
Fair Procedures with Naive Agents: Who Wants the Boston Mechanism?

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

Discussion Paper No. 222

December 18, 2019

Fair Procedures with Naive Agents: Who Wants the Boston Mechanism?*

Tobias König

Dorothea Kübler

Lydia Mechtenberg

Renke Schmacker

December 15, 2019

Abstract

We study preferences over procedures in the presence of naive agents. We employ a school choice setting following [Pathak and Sönmez \(2008\)](#) who show that sophisticated agents are better off under the Boston mechanism than under a strategy-proof mechanism if some agents are sincere. We use lab experiments to study the preferences of subjects for the Boston mechanism or the assortative matching. We compare the preferences of stakeholders who know their own role with agents behind the veil of ignorance and spectators. As predicted, stakeholders vote for the Boston mechanism if it maximizes their payoffs and vote for the assortative matching otherwise. This is in line with the model of [Pathak and Sönmez \(2008\)](#). Subjects behind the veil of ignorance mainly choose the Boston mechanism when the priority at schools is determined randomly. In a second experiment with priorities based on performance in a real-effort task, spectators whose payoff does not depend on the choice of the mechanism are split in their vote for the Boston mechanism and the assortative matching. According to the spectators' statements in the post-experimental questionnaire, the main reason for preferring the Boston mechanism is that playing the game well deserves a higher payoff. These findings provide a novel explanation for the widespread use of the Boston mechanism.

JEL-codes: D47; C92; I24; D72

Keywords: Matching markets; school choice; voting; Boston mechanism; naive agents; stable assortative matching

*Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged. We would like to thank Axel Ockenfels, Andy Schotter, and Leeat Yariv for helpful comments. The paper has benefited from discussions with participants of the WZB MC seminar.

Tobias König: Linnaeus University, [tobias.koenig\[at\]lnu\[dot\]se](mailto:tobias.koenig[at]lnu[dot]se)

Dorothea Kübler: Technical University Berlin and WZB Berlin Social Science Center, [dorothea.kuebler\[at\]wzb\[dot\]eu](mailto:dorothea.kuebler[at]wzb[dot]eu)

Lydia Mechtenberg: University Hamburg, [lydia.mechtenberg\[at\]uni-hamburg\[dot\]de](mailto:lydia.mechtenberg[at]uni-hamburg[dot]de)

Renke Schmacker: DIW and Free University Berlin, [rschmacker\[at\]diw\[dot\]de](mailto:rschmacker[at]diw[dot]de)

1 Introduction

Mechanisms or procedures to solve allocation problems are used in many domains of life. Examples are central clearinghouses for seats at schools or universities, procedures for the allocation of public housing, and online booking systems for appointments. The complexity of such procedures differs widely, and can affect their public acceptability.¹ We focus on procedures that differ with respect to how difficult it is to make an optimal choice. If mistakes are more often made by people with less information or less education, this can raise issues of equity and fairness. In this paper, we study how people evaluate procedures that differ with respect to the possibility of reaching a better outcome by making strategic choices.

To study preferences over procedures, we build on a debate in the school choice literature regarding the Boston and the deferred acceptance (DA) mechanism (Abdulkadiroğlu and Sönmez, 2003; Gale and Shapley, 1962; Abdulkadiroğlu *et al.*, 2011; Featherstone and Niederle, 2016). While the DA mechanism is strategy-proof, i.e., the best strategy is to report the true preferences over schools, the Boston mechanism often makes it optimal to manipulate the preference list. Pathak and Sönmez (2008) show that in theory the Boston mechanism benefits *sophisticated* students who act strategically at the expense of *sincere* students who always submit their true preferences. In other words, sophisticated parents can reach better outcomes under the Boston mechanism in the presence of sincere parents than under DA.

Apart from this argument, we know very little about people’s preferences over mechanisms.² In this paper, we study whether people prefer a mechanism such as DA that yields the stable assortative matching or the Boston mechanism. We design an experimental school choice problem where students have fully correlated preferences over schools and *vice versa*, and where there are enough seats for all students. Some students in a market are *sincere* as they are forced to submit their true preference ranking. The remaining students are free to submit any preference ranking, i.e., they can be strategic. In the Nash equilibrium of the Boston mechanism, the two sincere students are matched to the two lowest-ranked schools. This benefits all students that have a lower priority and who thereby receive better seats. In contrast, in the stable assortative matching (and the Nash equilibrium in DA), the two sincere students are matched to the schools of their rank. Note that for simplicity, we implemented the assortative matching directly, without relying on preference reports.

We elicit preferences over the two mechanisms by letting subjects vote on which mechanism should be implemented. We consider three different types of voters: subjects in the role of *stakeholders*, i.e., voters who know their rank when making their voting decision, subjects

¹For example, the criticism of the French centralized procedure for university admissions, Parcoursup, focused on its intransparency and complexity, among other issues. See https://www.lepoint.fr/education/parcoursup-2e-version-en-progres-16-05-2019-2313092_3584.php.

²One exception is Schmelzer (2016) who compares preferences regarding single and multiple tie-breaking in school choice.

behind a veil of ignorance, i.e., subjects who do not know their rank when voting, and *spectators*, i.e., subjects who vote for a mechanism that will be implemented for another group of subjects.³

We hypothesize that stakeholders tend to vote for the mechanism that yields the higher payoff for them, following [Pathak and Sönmez \(2008\)](#). Thus, participants who can choose their preference list freely are predicted to vote for the Boston mechanism while those who have to submit their true preferences vote for the assortative matching. Some participants may have procedural preferences that deflect their choices away from the selfish prediction. However, the precise direction of such deviations is hard to predict: On the one hand, people may dislike the strategic environment of the Boston mechanism since previous research has shown that people have a preference for truth-telling ([Abeler et al., 2019](#)). On the other hand, they may prefer the ability to make choices in the Boston mechanism, e.g., because they are overconfident and believe that they understand the mechanism better than others.

Regarding the spectators, we hypothesize that fairness concerns play a role for their choice. Since the rank of subjects matched to schools is determined by their performance in a real-effort task performed in the lab, spectators may have meritocratic reasons for preferring the stable assortative matching that gives higher payoffs to subjects who performed better in the task.

The veil of ignorance constitutes an intermediate situation. Given that the distribution of payoffs is fixed across mechanisms and given that subjects do not know their rank behind the veil of ignorance, the main difference between the Boston mechanism and the assortative matching is which payoffs are allocated to participants at different ranks. Thus, we hold constant outcome-based fairness in equilibrium but vary the source of inequality in the payoffs. We therefore expect subjects behind the veil of ignorance to vote for the assortative matching if they place some value on the ranks or abstain.

We find that the voting of stakeholders is in line with the predictions of [Pathak and Sönmez \(2008\)](#): those stakeholders who can choose freely tend to vote for the Boston mechanism. By contrast, those who have to submit their true preferences tend to vote for the assortative outcome. Thus, the preferences of stakeholders for the Boston mechanism are in line with their monetary incentives. More surprisingly, the majority of subjects behind the veil of ignorance also vote for the Boston mechanism instead of the assortative matching.

