofness.

Figure 1: Vote shares by treatment



(a) EFFORT (b) RANDOM (c) ASSORT

Notes: The figure illustrates the voting behavior of subjects across different treatments for the mechanism to be implemented for future participants. In EFFORT and RANDOM, subjects can choose between the Boston mechanism and SD, while in ASSORT they choose between the Boston mechanism and a direct implementation of the assortative matching. In EFFORT and ASSORT, priorities are based on real-effort task performance, while in RANDOM priorities are randomly determined. The lines indicate 95% confidence intervals.

4.2 Do preferences for mechanisms change when priorities are not based on merit?

In the experiments reported so far, we have investigated preferences for mechanisms when priorities are based on performance in a real-effort task. However, fairness preferences may interact with the perceived legitimacy of priorities. In the following, we compare the previous findings to a treatment in which priorities are determined at random rather than based on performance.

Figure 1b shows the voting behavior in the RANDOM treatment, that is, when priorities are assigned at random. Again, a large share of respondents (84.4%) participated in costly voting for the matching mechanisms. Regarding the voting decision, 71.9% of spectators (85.2% of voters) selected the Boston mechanism, while only 12.5% of spectators (14.8%

of voters) chose the serial dictatorship mechanism. This marks a considerable shift in voting behavior compared to the treatments where priorities are based on performance. Specifically, there is a 15.6 percentage point decrease in the proportion of spectators voting for SD ($p < 0.05$, Fisher's exact test) and a 26.1 percentage point increase in the proportion of those selecting the Boston mechanism ($p < 0.01$, Fisher's exact test).

Result 2. *When student priorities are determined randomly, the majority of spectators choose to participate in the costly voting process. The spectators primarily vote for the Boston mechanism (85%), while the remaining 15% vote for SD. This shift towards the Boston mechanism in the RANDOM treatment, compared to the EFFORT treatment, is statistically significant.*

In the next section, we use the the post-experimental survey to investigate what drives the spectators' decisions among the two mechanisms.

4.3 What explains voting decisions?

In the post-experimental survey, we include several statements related to the fairness properties of the matching mechanisms. We evaluate the relative predictive power of the fairness-related statements for the actual voting choice. This provides suggestive evidence of the factors that shape individuals' views on matching mechanisms.

We start by establishing that the voting behavior of spectators is indeed motivated by fairness considerations. We find that 82.3% of spectators weakly agree that "fairness played a role in their voting decision" (68.8% strongly agree). Moreover, 87.6% of spectators who vote for the Boston mechanism weakly agree that they "find the Boston mechanism fairer," while 82.1% of spectators who voted for SD disagree with that statement. This supports the idea that fairness is an important factor for the voting behavior of spectators.

To shed light on which fairness considerations motivate voting behavior, we asked spectators for their agreement with a battery of statements related to various aspects of fairness.[9] We select the statements that matter most for voting behavior by employing the least absolute shrinkage and selection operator (lasso).[10]

The left panel of Figure 2 displays the percentage of spectators who "agree" or "fully agree" with the statements selected by lasso. Meanwhile, the right panel illustrates the extent to which each statement predicts an individual's preference for the Boston mechanism over SD. To that end, we plot the coefficients derived from regressing an indicator of voting for the Boston mechanism on dummies indicating whether a subject agrees with the respective fairness item, using the sample of those who voted. Hence, the coefficients indicate how much higher the vote share for Boston is among individuals who agree with the statement.

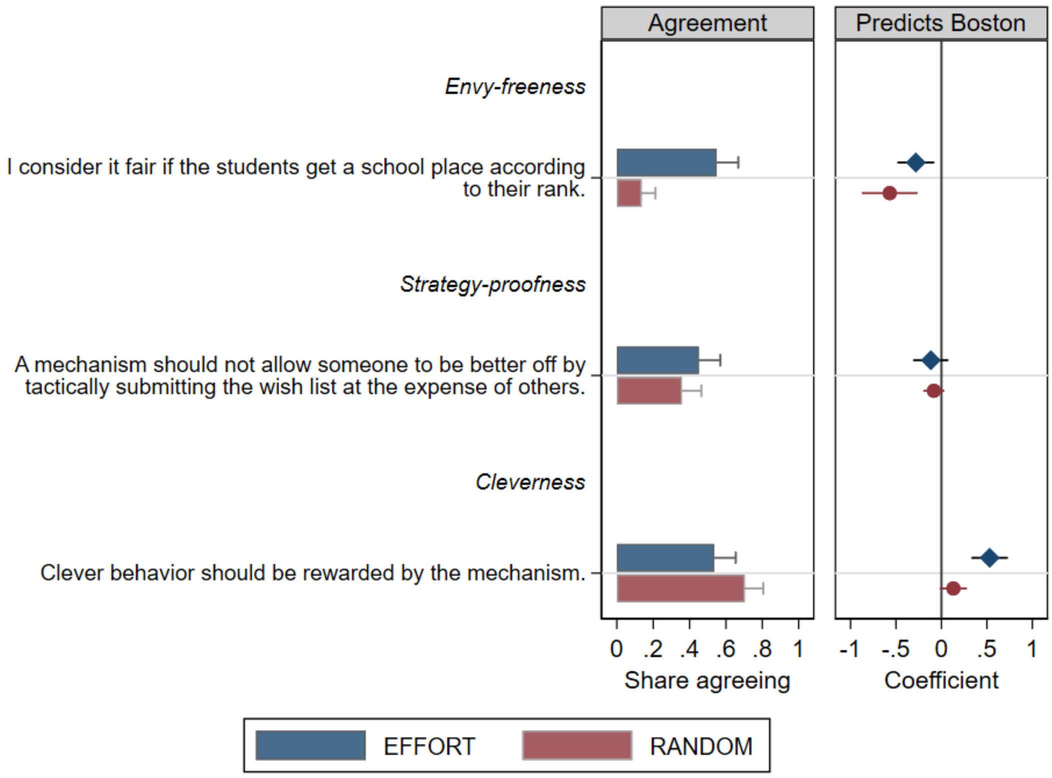
Our initial focus is on the EFFORT treatment. We first look at agreement with the statement "I consider it fair if the students get a school place according to their rank." This statement captures respondents' views on the fairness of the assortative matching and, therefore, relates to their support for an allocation that is free of justified envy in our setting. As shown in Figure 2, agreement with this statement is a significant predictor of preferences for SD. Participants who perceive the envy-free, assortative matching as fair are 28 percentage points more likely to vote for the serial dictator mechanism.

Preferences for SD could also be driven by the fact that it is strategy-proof. In the policy discussion, the idea that strategically sophisticated individuals can make themselves better off in the Boston mechanism is often viewed as undesirable (Pathak and Sönmez, 2008,

[9]The complete list of statements can be found in Appendix B.6.

[10]Lasso is a penalized regression method that selects the subset of variables that contribute the most to the predictive power of the model (Tibshirani, 1996). We use the Stata command `lasso` to implement the linear lasso and cross-validation to choose the penalty parameter. We implement the lasso for predicting votes for the Boston mechanisms (compared to votes for the serial dictatorship) using indicators of whether subjects agree or strongly agree with the respective statements. Lasso selects three out of ten items, which are discussed below.

Figure 2: Decomposition of preferences for mechanisms by treatment



Notes: The figure shows subjects' agreement with the statement in the post-experimental questionnaire and whether approval predicts voting behavior. In the left panel, we plot the share of subjects stating "Agree" or "Fully agree" with the respective statements. In the right panel, we plot coefficients from regressing a Boston vote indicator jointly on the z-standardized items. The Boston vote indicator is one if a subject voted for the Boston and zero if a subject voted for SD (abstentions excluded). The lines indicate 95% confidence intervals.