For the spectators, we find that a majority of them prefers the Boston mechanism, and only a small fraction abstains. In the experiments, we varied the performance rank of the students who are forced to submit their true preferences. We observe that the spectators react to this information. There is suggestive evidence that more spectators vote for the Boston

³Previous experiments have employed the veil of ignorance (e.g. [Schildberg-Hörisch, 2010](#)) and spectator designs (e.g. [Konow, 2000](#); [Cappelen et al., 2013](#); [Møllerstrom et al., 2015](#)) to study preferences for redistribution. To our knowledge, we are the first to employ these designs for the choice of mechanisms.

mechanism when those who are forced to be sincere are the lower-performing students. Thus, the perceived fairness of the allocation procedure seems to matter for spectators’ choices. Another aspect of fairness emerges from the free-form answers in the post-experimental questionnaire according to which many subjects consider the outcome of a strategic game such as the Boston mechanism as fair.

The experiment contributes to a small, but growing literature on behavioral effects in matching markets.⁴ [Pathak and Sönmez \(2008\)](#) show that sophisticated players have an incentive to support a non-strategy-proof matching algorithm, since this allows them to exploit naive players who do not act strategically. Our paper can be seen as the first empirical test of this idea: imposing the existence of non-strategic players in a controlled laboratory environment, we find evidence that the remaining players understand the incentives of the Boston mechanism to game the system, and prefer this system over a matching system that directly implements the assortative matching outcome. [Basteck and Mantovani \(2018\)](#) show with the help of experiments that payoff differences between high and low- cognitive-ability participants are smaller under DA than under Boston, and that DA is more effective in preventing low-ability individuals from ending up at the worst schools. In contrast, we let stakeholders and spectators vote over mechanisms—a feature that is absent in [Basteck and Mantovani \(2018\)](#). Moreover, we impose sincere preference reports by some subjects instead of measuring cognitive ability.

The paper builds on a recent literature in experimental economics which uses a spectator approach to study voting behavior in isolation of any self-interest and personal involvement (see, e.g. [Konow, 2000](#); [Cappelen et al., 2013](#); [Møllerstrom et al., 2015](#)). In these studies, individuals vote over different income allocations in order to address questions of *distributional* justice. In contrast, income distributions are held constant in our paper. Instead, we let individuals decide over the mechanisms governing how subjects are assigned the various school seats or monetary payoffs. We thus contribute to a better understanding of preferences about *procedural* justice. For such preferences over procedures, notions of transparency, impartiality, and equality of opportunity are well established. However, our results reveal that many individuals value mechanisms that leave room for clever behavior. To our knowledge, a preference for the possibility to make strategic choices has not been identified so far.

Our paper goes beyond [Pathak and Sönmez \(2008\)](#) by varying the source of priorities at schools and the degree to which decision makers are involved in the school choice procedure. First, we show that the preference of sophisticated players for the Boston mechanism survives in contexts where ranks are merit based, that is, earned in a competitive environment—which is not considered explicitly in the theoretical framework of [Pathak and Sönmez \(2008\)](#). Second, we test whether behavioral factors shape the preferences for matching mechanisms. We demonstrate that overconfidence regarding the ability to game the system cannot explain

⁴For a survey of matching experiments see [Hakimov and Kübler \(2019\)](#).

our results. We identify a direct preference for matching mechanisms, largely unnoticed so far: A considerable share of individuals vote for the Boston mechanism, since they consider the outcome of a strategic game as fair. This observation appears to be robust, since it also emerges when we completely switch off any personal stakes in the mechanisms: Impartial spectators want others to play the Boston mechanism because they consider strategic choices as a source of entitlement.

The finding that many participants prefer the strategic environment adds a new twist to explaining the widespread use of the Boston mechanism in spite of its problematic properties. [Pathak and Sönmez \(2008\)](#) argue that sophisticated players may form political coalitions (e.g, parent clubs) to lobby for the Boston mechanism. However, for this argument to work, sophisticated players need to find a political majority to enforce their interests, which is questionable if naive players learn that they tend to lose in the Boston mechanism. Our experiments demonstrate that individuals not directly affected by the matching market tend to support the Boston mechanism for reasons of fairness, making majorities in favor of it more likely.

We acknowledge that questions of external validity arise. In particular, we can only speculate about the preference for non-strategy-proof mechanisms with high-stake outcomes outside of the lab. However, our findings suggest that the strength of the preference for non-strategy-proof mechanisms interacts with performance-based entitlements: We observe less voting for the Boston mechanism the more it hurts individuals on higher performance ranks. Therefore, we would predict the preference for manipulable mechanisms to be more relevant in settings where matching priorities are less tied to merit, for example when priorities at schools are not based on ability. In such instances, strategic play can serve as a substitute for ability in the fairness perceptions.

In the next section, we introduce some basic concepts from the theory of matching. We conducted two experiments, Experiment 1 and 2, that are presented in Section 3. Section 4 concludes.

2 Matching markets: basic concepts

In this section, we introduce concepts of matching theory that will be used to characterize the experimental school choice problem. A matching is called *stable* if two conditions are satisfied: First, every agent prefers the assigned matching partner to remaining unmatched, i.e., the student is matched to a school that she prefers to being unmatched, and the school is only matched to acceptable students. Second, there is no school-student pair such that each prefers one another to their respective match. Thus, in a stable matching no student and school have an incentive to contract around the clearinghouse and form a different match. The concept of stability can be interpreted as a notion of fairness, namely that no student

experiences justified envy towards another student: it cannot be the case that a student prefers the school of another student and has a higher priority at that school.

In our setup, all students rank schools in the same way, i.e., they agree on the best school, the second-best school, etc. At the same time, all schools rank students based on their performance rank. 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.

The incentives of students to submit their rankings of schools are an important property of matching mechanisms. Moreover, it matters how complicated it is for agents to submit a ranking that is optimal. *Strategy-proofness* means that truthful preference revelation is a (weakly) dominant strategy for agents. If a mechanism is strategy-proof for the students, they cannot gain from misreporting their preferences, independent of the choices of other students.

Our experiments employ the *Boston mechanism* which works as follows:

Step 1: Each student applies to the school that she ranks first. Each school admits acceptable students up to its capacity, in the order of the performance ranks. 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 most-preferred acceptable school among the schools to which the student has not yet applied. Each school admits acceptable students up to its remaining capacity, according to their performance ranks. These assignments are final. The remaining students are rejected.

The algorithm terminates when no student is rejected, or all schools have filled the seats up to their capacity. All remaining students are unassigned.

The Boston mechanism is not strategy-proof for the students. 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. If students tell the truth, unstable outcomes can occur where a student prefers a seat at a school where she has a higher priority than a student who was admitted. In Nash equilibrium, the outcome of the Boston mechanism is stable (Ergin and Sönmez, 2006). However, the equilibrium requires strategic play by the students. On the other hand, if all students report their preferences truthfully, the Boston mechanism leads to a Pareto efficient and potentially unstable allocation.

3 Experiments

The experiments are designed to study preferences over mechanisms. We employ stylized markets in order to make it transparent that sincere subjects are harmed by the Boston mechanism if other players act strategically. The alternative to the Boston mechanism is the assortative matching that is implemented directly, without preference reports, for the sake of

simplicity.⁵ In our matching problem, the sophisticated players earn higher payoffs under the Boston mechanism than in the assortative matching while the opposite holds for the sincere players if the equilibrium is played in the Boston mechanism. By contrast, if all subjects report their preferences truthfully in the Boston mechanism, both mechanisms lead to the same allocation of payoffs. We also investigate the preferences of subjects behind the veil of ignorance in order to capture their ex ante preferences. Finally, we let spectators decide on one of the two mechanisms that is then applied in an experimental session.

3.1 Experiment 1

The first experiment studies the choices of stakeholders who know their rank and of stakeholders behind the veil of ignorance.