2013). We ask whether individuals agree that a mechanism should not allow individuals to benefit from tactically submitting the wish list. We find that 45.1% of all spectators report an aversion against a mechanism that is not strategy-proof. However, the approval to this statement does not significantly predict voting for the strategy-proof SD mechanism, when controlling for the other fairness items. This suggests that preferences for strategy-proofness do not independently influence voting behavior beyond preferences for envy-freeness.[11]

One possible explanation of the significant support for the Boston mechanism is the conviction that clever, strategic choices merit higher payouts. To elicit such potential views, we asked spectators whether "clever behavior should be rewarded by the mechanism." We find that 53.5% of spectators agree with this statement, and agreement is linked to a 53.0 percentage point higher vote share for the Boston mechanism.[12]

Statements capturing preferences for a simple and transparent mechanism are not selected by the lasso and do not significantly predict voting behavior. While we find that more than 90% of subjects agree that a "mechanism should be understandable and transparent for all participants," these views do not predict voting for either the Boston mechanism or SD.

Next, we focus on the results from the RANDOM treatment. Figure 2 displays that in both treatments, RANDOM and EFFORT, the view that ranks should matter for match-

[11]In a bivariate regression, the coefficient of the strategy-proofness item is significant (see Figure A.1), but loses its predictive power once we control for envy-freeness.

[12]In Figure A.2 in the appendix, we explore whether these findings are linked to broader attitudes about the type of performance that deserves merit: effort or cleverness. To test this, we generate (1) an effort merit index summarizing whether individuals agree that "diligence and effort should be rewarded" and that "those who work harder should have higher incomes" and (2) a cleverness merit index measuring whether subjects agree that "clever and strategic behavior should pay off" and that "those who behave cleverly and intelligently should have a higher income." The regression results show that attitudes toward merit strongly predict voting behavior in the EFFORT treatment. A one standard deviation higher effort merit index is associated with a 20.9 percentage points higher vote share for the serial dictator mechanism (relative to the Boston mechanism). In contrast, increasing the cleverness merit index by one standard deviation makes subjects 16.2 percentage points more likely to vote for the Boston mechanism instead of for SD.

ing students to schools strongly predicts voting for SD. However, there is a substantial treatment difference in the agreement with the statement that students should be matched based on their rank. In the RANDOM treatment, only 13.6% of individuals agreed that rank-based matching is fair, which is 41.3 percentage points lower than in the EFFORT treatment ($p < 0.01$, t-test), where 54.9% of spectators agreed. Thus, the preference for envy-free matching outcomes interacts with the perceived legitimacy of the ranks.

Preferences for rewarding strategic behavior also depend on the perceived legitimacy of priorities but to a lesser extent. When priorities are based on luck, agreement with the statement that a mechanism should reward clever behavior increases by 16.8 percentage points ($p < 0.05$, t-test). In turn, the coefficient of strategy-proofness to predict voting behavior is smaller and only marginally significant ($p = 0.09$), controlling for other fairness motives.[13] As a result, we conclude that the decrease in the vote share for SD in favor of the Boston mechanism is primarily due to the change in the perceived legitimacy of the ranks and the reduced importance of envy-freeness.

Result 3. *Voting decisions for the Boston mechanism among spectators show a positive correlation with the belief that clever choices should be rewarded. On the other hand, voting decisions for SD are correlated with the belief that fairness necessitates students receiving a school based on their rank. When controlling for the belief that envy-freeness matters, the value placed on strategy-proofness is not correlated with voting decisions. In situations where priorities are determined randomly, fewer individuals agree that an envy-free outcome is fair, while a larger number of people agree that clever behavior should be rewarded.*

Overall, our results suggest that fairness as absence of justified envy plays a significant role in the voting behavior of participants, in particular when priorities reflect earned entitlements. However, our findings challenge the conventional view that individuals value

[13]Figure A.1 in the appendix shows that preferences for cleverness significantly predict voting behavior when not controlling for the other motives.

the strategy-proofness of mechanisms. Instead, we observe that many participants find that strategic behavior merits a high payoff and, as a result, prefer the Boston mechanism.

5 Robustness

5.1 Voting over the Boston mechanism versus the assortative matching outcome

If individuals do not truthfully report their preferences, the serial dictatorship mechanism may not be the optimal benchmark for evaluating the preference for justified envy-freeness (see Hakimov and Kübler, 2020, for evidence on misreporting in strategy-proof mechanisms). Such behavior undermines the mechanism's ability to ensure the absence of justified envy.[14] As a result, spectators may not vote for the serial dictatorship mechanism, even if they value envy-freeness. In the following, we show that our voting results are robust to introducing a matching procedure that always implements the assortative matching outcome.

Design of the ASSORT treatment In this robustness treatment, we let spectators decide between the Boston mechanism and a direct implementation of the envy-free, assortative matching. The experimental design is identical to the design of the main study described in Section 3, with the exception that spectators can now choose between the Boston mechanism and a mechanism that does not require the future participants to submit rank-order lists. Instead, student 1 is automatically assigned to School A, student 2 to School B, student 3 to School C, etc. Priorities are based on performance in the same

[14]We show in Table A.5 that many lower-ranked individuals misreport their induced preferences. However, despite this, all our player sessions result in the assortative matching outcome.

real-effort task as before. We obtained a total of 96 spectators who made decisions for 96 future players.[15]

Results Figure 1c shows the results of the ASSORT treatment. Overall, 45.8% of spectators (55.0% of voters) vote for the Boston mechanism, while 37.5% (45.0% of voters) vote for the assortative matching; 16.7% of the spectators abstain from voting.

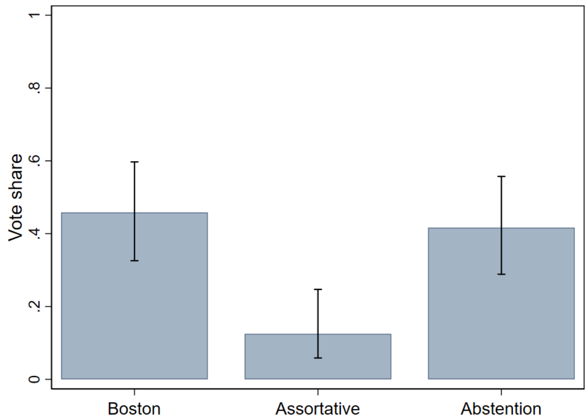
While we observe that slightly more subjects vote for the assortative matching than for SD in the EFFORT treatment, the difference is only 9.4 percentage points and not statistically significant ($p = 0.22$, Fisher's exact test). Similar to our previous findings, a majority of voters prefer the Boston mechanism and the vote share is the same as in the EFFORT treatment. Thus, we find that spectators express similar preferences when voting for the serial dictatorship mechanism and the envy-free matching outcome, relative to the Boston mechanism. This suggests that the high share of votes for the Boston mechanism is not driven by subjects doubting that serial dictatorship reaches the envy-free assortative outcome. Furthermore, those people who value envy-freeness tend to place greater importance on the ultimate outcomes rather than on the specific process that leads to them. This is also consistent with the results of our post-experimental questionnaire, which suggest that people do not value strategy-proofness per se.

5.2 Voting behind a veil of ignorance

It is possible that making decisions for others leads to different voting results compared with decisions that have the potential to impact oneself. For example, individuals may not be sufficiently interested in mechanisms that do not apply to themselves (e.g., they may

[15]Spectators received a show-up fee of 10.05 Euros and average earnings (including incentivized comprehension questions) were 17.05 Euros. Players received a show-up fee of 8.00 Euros and received average earnings of 16.88 Euros. Experimental sessions lasted around 60 minutes and were conducted in the WZB-TU Berlin lab.

Figure 3: Vote shares in the VEIL treatment



Notes: The figure shows the share of subjects (n=48) who vote for the Boston mechanism, for the assortative matching, and who abstain from voting by treatment. Subjects vote on behalf of their own group but do not know their rank yet. Ranks are randomly assigned.

not read the instructions carefully), which could potentially interfere with our observed fairness preferences regarding matching mechanisms.