3.1.1 Design of Experiment 1

Priorities of students At the beginning of the treatment, the students are randomly assigned their rank. These ranks determine the priority of students at schools where a better rank means a higher priority (rank 1 is best etc.). Students are informed about their rank.

Matching market We implement the simplest possible market where strategic applicants can benefit from sincere applicants in the Boston mechanism. Although the setup is stylized, this allows us to study preferences over mechanisms in a transparent setup. The experimental matching problem consists of eight students and five schools. Schools A and B have one seat each while schools C, D, and E have two seats each. Students have fully correlated preferences over schools. The priority of students at schools is determined by their rank, and is therefore also perfectly correlated. Students participate in a school admissions game. A participant earns 12 Euro if she receives a seat at school A, 10 Euro at school B, 8 Euro at school C, 6 Euro at school D, and 4 Euro at school E. The assignment of students to schools is either based on the Boston mechanism, or the stable assortative matching is implemented. The assortative matching implies that the student of rank 1 is admitted by school A, the student of rank 2 is admitted by school B etc.:

$$(1) \quad ((1, A), (2, B), (3, C), (4, C), (5, D), (6, D), (7, E), (8, E)).$$

Nash equilibrium There is a unique Nash equilibrium outcome of the game induced by the Boston mechanism if all eight students can submit their preference lists strategically. This equilibrium outcome is the assortative matching. To see this, assume that students who know

⁵Alternatively, we could have implemented a serial dictatorship mechanism which yields the assortative matching outcome in our market.

their rank list the school first that they would be matched to in the assortative matching. This is an equilibrium since no profitable deviation exists.⁶ However, if students 2 and 3 are forced to submit their preferences truthfully, student 4 can get into school B by listing it first, student 5 can get into school C by listing it first etc. such that students 2 and 3 end up at the least attractive school E.⁷

The resulting matching is

$$(2) \quad ((1, A), (2, E), (3, E), (4, B), (5, C), (6, C), (7, D), (8, D)).$$

Analogous equilibrium outcomes result if students 3 and 4 or 4 and 5 are sincere where all students with ranks below the sincere students move up to a better school while the sincere students are admitted to school E.

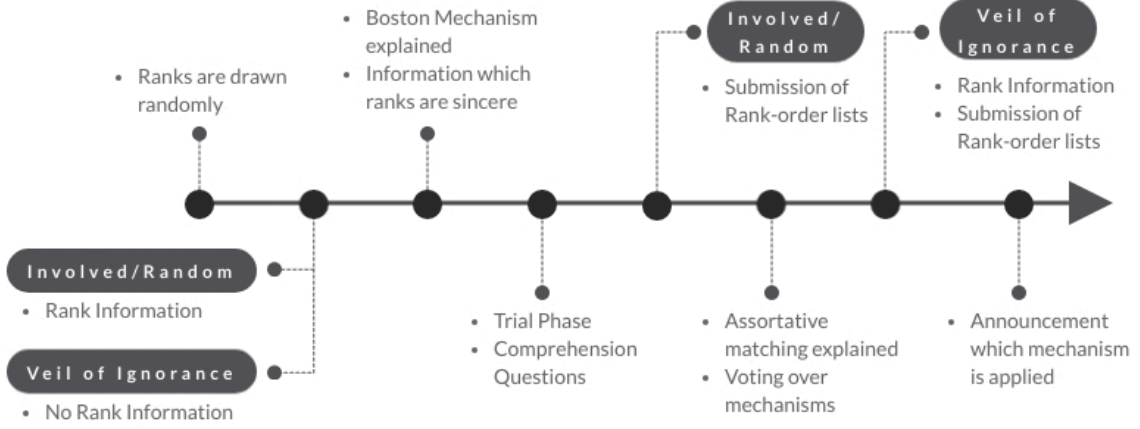
Mechanisms First, the Boston mechanism is introduced and explained. The students in ranks 2 and 3 cannot choose freely, but are forced to submit their true preferences. We decided to impose sincerity for students on ranks 2 and 3 in order to generate a predicted difference in outcomes between the two mechanisms for as many subjects as possible. Only for the student in rank 1 the equilibrium outcome in the Boston mechanism is equivalent to her assortative matching outcome. The students are informed that students in ranks 2 and 3 have to submit their true rank-order list. Then, the students can explore the Boston mechanism with the help of an on-screen tool that lets them simulate the outcomes for different rank-order lists. After this trial phase, they complete a quiz to make sure they have understood the Boston mechanism. Finally, they are asked to submit rank-order lists of the schools. Besides their own rank-order lists, we also elicit their beliefs about the first choice of other students. If the expectations about all other students' first choices are correct, they earn 2 Euros.

Afterwards, the assortative matching is displayed (student of rank 1 receives a seat at school A etc.). In the stable assortative matching, the two sincere students are matched to the schools according to their rank. However, under the Boston mechanism they are matched to the lowest-ranked school E if all other students play the equilibrium strategy. Thus, the two *sincere* students earn higher payoffs in the assortative matching than in the outcome reached under the Boston mechanism. On the other hand, the students who can report their preferences freely either earn the same payoff as in the Boston mechanism (those with a higher priority than the sincere students), or are better off in the Boston mechanism (those with a worse priority than the sincere students).

⁶Note that there are multiple equilibrium strategies since the ranking of schools after the first do not matter. Moreover, students 7 and 8 can submit any ranking as long as school E is on their list.

⁷Students 7 and 8 can submit any list as long as school D is on the first, second, or third position of their rank-order list.

Figure 1: Timeline of Experiment 1



Voting Finally, students vote on which outcome to implement: the outcome implied by the game under the Boston mechanism that they just participated in, or the assortative matching. The voting outcome is determined by the random-dictator rule. Subjects are informed which mechanism is chosen to determine the payoffs.

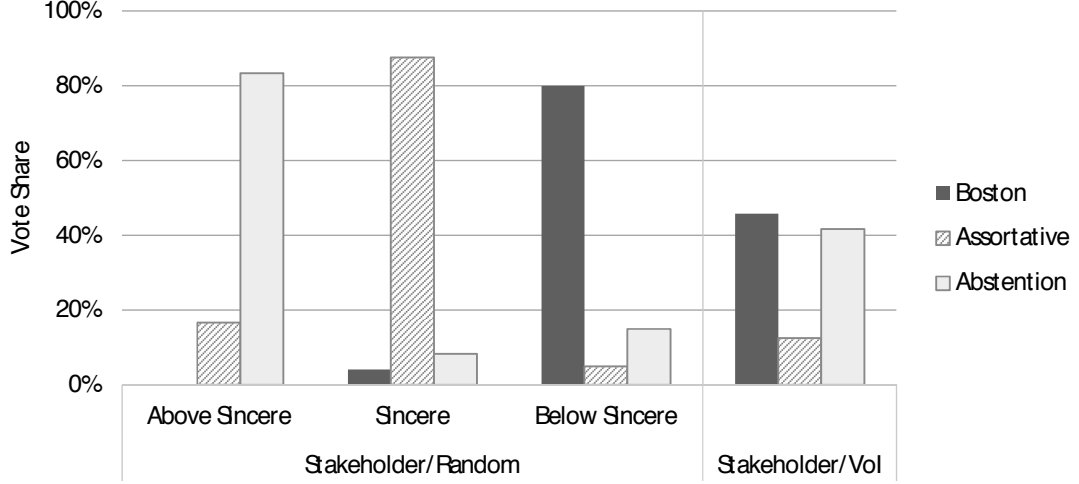
Treatments and hypotheses Figure 1 shows the timeline of Experiment 1. The only difference between the treatments concerns the time at which subjects receive information about their rank and at which they submit their rank-order list. Treatment Stakeholder/VoI differs from Stakeholder/Random with regard to the order of steps: the announcements of the subjects' ranks, the elicitation of beliefs regarding other subjects' first choices, and the submission of preference lists occur after the voting stage. Hence, in Stakeholder/VoI the subjects vote behind a veil of ignorance.