In this section, we show the results from an alternative approach to elicit fairness views, namely a veil of ignorance (treatment VEIL). The main difference to the spectator design is that subjects participate in the matching mechanism, meaning that their voting decision will be implemented for their own group (instead of for future subjects). However, individuals are unaware of their future role, that is, priority at the schools, at the voting stage. The reasoning behind this is that without knowledge of their position, individuals can make decisions that are not influenced by their own self-interest. However, as subjects at the voting stage anticipate that they will have to play the mechanisms themselves, they might be affected by their (over-)confidence in their understanding of the mechanism and their ability to strategize and outsmart others. For this reason, we believe that the spectator design is more suitable for providing evidence on fairness concerns.

The VEIL treatment is only implemented for random priorities, since ranks based on real effort would allow subjects to form expectations about their priorities.[16] As shown in Figure 3, we find that 45.8% of subjects (78.6% of voters) voted for the Boston mechanism, while 12.5% (21.4% of voters) voted for the assortative matching. It should be noted that although the experimental designs are not directly comparable, these vote shares are similar to those in the RANDOM treatment mentioned earlier, where 85.2% of voters chose the Boston mechanism and 14.8% chose the serial dictator mechanism. These results suggest that the voting behavior we observe in our paper is not an artifact of using a spectator design.[17]

6 Conclusion

With the help of lab experiments, we explored the fairness preferences of impartial spectators concerning matching procedures. Our findings reveal a new underlying motive behind people's endorsement of the Boston mechanism. Despite the involvement of sincere students who experience adverse effects due to the Boston mechanism, a significant majority of spectators still exhibit a preference for it over the strategy-proof serial dictatorship (SD), which produces an envy-free matching. This preference stems from their perception of the strategic environment as fair, despite the justifiable envy it creates, as they ascribe value to clever strategic decision-making.

When priorities are effort-based, about one-third of spectators prefer the serial dictatorship mechanism, which leads to the envy-free assortative matching in equilibrium. We

[16] There are some additional design differences, precluding us from performing a treatment comparison with our main treatments. In particular, only ranks 2 and 3 were sincere, voting costs were 0.10 Euros, and the order of mechanisms was not randomized. For a detailed discussion of this treatment, see König *et al.* (2019).

[17] In an earlier version of this paper, we additionally present results from a stakeholder treatment in which subjects know their own priority when voting. Since we focus on fairness preferences in this paper, we refer the interested reader to König *et al.* (2019).

find that preferences for envy-freeness are influenced by how priorities are determined, with less emphasis placed on envy-freeness when priorities are randomly assigned. These results indicate that for applications in school choice when priorities do not reflect earned entitlements, the Top Trading Cycles (TTC) mechanism could be advisable. This mechanism guarantees efficiency while sacrificing envy-freeness.

We do not interpret our findings as a recommendation to use the Boston mechanism, but we believe that the study sheds light on the prevalence of the mechanism. In addition to the argument that sophisticated parents may find the Boston mechanism attractive in the presence of naive participants (Pathak and Sönmez, 2008), we identify another reason for the widespread use of the Boston mechanism. We find that clever behavior is perceived as justifying better outcomes, particularly when priorities are not perceived as being based on merit.

References

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E. and Sönmez, T. (2006). Changing the Boston school choice mechanism. *NBER Working Paper Series*, 11965.

— and Sönmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, **93** (3), 729–747.

Abeler, J., Falk, A., Goette, L. and Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, **101** (2), 470–492.

Almås, I., Cappelen, A. W. and Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy*, **128** (5), 1753—-1788.

Andre, P. (2022). Shallow meritocracy, `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3916303` (retrieved 01/24/2023).

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: Student placement. *Journal of Economic Theory*, **84** (1), 73–94.

Bartling, B., Cappelen, A. W., Ekström, M., Sørensen, E. Ø. and Tungodden, B. (2018). Fairness in winner-take-all markets. *NHH Dept. of Economics Discussion Paper*, no. 08/2018.

Basteck, C. and Mantovani, M. (2018). Cognitive ability and games of school choice. *Games and Economic Behavior*, **109** (1), 156–183.

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø. and Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, **97** (3), 818–827.

—, Konow, J., Sørensen, E. Ø. and Tungodden, B. (2013). Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, **103** (4), 1398–1413.

Cettolin, E. and Riedl, A. (2017). Justice under uncertainty. *Management Science*, **63** (11), 3739–3759.

de Haan, M., Gautier, P., Oosterbeek, H. and van der Klaauw, B. (2023). The performance of school assignment mechanisms in practice. *Journal of Political Economy*, **131** (2), 388–455.

Deffains, B., Espinosa, R. and Thöni, C. (2016). Political self-serving bias and redistribution. *Journal of Public Economics*, **134**, 67–74.

Dur, U., Hammond, R. G. and Morrill, T. (2018). Identifying the harm of manipulable school-choice mechanisms. *American Economic Journal: Economic Policy*, **10** (1), 187–213.

Durante, R., Putterman, L. and van der Weele, J. (2014). Preferences for redistribution and perception of fairness: an experimental study. *Journal of the European Economic Association*, **12**, 1059–1086.

Ergin, H. and Sönmez, T. (2006). Games of school choice under the Boston mechanism. *Journal of Public Economics*, **90** (1-2), 215–237.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, **10** (2), 171–178.

Gerber, A., Nicklisch, A. and Voigt, S. (2019). The role of ignorance in the emergence of redistribution. *Journal of Economic Behavior & Organization*, **163**, 239–261.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, **1** (1), 114–125.

Hakimov, R. and Kübler, D. (2020). Experiments in matching markets: A survey. *Experimental Economics*.

He, Y. (2017). Gaming the Boston school choice mechanism in Beijing, `https://sites.google.com/site/yinghuahe/files/BM.pdf` (retrieved 04/24/20).

Kapor, A. J., Neilson, C. A. and Zimmerman, S. D. (2020). Heterogeneous beliefs and school choice mechanisms. *American Economic Review*, **110** (5), 1274–1315.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, **90** (4), 1072–1091.

— (2009). Is fairness in the eye of the beholder? an impartial spectator analysis of justice. *Social Choice and Welfare*, **33**, 101–127.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, **94**, 131–141.

König, T., Kübler, D., Mechtenberg, L. and Schmacker, R. (2019). Fair procedures with naive agents: Who wants the Boston mechanism? *CRC TRR 190 Discussion Paper No. 222*, no. 222.

Mollerstrom, J., Reme, B.-A. and Sørensen, E. Ø. (2015). Luck, choice and responsibility. An experimental study of fairness views. *Journal of Public Economics*, **131**, 33–40.

Pathak, P. A. and Sönmez, T. (2008). Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review*, **98** (4), 1636–1652.

— and Sönmez, T. (2013). School admissions reform in chicago and england: Comparing mechanisms by their vulnerability to manipulation†. *American Economic Review*, **103** (1), 80–106.

SANDEL, M. J. (2020). *The Tyranny of Merit: What's Become of the Common Good?* London: Allen Lane.

SCHILDBERG-HÖRISCH, H. (2010). Is the veil of ignorance only a concept about risk? An experiment. *Journal of Public Economics*, **94**, 1062–1066.

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, **58** (1), 267–288.

Online Appendix

A Additional Tables and Figures

Table A.1: Descriptive statistics of the main treatments

	EFFORT	RANDOM
Age (Mean)	25.281	25.260
	(5.945)	(6.399)
Female (Share)	0.448	0.500
Studying (Share)	0.927	0.938
Politically left	0.594	0.604
Observations	96	96

Notes: Table shows descriptive statistics of the sessions from the main experiment. Standard deviations for continuous variables are in parentheses.

Table A.2: Descriptive statistics of the robustness treatments

	ASSORT	VEIL
Age (Mean)	22.802	22.854
	(3.272)	(4.946)
Female (Share)	0.500	0.604
Studying (Share)	0.958	0.917
Observations	96	48

Notes: Table shows descriptive statistics of the robustness treatments. Standard deviations for continuous variables are in parentheses.

Table A.3: Descriptive statistics of the player sessions

	PLAYER EFFORT	PLAYER RANDOM	PLAYER ASSORT
Age (Mean)	24.823	25.521	22.646
	(5.620)	(7.624)	(3.836)
Female (Share)	0.615	0.615	0.500
Studying (Share)	0.938	0.906	0.969
Politically left	0.583	0.604	
Observations	96	96	96

Notes: Table shows descriptive statistics of the player sessions. Standard deviations for continuous variables are in parentheses.

Table A.4: Equilibrium preference submission and outcomes in the Boston mechanism (player sessions)

	PLAYER RANDOM	PLAYER EFFORT	PLAYER ASSORT
Panel A: Equilibrium rank-order list			
Above Sincere	20/23 (87.0%)	11/12 (91.7%)	10/11 (90.9%)
Below Sincere	31/43 (72.1%)	19/24 (79.2%)	18/25 (72.0%)
Total	51/66 (77.3%)	30/36 (83.3%)	28/36 (77.8%)
Panel B: Equilibrium outcome			
Above Sincere	21/23 (91.3%)	11/12 (91.7%)	10/11 (90.9%)
Sincere	13/22 (59.1%)	9/12 (75.0%)	8/12 (66.7%)
Below Sincere	27/43 (62.8%)	14/24 (58.3%)	14/25 (56.0%)
Total	61/88 (69.3%)	34/48 (70.8%)	32/48 (66.7%)

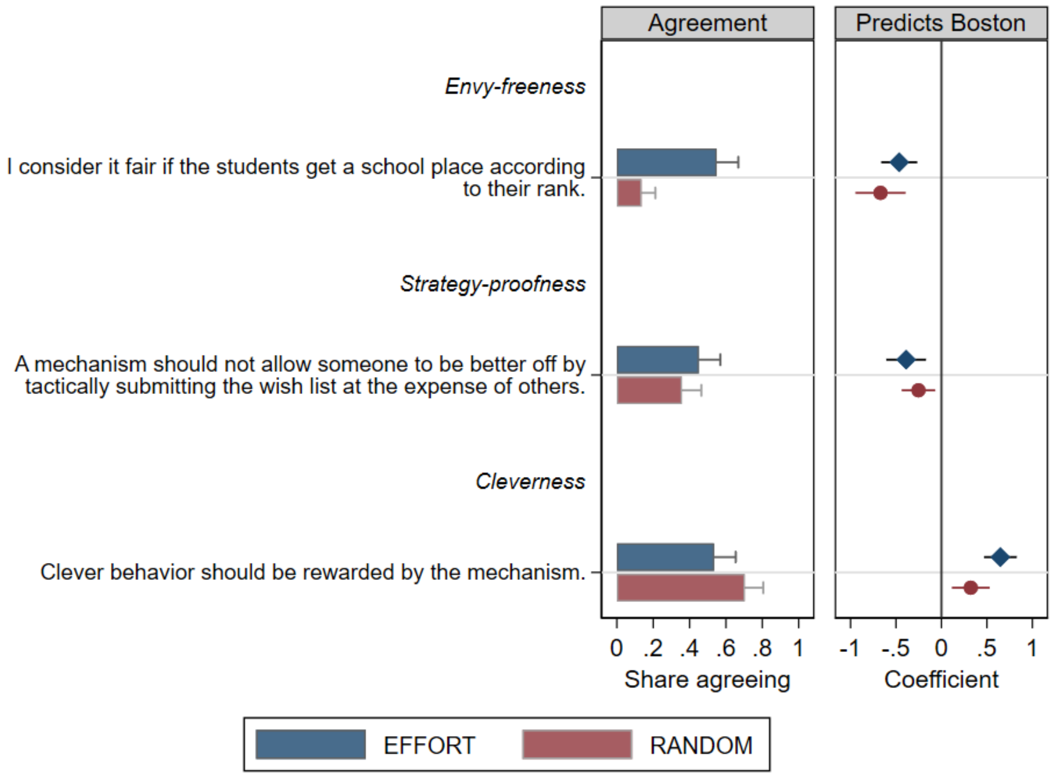
Notes: Panel A of the table shows the fraction of subjects who play the equilibrium strategy in the Boston mechanism by treatment and Panel B shows how many subjects receive the Nash equilibrium outcome. The equilibrium depends partly on who are the sincere students. In all cases, students in rank 1 list school A first and students in ranks 7 and 8 list school C first, second or third. With students in ranks 2 and 3 being sincere, students in rank 4 list school B first, while students in ranks 5 and 6 list school C first. With students in ranks 3 and 4 being sincere, students in rank 2 list school B first while students in ranks 5 and 6 list school C first. With students in ranks 4 and 5 being sincere, students in rank 2 list school B first while students in ranks 3 and 6 list school C first.

Table A.5: Truthtelling and equilibrium outcomes in SD (player sessions)

	PLAYER RANDOM	PLAYER EFFORT
Panel A: Truthtelling		
Above Sincere	1/1 (100.0%)	12/12 (100.0%)
Below Sincere	1/5 (20.0%)	10/24 (41.7%)
Total	2/6 (33.3%)	22/36 (61.1%)
Panel B: Equilibrium outcome		
Above Sincere	1/1 (100.0%)	12/12 (100.0%)
Sincere	2/2 (100.0%)	12/12 (100.0%)
Below Sincere	5/5 (100.0%)	24/24 (100.0%)
Total	8/8 (100.0%)	48/48 (100.0%)

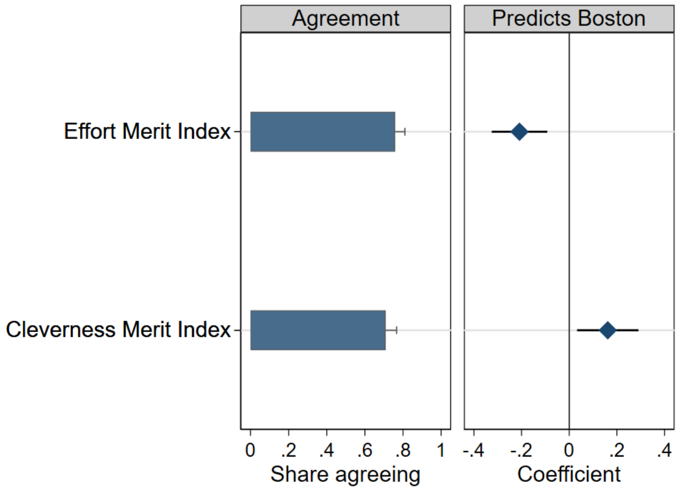
Notes: Panel A of the table shows the fraction of subjects who report their true preference list. Panel B shows how many subjects receive an outcome that coincides with the equilibrium outcome of the SD mechanism (the assortative matching).

Figure A.1: Decomposition of preferences for mechanisms by treatment (bivariate regressions)



Notes: The figure shows subjects' agreement with the statement in the post-experimental questionnaire and whether approval predicts voting behavior. In the left panel, we plot the share of subjects stating "Agree" or "Fully agree" with the respective statements. In the right panel, we plot coefficients from regressing a Boston vote indicator on the z-standardized item (bivariate regressions). The Boston vote indicator is one if a subject voted for the Boston and zero if a subject voted for SD (abstentions excluded). The lines indicate 95% confidence intervals.

Figure A.2: Decomposition of voting behavior by attitudes toward merit (only EFFORT)



Notes: The figure shows subjects' agreement with the merit indices in the post-experimental questionnaire and whether approval predicts voting behavior. The effort index includes the agreement with "diligence and effort should be rewarded" and that "those who work harder should have higher incomes," the cleverness index with "clever and strategic behavior should pay off" and that "those who behave cleverly and intelligently should have a higher income" on a 5-point Likert scale. In the left panel, we standardize the indices on a scale from 0 to 1 by dividing by the possible range. In the right panel, we plot coefficients from regressing an indicator of whether an individual votes for the Boston mechanism (rather than for SD) jointly on the z-standardized indices. Hence, The coefficients of the z-standardized index show by how many percentage points the vote share for Boston increases when the agreement with the respective statement increases by one standard deviation. The lines indicate 95% confidence intervals.

B Instructions and Screenshots

In this section, we provide the experimental instructions (translated from German) and accompanying screenshots. The order of presentation follows that of the actual experiment. Square brackets are used in the paper instructions to denote treatment-specific instructions. If screenshots are uniform across treatments, a single example is provided with annotations describing any variations.