We formulate the following hypothesis:

Hypothesis 1: *Subjects in Stakeholder/Random vote for the mechanism that yields a higher payoff for them. Thus, sincere subjects vote for the assortative matching while subjects at ranks 4 to 8 vote for the Boston mechanism.*

Note that behind the veil of ignorance, both mechanisms have the same distribution of payoffs ex ante, since ranks are determined randomly. However, we hypothesize that subjects either abstain from costly voting or prefer the assortative matching that honors the priorities of students at the schools.

Figure 2: Vote shares in Stakeholder/Random and Stakeholder/VoI



Notes: The figure shows the vote share for the respective procedure. In treatment Stakeholder/Random we divide the sample into ranks above sincere (n=12), sincere (n=24), and below sincere (n=60), and for Stakeholder/VoI we show the full sample (n=48).

Hypothesis 2: *Subjects behind the veil of ignorance vote for the assortative matching or abstain from voting.*

Overall, 96 subjects participated in the Stakeholder/Random treatment and 48 subjects in the Stakeholder/VoI treatment.⁸ Subjects received 5 Euro show-up fee. Voting costs were 10 cents. Subjects earned on average 12.28 Euro with the sessions lasting on average 60 minutes.

3.1.2 Results of Experiment 1: Voting

In Figure 2, we show the vote shares in both the Stakeholder/Random and the Stakeholder/VoI treatment. Since in Stakeholder/Random the incentives to vote depend on the rank, we show the vote shares separately for subjects who have to submit their true preference lists (sincere), and for subjects who are on the ranks above and below the sincere students.

In Stakeholder/Random we observe that subjects vote in line with their self-interest. The majority of subjects in the first rank who are indifferent (“above sincere”) abstain from voting and do not pay the voting costs. The majority of sincere students who benefit from the assortative matching in equilibrium, vote for the assortative matching. Moreover, the major-

⁸Table A.1 in the appendix presents some descriptive statistics of the sample.

Table 1: Vote shares in Stakeholder/VoI and Stakeholder/Random

Treatment	Vote share			H0: Vote shares equal to		
	Boston	Assortative	Abstention	VoI	Above sincere	Sincere
Stakeholder/VoI						
Total	45.8%	12.5%	41.7%			
Stakeholder/Random						
Above Sincere	0.0%	16.7%	83.3%	p=0.005		
Sincere	4.2%	87.5%	8.3%	p=0.000	p=0.000	
Below Sincere	80.0%	5.0%	15.0%	p=0.001	p=0.000	p=0.000

Notes: This table shows the fraction of subjects who vote for the Boston mechanism, the assortative matching, or abstain. For the Involved/Random Treatment, we show separate vote shares for sincere students and students whose ranks are above or below the sincere students. In the right panel, we show results from hypothesis tests that compare the vote shares between the groups using Fisher’s exact tests.

ity of subjects who benefit from the Boston mechanism because they have ranks below the sincere students (“below sincere”), vote for the Boston mechanism. The Fisher’s exact tests in Table 1 confirm that the vote shares between the three groups are significantly different from each other. Thus, the findings support Hypothesis 1:

Result 1: *A large majority of subjects in Stakeholder/Random vote for the mechanism with the higher individual payoff. Thus, most sincere subjects vote for the assortative matching while most subjects at ranks 4 to 8 vote for the Boston mechanism.*

In the treatment with a veil of ignorance (Stakeholder/VoI), we observe that 45.8% of subjects (78.6% of voters) vote for the Boston mechanism while only 12.5% (22.4% of voters) vote for the assortative matching. The remaining 41.7% of subjects abstain from voting. A binomial test confirms that significantly more subjects vote for the Boston mechanism than for the assortative matching ($p=0.004$). Moreover, the hypothesis tests in Table 1 show that the distribution of votes in Stakeholder/VoI is significantly different from the three groups in the Stakeholder/Random Treatment. Thus, in contrast to Hypothesis 2 we find:

Result 2: *Behind the veil of ignorance, almost half of the subjects vote for the Boston mechanism. A slightly smaller proportion abstains, and only 12.5% vote for the assortative matching.*

Hence, in Stakeholder/VoI we observe that 58.3% of subjects are willing to pay the voting cost to affect which mechanism is implemented, and that the majority of voters prefer the Boston mechanism. Note that behind the veil of ignorance the distribution of payoffs is exactly identical in the assortative matching and under the Boston mechanism. However,

Table 2: Equilibrium choices in Stakeholder/Random and Stakeholder/VoI

Rank	Stakeholder/Random	Stakeholder/VoI	Fisher’s exact test
Above Sincere (Rank 1)	12/12 (100.0%)	5/6 (83.3%)	p=0.333
Below Sincere (Rank 4-6)	27/36 (75.0%)	12/18 (67.0%)	p=0.536
Below Sincere (Rank 7-8)	21/24 (87.5%)	11/12 (91.7%)	p=1.000
Total	60/96 (83.3%)	28/36 (77.8%)	p=0.600

Notes: This table shows the fraction of subjects who play the equilibrium strategy in the Boston mechanism by treatment. Best response means for students in rank 1 to list school A as first choice, for students in rank 4 to list school B first, for students in ranks 5 and 6 to put school C first, and for students in ranks 7 and 8 to list school C as first, second or third choice.

subjects may believe that they understand the Boston mechanism better than other participants, therefore expecting higher payoffs than in the assortative matching. To investigate this argument further, we analyze how subjects actually play the Boston mechanism and their beliefs regarding other subjects’ choices.

3.1.3 Results of Experiment 1: Preference lists and outcomes of Boston mechanism

Table 2 shows how many students choose the equilibrium strategy.⁹ We observe a large fraction of subjects exhibiting equilibrium behavior: 83.3% in Stakeholder/Random and 77.8% in Stakeholder/VoI with no significant differences between the treatments. Among the subjects with ranks below the sincere students, only one out of 60 in Stakeholder/Random and two out of 30 subjects in Stakeholder/VoI reported their preferences truthfully. Hence, the instructions and the trial phase were successful in explaining the game, and subjects do not display a strong preference for truthtelling.

Next, we investigate the expectations about others’ behavior in the Boston mechanism. If subjects believe that others make mistakes and do not play the equilibrium strategy, this could potentially explain the high vote share for Boston behind the veil of ignorance. In Table 3, we show the fraction of subjects who expect other students to list the first choice predicted in equilibrium.¹⁰ The majority of subjects in both treatments expect other students to submit first preferences that are in line with Nash equilibrium. We can also consider only those subjects who play the equilibrium strategy themselves, thereby exhibiting that they

⁹Since students in ranks 2 and 3 are forced to submit their true preferences, it is the best response by the subject in rank 1 to list A first, in rank 4 to list B first, in ranks 5 and 6 to list C first and in ranks 7 and 8 to list D at the first, second, or third position on the list.

¹⁰We have elicited expectations about the first choice of other students. For students on ranks 1 to 6 the first choice determines equilibrium behavior, hence, we show these ranks in Table 3. To determine if subjects on ranks 7 and 8 exhibit equilibrium behavior, we would also need expectations about their second and third choice.