B.1 Instructions for spectators

In this experiment, we simulate two procedures for allocating school places to students. Your task is to thoroughly understand these school place allocation procedures. At the end of the experiment, you will vote on which of these two procedures should be applied to a group of future experiment participants who will take on the role of students. This means that in the future, we will invite participants who will act as students, and today you will decide which school allocation procedure these students will be assigned to hypothetical schools by.

The two allocation procedures will be explained in the instructions given to future experimental participants. We will distribute these instructions now and ask you to read them carefully. You will also be required to answer the comprehension questions of the future participants. If you answer all the comprehension questions correctly, you can earn an additional 7 euros. If you have a question, please raise your hand. Someone will come to assist you.

[Treatment EFFORT & ASSORT: The experiment for the future participants consists of two parts. First, we will distribute the instructions for the first part, in which the participants will be given a task to complete. Then, the instructions for the second part of

the experiment, which explain the school place allocation procedures, will be distributed to the other participants.*]*

At the end of today's session, you will vote on which school place allocation procedure will be implemented for the future participants. The future participant session will take place *[Treatment EFFORT & RANDOM:* within the next 8 weeks. If you want us to inform you of the date of the future participant session, please send an email to Renke Schmacker (renke.schmacker@unil.ch) after the experiment.*] [Treatment ASSORT:* today in the afternoon*]*.

[Next, spectators are informed that they will now receive the original instructions of the future subjects (the players).]

B.2 Real-effort task instructions (only EFFORT & ASSORT)

In today's experiment, you will be divided into groups of eight participants. The composition of the groups stays the same for the entire course of the experiment.

Work Task

In this part of the experiment you have to complete a work task. This work task determines your rank from 1 to 8 in your group. The rank in your group is important for the further course of the experiment: on average, participants with a better rank receive a higher payout. It is therefore worthwhile to provide effort in the work task.

The work task is a counting task. You will be presented a series of tables with zeros and ones for a duration of 10 minutes. Your task is to count the number of zeros in each table. If you enter the correct number of zeros, you will receive a point. If you give the wrong number of zeros, you will be deducted one point. Whoever has the most points will be in rank 1; whoever has the second most points will be in rank 2; whoever has the third most points will be in rank 3, etc. After the counting task, you will be told which rank you have achieved in your group.

During the counting task, you can open the internet browser by clicking on the red button at the bottom of the screen and surf the internet in the remaining time. With the key combination Alt+Tab on the keyboard you can return to the work task at any time.

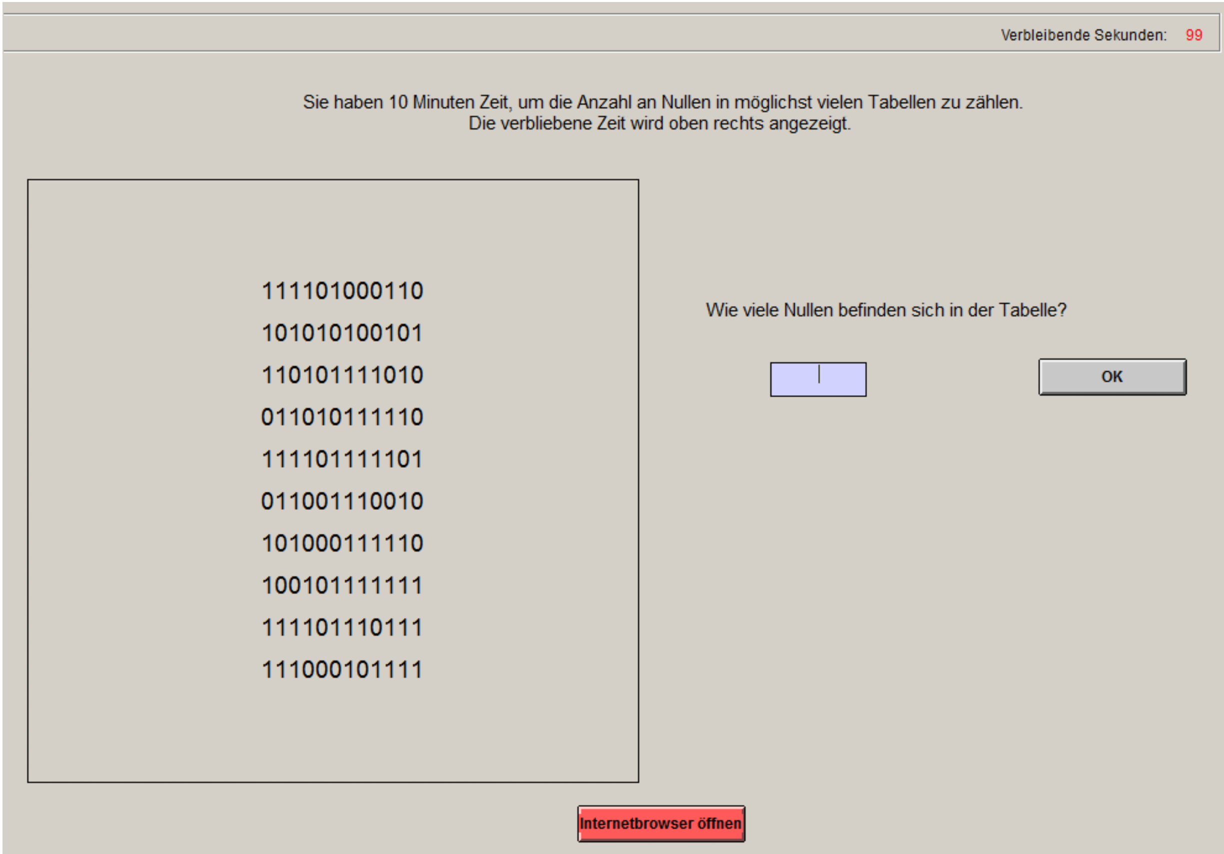
The figure below[18] illustrates the choice situation. Please wait until the work task is started on the screen.

Figure B.3: Screenshot of welcome screen (only EFFORT & ASSORT)

Notes: The screenshot displays the welcome screen in the Spectator treatment. Subject are informed about the real effort task on the next screen. In particular, they are told that they have to count the number of zeros in as many tables as possible within 10 minutes. In the top left corner of the screen, subjects in the Spectator treatment are informed that this is the original screen that subjects in the Player session will see. This note is omitted in the Stakeholder/Effort and Player treatments.

[18] The printed instructions contain a screenshot that is the same as Figure B.4.

Figure B.4: Screenshot of real effort task (only EFFORT & ASSORT)



Notes: The screenshot displays the real effort task. Subjects are supposed to count the number of zeros in the table on the left. (For subjects in the Spectator treatment the screen is only for illustrative purposes and their performance is without consequence.) After submitting the counted number, a new table is generated. The red button at the bottom of the screen opens a web browser, in which subjects can browse the internet. The browser is closed automatically after the time has run up.

Figure B.5: Screenshot of real effort task feedback (only EFFORT & ASSORT)

Notes: On this screen, subjects receive feedback on how many tables they counted correctly and which rank they achieved. In the Spectator treatment, the same screen is displayed for illustrative purposes, but instead of the number of tables and the rank three dots ("...") are displayed. Moreover, subjects are told to read the instructions for the second part of the experiment, which are handed out by the experimenter. After everyone finished reading, subjects need to enter a number announced by the experimenter to proceed.

B.3 Instructions on matching mechanisms

General Information

In this part of the experiment, we are simulating two procedures that allocate school seats to students. You and your group take on the role of students. Your task is to carefully understand these allocation procedures. At the end of the experiment you will learn which of the two procedures is used for your group of participants. Your payment for this experiment will be determined by the school at which you are placed.

Scenario and Payout

The scenario for the two procedures is as follows. Each member of your group is assigned a number from 1 to 8, representing student 1 to student 8.

[Treatment EFFORT & ASSORT: Your assigned number corresponds to the rank you achieved in the first part of the experiment. Example: If you have achieved rank 1 in the work task, your student number is 1; if you have achieved rank 8, your student number is 8.*]*

[Treatment RANDOM: Each participant in your group will be randomly assigned a rank from 1 to 8, which represents student 1 through student 8. Example: If you are randomly assigned rank 1, your student number is 1; if you are assigned rank 8, your student number is 8.*]*

There are five schools A, B, C, D and E, of which schools A and B each have one and schools C, D and E each have two seats to allocate. Hence, there are eight seats in total.

Your payout depends on which of the five schools you get a seat at. The payouts you receive at a respective school are summarized in the following table:

	School A	School B	School C	School D	School E
Payout	16 Euro	13 Euro	10 Euro	7 Euro	4 Euro

Example: Suppose you get a seat at school A. Your resulting payout is 16 Euro. If you get a seat at school E, you will receive 4 Euro.

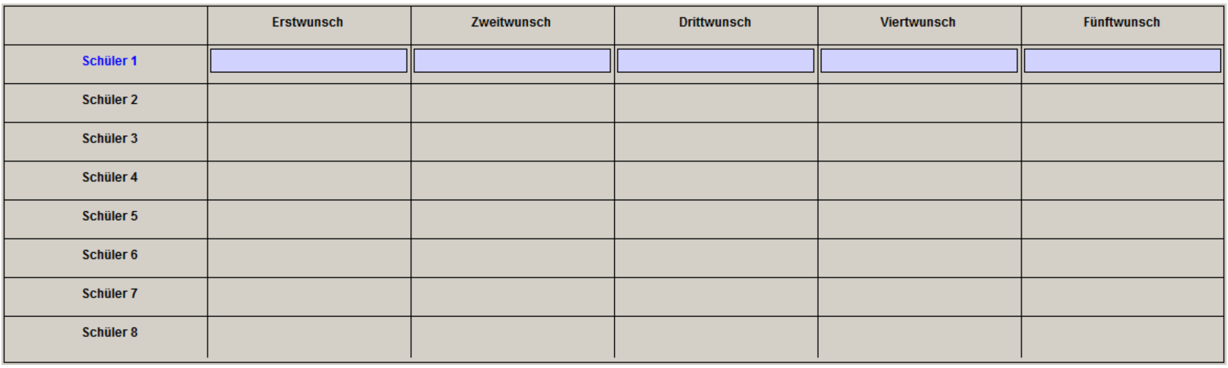
The payout table is the same for all students, i.e. each student receives the highest payout at School A and the lowest payout at School E.

The procedures

Seats at school are allocated as follows. First, you have to enter on-screen which school is your first, second, third, fourth and fifth choice.

The following sample screen shows you how to enter your rank-order list.

Example: Submission of rank-order list



	Erstwunsch	Zweitwunsch	Drittwunsch	Viertwunsch	Fünftwunsch
Schüler 1					
Schüler 2					
Schüler 3					
Schüler 4					
Schüler 5					
Schüler 6					
Schüler 7					
Schüler 8					

In the example, you are student 1. You can click in the row of that respective student on each column field and type in one of the letters A, B, C, D or E, which represent the corresponding schools. You can rank the schools according to their payout, i.e. the rank-order list "first: A, second: B, third: C, fourth: D, fifth: E" or choose any other ranking.

For two randomly selected students on ranks 2 to 5, rank-order lists are generated automatically according to payout, i.e. "first: A, second: B, third: C, fourth: D, fifth: E".

These students do not need to enter anything on the screen; the computer will automatically generate their rank-order list. You will be informed later on the screen for which students the rank-order list will be created automatically according to the payout amount.

Admission to the five schools is based on the submitted rank-order lists and the student numbers.

Procedure 1[19]

Allocation procedure 1 works as follows:

Round 1:

- Application: Each of the students 1 to 8 applies to the school that they have named as their first choice on their rank-order list.
- Admission: If the same number or fewer students apply to a school than there are seats available, all of them receive a seat. If more students have applied to a school than the seats that the school offers, the applicants with the lowest student numbers will be admitted first. For the applicants who have received a seat, the allocation procedure is over.

Round 2:

- Application: Each of the students 1 to 8, who did not get a seat in round 1, applies to the school they have named on their rank-order list as second choice.
- Admission: A school only admits participants in round 2 if there are still free seats after round 1. If the same number or fewer students apply to a school than there are seats available, they will all receive a seat. If more students apply to a school than that school's available seats, the applicants with the lowest student numbers will be

[19] Which procedure is Procedure 1 varies on the session level.

admitted first. For the applicants who have received a seat, the allocation procedure is over.

Round 3

- Application: Each of the students 1 to 8, who did not get a seat in round 2, applies to the school they have named on their rank-order list as a third choice.
- Admission: A school only admits participants in round 3 if there are still free seats after round 2. If the same number or fewer students apply to a school than there are seats available, they will all receive a seat. If more students apply to a school than that school's available seats, the applicants with the lowest student numbers will be admitted first. For the applicants who have received a seat, the allocation procedure is over.

etc.

The process ends when all eight school seats have been allocated.

Procedure 2 (only EFFORT & RANDOM)

In contrast, procedure 2 works as follows:

Round 1:

- The student with the lowest student number (student 1) will be admitted to the school that he has listed as their first choice on their wish list. The allocation process for this applicant is complete.

Round 2:

- Student 2 will be admitted to the school that is highest on their wish list and still has available spots. If all spots at their first choice school are already taken, student 2 will be admitted to their second choice school. The allocation process for this applicant is complete.

Round 3

- Student 3 will be admitted to the school that is highest on their wish list and still has available spots. If all spots at their first and second choice school are already taken, student 3 will be admitted to their third choice school. The allocation process for this applicant is complete.

etc.

The process ends when all eight school seats have been allocated.

Procedure 2 (only ASSORT)

In Procedure 2, students are allocated to schools according to their rank. The allocation is as follows:

Table: Result of Procedure 2

	Seat at school	Payout
Student 1	A	16 Euro
Student 2	B	13 Euro
Student 3	C	10 Euro
Student 4	C	10 Euro
Student 5	D	7 Euro
Student 6	D	7 Euro
Student 7	E	4 Euro
Student 8	E	4 Euro

The table can be read as follows: Student 1 receives a seat at School A and, thus, a payment of 16 Euro, Student 2 receives a seat at School B and 13 Euro, etc.

Further Steps

First, you will be informed on-screen for which of the two randomly selected participants the rank-order lists will be generated automatically. Then you will have the opportunity to try out Procedure 1 *[Treatment EFFORT & RANDOM:* and 2*]* on the screen. Afterwards, you will be able to vote for the allocation procedure to be used. This procedure will then be implemented for your group and your payment will depend on which school you are allocated to.

Before the experiment continues on-screen, please answer the following comprehension questions.

Comprehension Questions (EFFORT & RANDOM)

- How many seats do schools A and B have to offer each?
- How many seats are there at schools C, D, and E?
- What is your payment if you are allocated to school A?
- What is your payment if you are allocated to school E?
- True or false? If the students are assigned a school placement according to their random rank, the allocation looks as follows: Student 1, School A, 16 EUR; (...) Student 8, School E, 4 EUR *[shown in table form]*
- Assuming that all students submit their wish list based on payout amount. That is, all students from 1 to 8 create a wish list "First choice: A, Second choice: B, Third choice: C, Fourth choice: D, Fifth choice: E".
  - In Procedure 1, in this case, the allocation to the schools is done according to the rank. That is, student 1 gets a school placement at School A, student 2 at School B, student 3 at School C, and so on. True/False
  - In Procedure 2, in this case, the allocation to the schools is done according to the rank. True/False
- Assuming that students 1 and 2 submit their wish list according to payout ("First choice: A, Second choice: B, Third choice: C, Fourth choice: D, Fifth choice: E").

Where will student 3 end up if they indicate the wish list ”First choice: B, Second choice: A, Third choice: C, Fourth choice: D, Fifth choice: E”?