Table 3: Expectations of other subjects’ equilibrium choices

Rank	Stakeholder/Random	Stakeholder/VoI	
		Total	Own Best Response
Above Sincere (Rank 1)	75/84 (89.3%)	38/42 (90.5%)	22/23 (95.7%)
Below Sincere (Rank 4)	63/84 (75.0%)	32/42 (76.2%)	22/24 (91.7%)
Below Sincere (Rank 5)	58/84 (69.0%)	32/42 (76.2%)	22/25 (88.0%)
Below Sincere (Rank 6)	57/84 (67.9%)	32/42 (76.2%)	21/23 (91.3%)

Notes: This table shows the reported expectations about other subjects’ equilibrium behavior regarding their first choice. This means for students in rank 1 to list school A first, for students in rank 4 to list school B first, and for students in ranks 5 and 6 to put school C first. The last column is based on subjects in the Veil of Ignorance treatment who play the equilibrium strategy themselves. Note that students 7 and 8 list school D first, second, or third in equilibrium but we did not elicit beliefs about second or third choices.

understand the equilibrium. Among those subjects (“Own Equilibrium Strategy”) on average about 87% expect others to choose the equilibrium strategy.

The elicited beliefs limit the scope for explaining the high vote share for Boston in the Stakeholder/VoI treatment with the expectation that other students fail to play the equilibrium strategy. Instead, subjects could have a procedural preference for the Boston mechanism. A possible reason is that the ranks in Experiment 1 are determined randomly, which does not create strong entitlements, and that gaming the Boston mechanism is seen as meriting a high payoff. In order to investigate this possibility more thoroughly, we run Experiment 2.

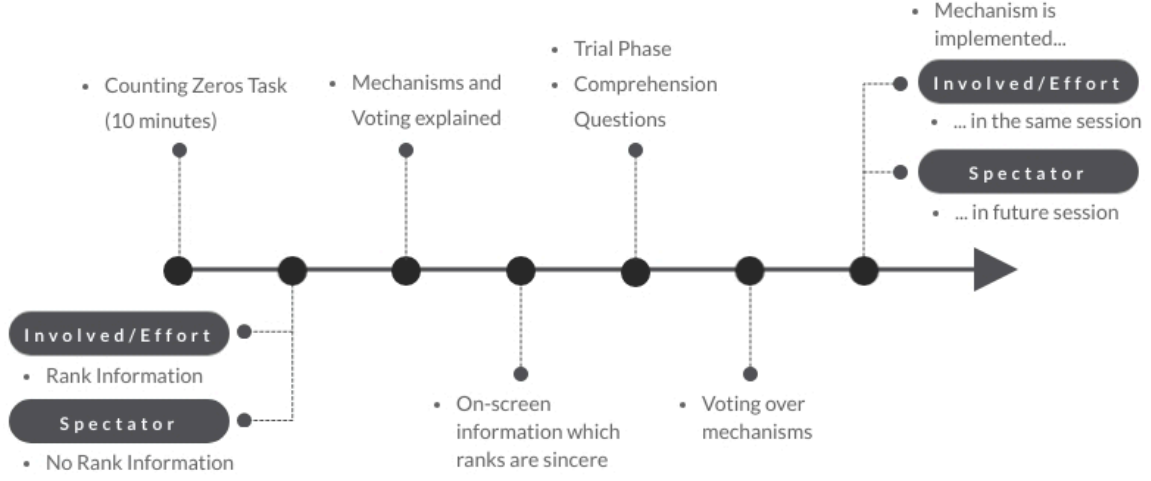
3.2 Experiment 2

It is conceivable that the ranks of students in Experiment 1 that were determined randomly are not perceived as justifying the higher payoffs of higher-ranked students in the assortative matching. This could explain why the Boston mechanism is chosen by the majority of stakeholders behind the veil of ignorance. In Experiment 2, the students’ ranks are based on the performance in a real-effort task. Moreover, we randomly determined which pairs of students are sincere, either 2 and 3, 3 and 4, or 4 and 5. This ensures that subjects have an incentive to perform well in the real-effort task, and provides some variation in the ranks of subjects who are harmed by the Boston mechanism vis-à-vis the assortative matching.

3.2.1 Design of Experiment 2

Priorities of students The rank is no longer determined randomly. Instead, the students have to count zeros in matrices consisting of zeros and ones for 10 minutes in the first part of the experiment. The task is based on [Abeler *et al.* \(2011\)](#), and a screenshot is provided in Figure B.2 in the Appendix. During these 10 minutes, subjects are allowed to browse the internet by clicking an on-screen button. This outside option was meant to make the

Figure 3: Timeline of Experiment 2



ranks even more informative and legitimate, since subjects had an attractive alternative to working. Upon completion of the task, participants are informed of their performance rank in the group.

Mechanisms and voting The Boston mechanism and the assortative matching are presented in counterbalanced order. After the Boston mechanism and the assortative matching are explained, the subjects can vote for one of the two mechanisms. If the Boston mechanism is selected with the random dictator rule, the subjects submit their preference lists for the Boston mechanism.

Treatments and hypotheses As in Experiment 1, there are two treatments: Stakeholder/Effort and Spectator. In Stakeholder/Effort, subjects vote on the mechanism that will assign them to schools, similarly to subjects in Stakeholder/Random (Experiment 1). Figure 3 shows the timeline of Experiment 2. After being introduced to the Boston mechanism with the help of the simulation and the quiz, subjects are asked to vote for the Boston mechanism or the assortative matching. Voting costs 0.05 Euro while subjects have to pay nothing in case they abstain. As in Experiment 1, this small cost of voting prevents those subjects from voting who are indifferent between the two options. The computer randomly picks one of the subjects and implements her choice. If a subject is picked who abstained, a random draw selects the procedure. If the Boston mechanism is selected, subjects in the player session are asked to submit their preference lists for the Boston mechanism. Finally, participants are informed about the result of the vote and their payoffs.

For treatment Stakeholder/Effort we hypothesize as in Experiment 1:

Hypothesis 3: *Subjects in Stakeholder/Effort vote for the mechanism that gives them a higher payoff in equilibrium. Thus, sincere subjects vote for the assortative matching while subjects at ranks 4 to 8 vote for the Boston mechanism.*

The Spectator treatment differs from Stakeholder/Effort in that participants vote on the mechanism that will assign *other* subjects to schools in a later session. These sessions are called Player sessions. Thus, the voting outcome in the Spectator session affects a group of experimental subjects who do not vote themselves. For each group of eight spectators, we ran a session of eight subjects. After the spectators made their decisions, we picked one of them randomly to determine the allocation mechanism in the session that took place afterwards. Thus, a spectator was pivotal with a probability of $\frac{1}{8}$, just as the applicants in the Stakeholder/Effort treatment. Since the mechanism that is selected in the Spectator treatment is not payoff-relevant for the spectators, we interpret their decisions as reflecting fairness preferences. Since ranks are earned in a real-effort task, merit-based fairness leads to the following hypothesis:

Hypothesis 4: *Spectators vote for the assortative matching or abstain from voting.*

In total, there are 96 subjects each in the Stakeholder/Effort, Spectator, and Player sessions.¹¹ Participants received a show-up fee of 8.05 Euro in the Stakeholder/Effort treatment and 10.05 Euro in the Spectator treatment. The experimental sessions of both treatments lasted on average 60 minutes. Average earnings (including the show-up fee) were Euro 16.92 in the Stakeholder/Effort and 17.05 Euro in the Spectator treatments.