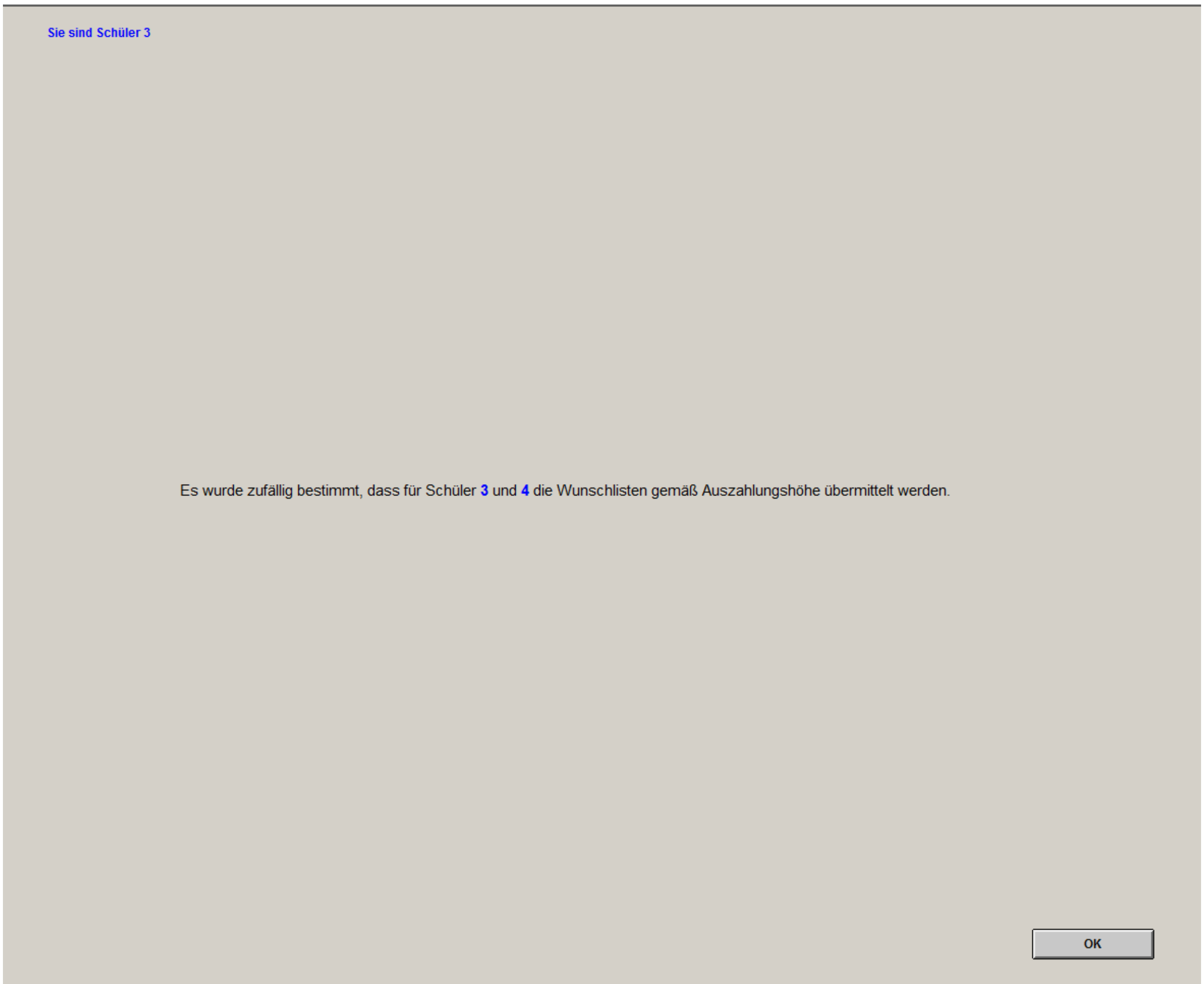
  - In Procedure 1, student 3 will get a spot at school:
  - In Procedure 2, student 3 will get a spot at school:

- Assuming again that student 3 submits the preference list ”First choice: B, Second choice: A, Third choice: C, Fourth choice: D, Fifth choice: E”. Student 1 submits the preference list according to payout.
  - Can student 2 improve their position in Procedure 1 by submitting a different preference list than the one based on payouts at schools? Yes/No
  - Can student 2 improve their position in Procedure 2 by submitting a different preference list than the one based on payouts at schools? Yes/No

Comprehension Questions (ASSORT)

- How many seats do schools A and B have to offer each?
- How many seats are there at schools C, D, and E?
- What is your payment if you are allocated to school A?
- What is your payment if you are allocated to school E?
- In Procedure 2, schools are allocated solely on the basis of the rank achieved in the work task. True/False
- Consider Procedure 1. Suppose students 1 and 2 give their rank-order list according to payout (“first: A, second: B, third: C, fourth: D, fifth: E”). At which school does student 3 end up if he submits “first: B, second: A, third: C, fourth: D, fifth: E” as his rank-order list?

Figure B.6: Screenshot of announcement of ranks

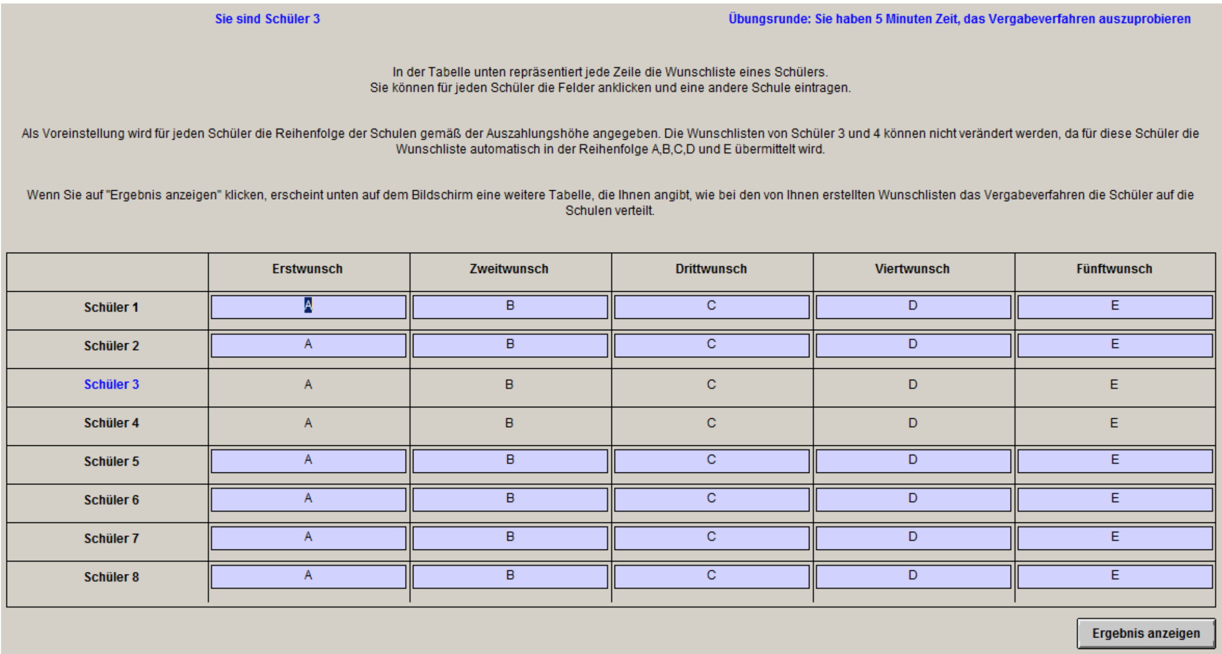


Notes: On this screen, subjects are informed about the students who will automatically submit the rank order lists according to payout amount (i.e., the sincere students). In the Spectator treatment, the subject's rank in the top left corner of the screen is omitted.

Figure B.7: Screenshot of comprehension quiz

Notes: On this screen, subjects have to answer a comprehension quiz about the two procedures (in this example, procedure 1 is the assortative matching and procedure 2 is the Boston mechanism). The comprehension quiz asks for outcomes in different scenarios conditional on two students (in this example, ranks 2 and 3) submitting their sincere rank order list. In the top right corner of the screen, subjects in the Spectator treatment are informed that they receive an additional 5 Euros for answering the quiz correctly. In the top left corner, subjects are informed that this is the original screen that the future subjects will see.

Figure B.8: Screenshot of Trial Period



Notes: The screenshot displays the trial period. In the example, the subject is in rank 3 and is in a group, in which student 3 and 4 have to submit their true preferences. Subjects have five minutes to try out different rank-order lists for non-sincere students. When clicking the button the resulting matching is shown. Subjects in the Spectator treatment see the same screen but without a subject's rank displayed in the top left corner and without her rank being colored in the table. Subjects in EFFORT and RANDOM do the trial period for the Boston and SD mechanisms, subjects in ASSORT only for the Boston mechanism.

B.4 On-screen instructions about voting stage (only spectators)

Now you can decide which school matching procedure will be applied for the future subjects.

The vote works as follows. All subjects in today's session decide on-screen whether they would like to participate in the vote or abstain. Those who take part in the vote can indicate whether Procedure 1 or Procedure 2 is applied for the future subjects.

In a future session, there will be a group of subjects, in which subjects *[2 and 3; 3 and 4; 4 and 5]* automatically submit their rank-order list. This corresponds to the case that you have seen on the previous screens. In today's session you and seven other subjects will take a vote on whether Procedure 1 or Procedure 2 will be applied for this future group of subjects.

In the vote you or one of the other seven subjects are selected at random. If the selected subject voted for one of the procedures, the result of the allocation procedure that he chose will be applied for the future group of subject. His (and only his) payout is reduced by 5 Cent. If the randomly selected subject decided to abstain from the vote, it is determined randomly whether Procedure 1 or Procedure 2 is applied for the future subjects. In this case of abstention, the randomly selected participant does not have to pay the 5 Cent.

When all participants are finished reading the instructions, the experiment continues with a comprehension quiz.

B.5 On-screen instructions about voting stage (only players)

In a previous session there was a vote on which procedure is applied for you and the other subjects.

On the next screen you will be informed which procedure is applied.

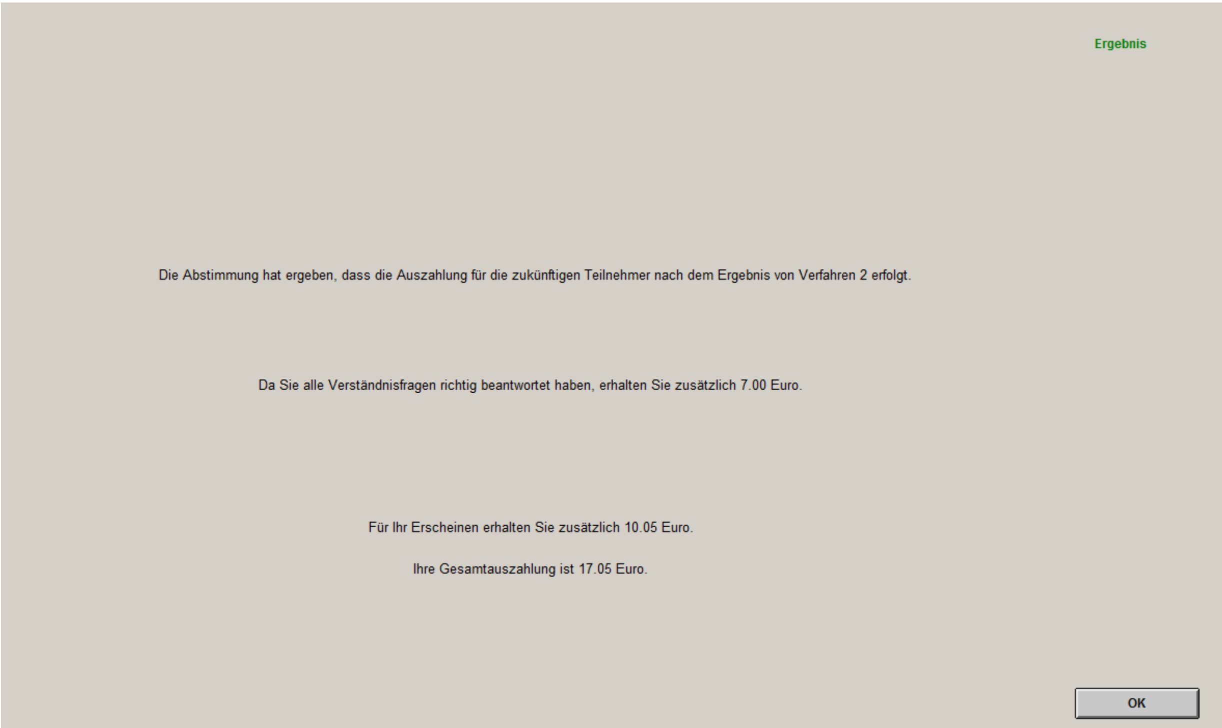
Figure B.9: Screenshot of comprehension quiz about voting

Notes: On this screen, spectators answer a comprehension quiz about the voting stage. The quiz tests comprehension of the random dictator rule and the voting costs of 5 Cents. In the top right corner of the screen, subjects in the Spectator treatment are informed that they can earn an additional 2 Euros for answering the questions correctly.

Figure B.10: Screenshot of voting decision

Notes: On this screen, spectators vote whether the payout for the future subjects is according to Procedure 1 or Procedure 2. Subjects are asked whether they would like to participate in the vote and the options are "Yes, I would like the matching to be according to Procedure 1", "Yes, I would like the matching to be according to Procedure 2", or "No, I do not want to participate in the vote".

Figure B.11: Screenshot of payout screen



Notes: On this screen, spectators are informed which mechanism was selected for the future player session. In addition, they are informed that they receive 10.05 Euros show-up fee and 7.00 Euros for answering all comprehension questions correctly. For spectators, the experiment ends at this point. After players are informed about the selected mechanism, they have to submit their rank-order lists (as described above) before seeing the payout screen.

B.6 Post-experimental questionnaire

We now kindly ask you to answer a few questions. The responses will, of course, be evaluated in an anonymized form.

- Please take a moment to recall the vote regarding whether Procedure 1 or 2 should be implemented. What were your main considerations for voting in favor of one procedure over the other? [Free-text]
- Please think again about the question of whether Procedure 1 or Procedure 2 should be implemented. To what extent do you agree with the following statements? [Do not agree at all / Completely agree, 5 point scale]
  - Fairness played a role in my voting decision on the procedures.
  - I find Procedure 1 [2] to be fairer than Procedure 2 [1].[20]
  - In Procedure 2 [1], students are more likely to receive the payout they deserve compared to Procedure 1 [2].
  - Students with a lower student number should receive a higher payout.
  - I find Procedure 1 [2] easier to understand than Procedure 2 [1].
  - A mechanism should not allow someone to be better off by tactically submitting the wish list at the expense of others.
  - I consider it fair if the students get a school place according to their rank. That means: Student 1 receives a place at School A, Student 2 at School B, Student 3 at School C... and Student 8 at School E.
  - A mechanism should be understandable and transparent for all participants.
  - In Procedure 1 [2], clever behavior is rewarded more than in Procedure 2 [1].

[20] The order of mechanisms was balanced on the session level.

- The ranks of the students are fair.
- In Procedure 1 [2], compared to Procedure 2 [1], those students for whom the preference lists are automatically submitted are unfairly disadvantaged.
- Clever behavior should be rewarded by the mechanism.

• In general, to what extent do you agree with the following statements? [Do not agree at all / Completely agree, 5 point scale]

- Diligence and effort should be rewarded.
- Clever and strategic behavior should pay off.
- Those who work harder should have higher incomes.
- Those who behave cleverly and intelligently should have a higher income.

• Finally, we have some questions about your personal information.

- How old are you?
- Your gender?
- Are you studying?
- What are you studying? (If not studying, please enter 0)
- Different political attitudes are often labeled as “left” and “right.” When thinking about your own political views, where would you classify these views? [Far left / Far right, 10 point scale]