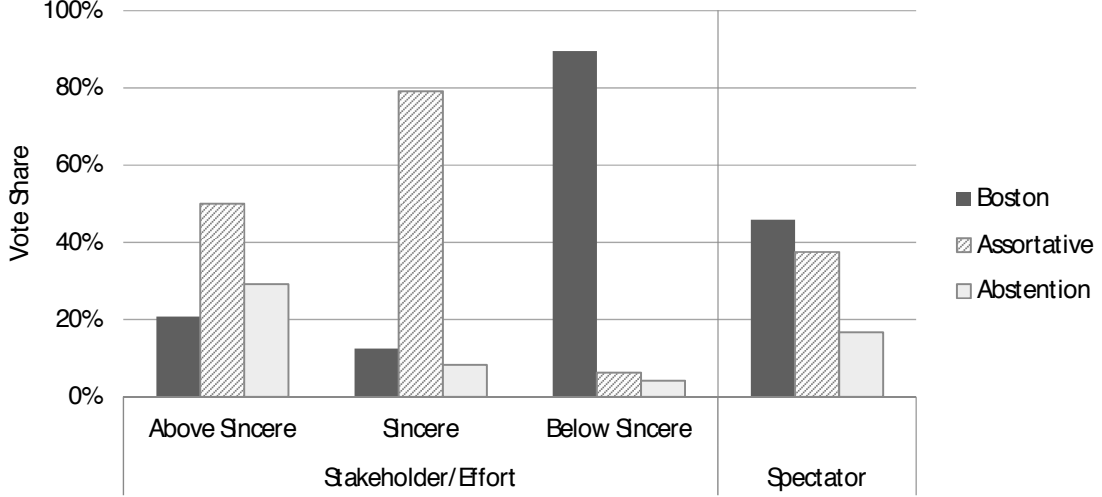
3.2.2 Results of Experiment 2: Voting

The share of votes for the Boston mechanism and the assortative matching are displayed in Figure 4, as well as abstentions. Similar to Experiment 1, the stakeholders vote mainly for the mechanism that yields the highest payoff for them. While almost 79.2% of those who are forced to be sincere vote for the assortative matching, 89.6% of those who are on the lower ranks than the sincere students vote for the Boston mechanism. In Experiment 2, only 29.2% of students with better ranks than the sincere students (“above sincere”) abstain although their payoffs are the same under both mechanisms, 50.0% of them vote for the assortative matching while 20.8% vote for the Boston mechanism.

Hypothesis tests suggest that the vote shares of “below sincere” students are significantly different from the other two groups but that the vote shares of “sincere” and “above sincere” are only marginally significantly different from each other. In sum, our findings support Hy-

¹¹Table A.1 in the appendix presents descriptive statistics of the sample.

Figure 4: Vote shares in Stakeholder/Effort and Spectator



Notes: The figure shows the vote share for the respective procedure. In Stakeholder/Effort we divide the sample into ranks above sincere (n=24), sincere (n=24), and below sincere (n=48), and for Spectator we show the total sample (n=96).

pothesis 3:

Result 3: *A large majority of subjects in Stakeholder/Effort vote for the mechanism that yields a higher payoff for them. Thus, most sincere subjects vote for the assortative matching while most subjects at ranks 4 to 8 vote for the Boston mechanism.*

Spectators favor the Boston mechanism and the assortative matching in similar proportions. Note that only 16.7% abstain and save the voting costs. 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. The hypothesis tests in Table 4 show that the voting pattern in Spectator is significantly different from the “sincere” and “below sincere” students in treatment Stakeholder/Random but there is only a marginally significant difference to the subjects ranked “above sincere.”

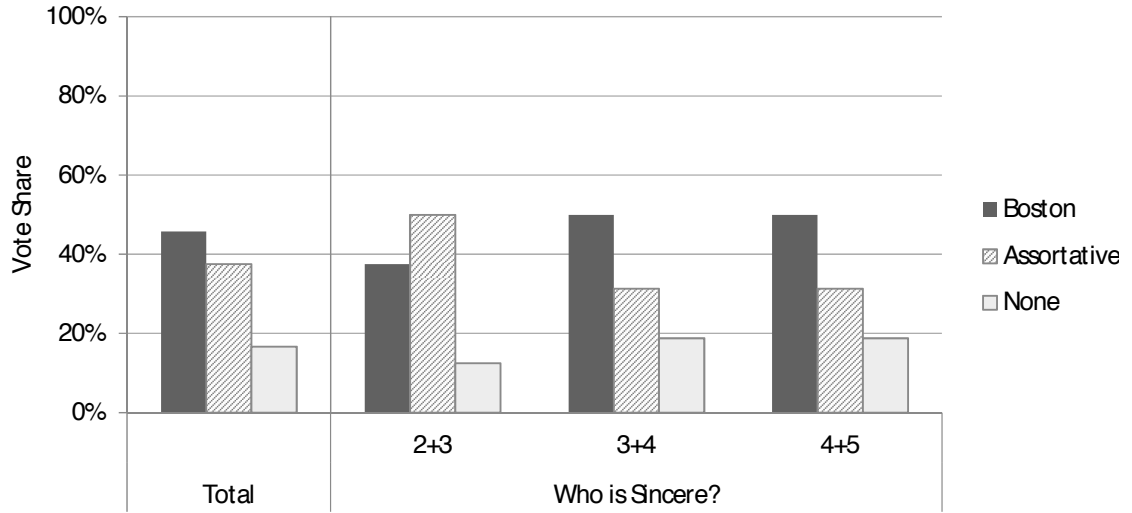
Since we varied which subjects are forced to be sincere, we can study the spectators’ reaction to this. In Figure 5, we show suggestive evidence that spectators are more likely to vote for the assortative matching when sincere students have ranks 2 and 3 than when they have lower ranks: In the treatment where sincere students have ranks 2 and 3, the share of spectators voting for Boston is 37.5%, and thus 12.5 percentage points lower, compared to

Table 4: Vote shares in Stakeholder/Effort and Spectator

Treatment	Vote share			H0: Vote shares equal to		
	Boston	Assortative	Abstention	Spectator	Above sincere	Sincere
Spectator						
Total	45.8%	37.5%	16.7%			
Stakeholder/Effort						
Above Sincere	20.8%	50.0%	29.2%	p=0.062		
Sincere	12.5%	79.2%	8.3%	p=0.001	p=0.092	
Below Sincere	89.6%	6.3%	4.2%	p=0.000	p=0.000	p=0.000

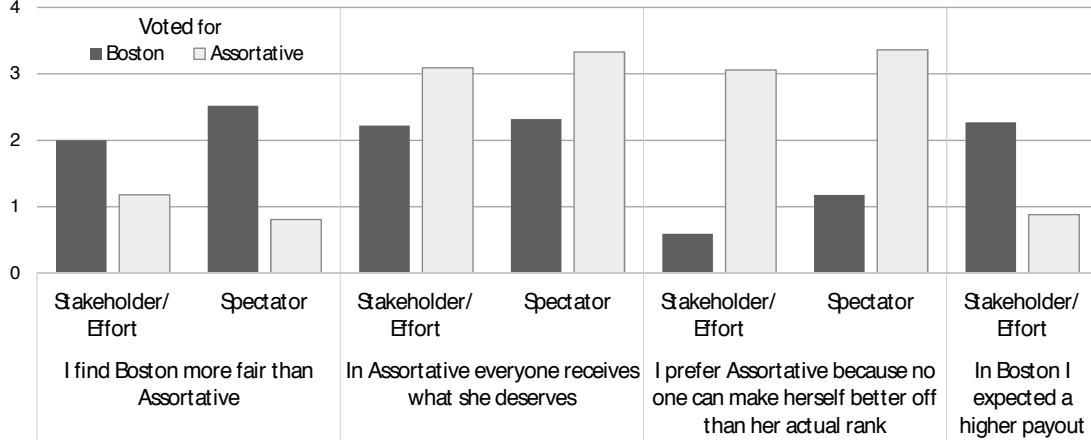
Notes: This table shows the fraction of subjects who vote for the Boston mechanism, the assortative matching, or abstain. For the Stakeholder/Effort treatment, we show separate vote shares for sincere students and students whose ranks are above or below the sincere students. In the right panel, we show results from hypothesis tests that compare the vote shares between the groups using Fisher's exact tests.

Figure 5: Vote shares in Spectator treatment by who is sincere



Notes: The figure shows the vote share for the respective procedure. "Total" comprises all subjects in the Spectator treatment (n=96), whereas "Who is sincere" divides the sample into subgroups depending on who is sincere (n=32 each).

Figure 6: Post-experimental questionnaire



Notes: The figure shows mean approval to each of the statements in Stakeholder/Effort and Spectator. The scale ranges from 0 (fully disagree) to 4 (fully agree).

when sincere students have ranks 3 and 4, or 4 and 5. In contrast, when students have ranks 2 and 3, the proportion of them voting for the assortative matching is 50%, compared to only 31% in the benchmark case. This suggests that the support for the Boston mechanism interacts with who is hit by strategic choices of others, that is, whether the disadvantaged students are ranked high or low. However, due to the low number of observations in the corresponding cells, we do not have enough statistical power to estimate these treatment differences precisely.

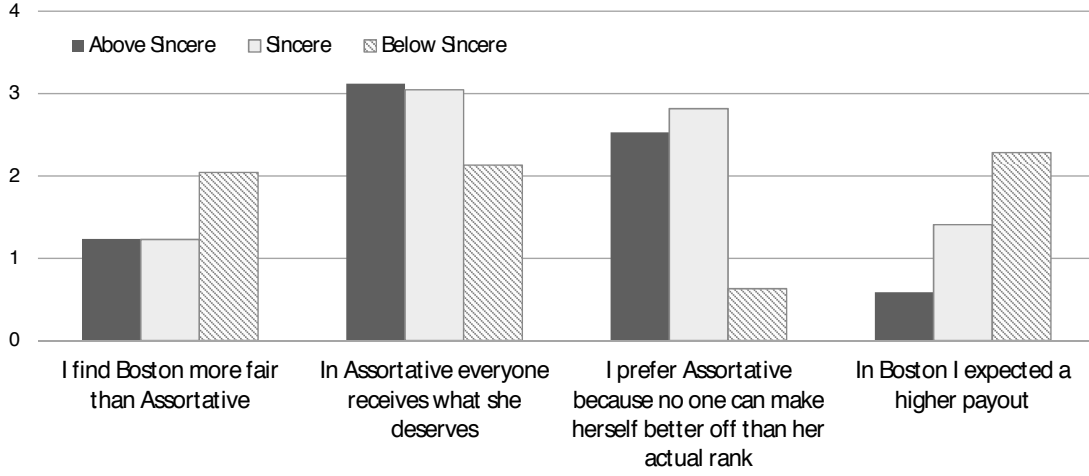
To sum up, in the Spectator treatment where neither self-interest nor overconfidence can play a role by design, a significant fraction of subjects favor the Boston mechanism. This fraction is almost identical in size as the fraction of voters for the Boston mechanism behind the veil of ignorance. Thus, we do not find support for Hypothesis 4.

Result 4: *Almost half of the spectators vote for the Boston mechanism. A slightly smaller proportion votes for the assortative matching, and around 16% abstain.*

3.2.3 Reasons for voting decisions

Unlike in treatment Stakeholder/VoI, overconfidence cannot explain the spectators' preference for the Boston mechanism. In order to shed more light on why subjects vote for the Boston mechanism, we turn to the responses in the post-experimental questionnaire.

Figure 7: Post-experimental questionnaire (only Stakeholder/Effort)



Notes: The figure shows mean approval to each of the statements in the Stakeholder/Effort treatment. The scale ranges from 0 (fully disagree) to 4 (fully agree).

Figure 6 displays the mean approval to statements regarding the mechanisms by treatment and by voting decision. A clear pattern emerges: those who vote for a mechanism agree more strongly to statements regarding their positive attributes than others. More strikingly, this pattern is very similar in both treatments. Despite shutting down self-interest in the Spectator treatment, the heterogeneity in fairness evaluations resembles the heterogeneity among stakeholders. These results suggest that spectators who vote for the Boston mechanism find the Boston mechanism fairer while the opposite is true for spectators who vote for the assortative matching.

This interpretation is also supported by the motivations that spectators report in the free-text part of the post-experimental questionnaire.¹² For example, they state that they voted for the Boston mechanism because “even with a bad result in the preceding test, a smart approach to the second part can get you into a better school,” “it is fairer than letting the allocation only depend on performance in the first part,” or “performance in the first part depends on certain capabilities and some people find it easier than others.”

On the other hand, spectators vote for the assortative matching because “the allocation is based on merit and those who deserve it have a 100% chance to get into the school they deserve,” “because in the other mechanism two random students are disadvantaged,” or “when subjects in lower ranks act strategically, Students 3 and 4 end up in School E, which I do not consider fair.”

¹²All statements were translated from German.

The last two columns of Figure 6 show that most subjects vote for the mechanism that maximizes their payoff. This emerges also from Figure 7 where we distinguish between subjects in the role of students who profit or lose from the Boston mechanism. Subjects who benefit from the Boston mechanism (“below sincere”) are more likely to find the Boston mechanism fairer than the assortative matching, they agree less to the statement that everyone gets what they deserve in the assortative matching, and they do not agree that the assortative outcome avoids that people can get more than they deserve according to their rank. Fairness evaluations in the Stakeholder treatment are self-serving.

3.2.4 Results of Experiment 2: Preference lists and outcomes of Boston mechanism

The Boston mechanism was selected in ten sessions in the Stakeholder/Random treatment and in six sessions in the Spectator treatment. The results are consistent with the findings of Experiment 1 in that the vast majority of subjects understood the mechanism well and exhibited equilibrium behavior. The shares of equilibrium behavior under the Boston mechanism are reported in Table A.2 in the appendix. Overall, in Stakeholder/Random 85.0% of subjects choose the equilibrium strategy and 77.8% of subjects in the Player sessions. The treatment difference is not significant.

Next we consider the share of subjects who expect others to submit a first choice that is consistent with equilibrium behavior. As in Experiment 1, by far the majority of subjects expect other subjects to rank the school first which is prescribed by the equilibrium strategy, see Table A.3 in the appendix.

4 Conclusions

Both studies show that stakeholders vote in line with their self-interest. This is true independently of whether the subjects’ ranks (or priority at schools) are determined randomly or with a real-effort task. This supports the work by [Pathak and Sönmez \(2008\)](#) who assume that strategically sophisticated parents support the Boston mechanism.

In the first study with random ranks, the majority of subjects, who do not know what their rank will be, prefer the Boston mechanism. Since ranks are determined randomly, the Boston mechanism and the assortative matching have the same distribution of payoffs ex ante. Nevertheless, many subjects are willing to give up a fraction of their payoff to vote for the Boston mechanism.

In the second study, the ranks are based on a real-effort task. Spectators, who decide on which mechanism is implemented in a future session, have split preferences between the assortative matching and the Boston mechanism. Thus, the proportion of votes for the

Boston mechanism decreases but it is still sizable. In a post-experimental questionnaire, the spectators state that playing the game well justifies higher payoffs.

The experiment identifies a novel reason for the frequent use of the Boston mechanism. In spite of the presence of sincere students who are harmed by the Boston mechanism, the majority of subjects behind a veil of ignorance and of spectators display a preference for the Boston mechanism over the assortative matching. A significant share of individuals make choices consistent with the idea that an environment in which one has to be fair. Given the often voiced concern that strategically complex mechanisms increase inequality and therefore tend to be unfair, this finding is surprising; it suggests that fairness preferences over matching mechanisms are more complex than typically assumed. Future research needs to assess the importance of procedural preferences over matching mechanisms for real-world applications.

References

- ABDULKADIROĞLU, A., CHE, Y.-K. and YASUDA, Y. (2011). Resolving conflicting preferences in school choice: The “Boston Mechanism” reconsidered. *American Economic Review*, **101** (1), 399–410.
- 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. *The American Economic Review*, **101** (2), 470–492.
- , NOSENZO, D. and RAYMOND, C. (2019). Preferences for truth-telling. *Econometrica*, **87** (4), 1115—1153.
- BASTECK, C. and MANTOVANI, M. (2018). Cognitive ability and games of school choice. *Games and Economic Behavior*, **109** (1), 156–183.
- CAPPELEN, A. W., 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.
- ERGIN, H. and SÖNMEZ, T. (2006). Games of school choice under the boston mechanism. *Journal of Public Economics*, **90** (1-2), 215–237.
- FEATHERSTONE, C. R. and NIEDERLE, M. (2016). Boston versus deferred acceptance in an interim setting: An experimental investigation. *Games and Economic Behavior*, **100**, 353–375.
- GALE, D. and SHAPLEY, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, **69** (1), 9–15.
- HAKIMOV, R. and KÜBLER, D. (2019). Experiments in matching markets: A survey, Working Paper, https://rationality-and-competition.de/wp-content/uploads/discussion_paper/153.pdf.
- KONOW, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, **90** (4), 1072–1091.
- 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.

- SCHILDBERG-HÖRISCH, H. (2010). Is the veil of ignorance only a concept about risk? An experiment. *Journal of Public Economics*, **94**, 1062–1066.
- SCHMELZER, A. (2016). Single versus multiple randomization in matching mechanisms, Discussion Paper Series of the Max Planck Institute for Research on Collective Goods 2016_08.

A Additional Tables and Figures

Table A.1: Descriptive Statistics

	Stakeholder/ Random	Stakeholder/ VoI	Stakeholder/ Effort	Spectator	Player
Age (Mean)	23.708 (3.854)	22.854 (4.946)	23.094 (4.496)	22.802 (3.272)	22.646 (3.836)
Female (Share)	0.521	0.604	0.552	0.500	0.500
Studying (Share)	0.948	0.917	0.927	0.958	0.969
Observations	96	48	96	96	96

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

Table A.2: Equilibrium choices in Stakeholder/Effort and Player (Experiment 2)

Rank	Stakeholder/Effort	Player	Fisher's exact test
Above Sincere	18/20 (90.0%)	10/11 (90.9%)	p=1.000
Below Sincere	33/40 (82.5%)	18/25 (72.0%)	p=0.363
Total	51/60 (85.0%)	28/36 (77.8%)	p=0.415

Notes: Table shows the fraction of subjects who play the equilibrium strategy in the Boston mechanism by treatment. Best response 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.3: Expectations of other subjects' equilibrium choices (Experiment 2)

Rank	Stakeholder/Effort	Player	Fisher's exact test
Above Sincere (Rank 1)	62/70 (88.6%)	37/42 (88.1%)	p=1.000
Above Sincere (Rank 2)	40/49 (81.6%)	19/21 (90.5%)	p=0.485
Above Sincere (Rank 3)	13/21 (61.2%)	12/14 (85.7%)	p=0.252
Below Sincere (Rank 4)	16/21 (76.2%)	18/21 (85.7%)	p=0.697
Below Sincere (Rank 5)	34/49 (69.4%)	21/28 (75.0%)	p=0.794
Below Sincere (Rank 6)	52/70 (74.3%)	28/42 (66.7%)	p=0.397

Notes: This table shows the reported expectations about other subjects' equilibrium behavior regarding their first choice. The equilibrium strategy depends partly on who is sincere. In all cases, students in rank 1 list school A first and in rank 6 list school C first. With students in ranks 2 and 3 being sincere, students in rank 4 list school B first, while students in ranks 5 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 list school C first. With students in ranks 4 and 5 sincere, students in rank 2 list school B first, while students in rank 3 list school C first. Note that students 7 and 8 list school D first, second, or third in equilibrium but we did not elicit beliefs about second or third choices.

B Screenshots

Figure B.1: Screenshot of Trial Period

Sie sind Schüler 3

Übungsrunde: Sie haben 5 Minuten Zeit, das Vergabeverfahren auszuprobieren

In der Tabelle unten repräsentiert jede Zeile die Wunschliste eines Schülers.
Sie können für jeden Schüler die Felder anklicken und eine andere Schule eintragen.

Als Voreinstellung wird für jeden Schüler die Reihenfolge der Schulen gemäß der Auszahlungshöhe angegeben. Die Wunschlisten von Schüler 3 und 4 können nicht verändert werden, da für diese Schüler die Wunschliste automatisch in der Reihenfolge A,B,C,D und E übermittelt wird.

Wenn Sie auf "Ergebnis anzeigen" klicken, erscheint unten auf dem Bildschirm eine weitere Tabelle, die Ihnen angibt, wie bei den von Ihnen erstellten Wunschlisten das Vergabeverfahren die Schüler auf die Schulen verteilt.

	Erstwunsch	Zweitwunsch	Drittwunsch	Viertwunsch	Fünftwunsch
Schüler 1	A	B	C	D	E
Schüler 2	A	B	C	D	E
Schüler 3	A	B	C	D	E
Schüler 4	A	B	C	D	E
Schüler 5	A	B	C	D	E
Schüler 6	A	B	C	D	E
Schüler 7	A	B	C	D	E
Schüler 8	A	B	C	D	E

Ergebnis anzeigen

Notes: The screenshot displays the trial period of student 3, who is in a group where 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.

Figure B.2: Screenshot of Real Effort Task

Verbleibende Sekunden: 99

Sie haben 10 Minuten Zeit, um die Anzahl an Nullen in möglichst vielen Tabellen zu zählen.
Die verbleibende Zeit wird oben rechts angezeigt.

1	1	1	1	0	1	0	0	0	1	1	0
1	0	1	0	1	0	1	0	0	1	0	1
1	1	0	1	0	1	1	1	0	1	0	
0	1	1	0	1	0	1	1	1	1	1	0
1	1	1	1	0	1	1	1	1	0	1	
0	1	1	0	0	1	1	0	0	1	0	
1	0	1	0	0	0	1	1	1	1	1	0
1	0	0	1	0	1	1	1	1	1	1	1
1	1	1	1	0	1	1	0	1	1	1	1
1	1	1	0	0	1	0	1	1	1	1	1

Wie viele Nullen befinden sich in der Tabelle?

Notes: The screenshot displays the real effort task in the Stakeholder/Effort and Spectator treatment. Subjects are supposed to count the number of zeros in the table on the left. 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.