Exp Econ (2013) 16:1–22
DOI 10.1007/s10683-012-9329-5

Preference intensities and risk aversion in school choice: a laboratory experiment

Flip Klijn · Joana Pais · Marc Vorsatz

Received: 19 September 2011 / Accepted: 5 June 2012 / Published online: 21 June 2012

Abstract We experimentally investigate in the laboratory prominent mechanisms that are employed in school choice programs to assign students to public schools and study how individual behavior is influenced by preference intensities and risk aversion. Our main results show that (a) the Gale–Shapley mechanism is more ro-

We are very grateful for comments and suggestions from the Editor, two referees, Eyal Ert, Bettina Klaus, Muriel Niederle, Al Roth, and the seminar audiences in Alicante, Braga, Évora, Maastricht, and Málaga. F. Klijn gratefully acknowledges a research fellowship from Harvard Business School for academic year 2009–2010 when he was visiting HBS and the first draft of the paper was written. He also gratefully acknowledges support from Plan Nacional I+D+i (ECO2008-04784 and ECO2011-29847), Generalitat de Catalunya (SGR2009-01142), and the Consolider-Ingenio 2010 (CSD2006-00016) program. J. Pais gratefully acknowledges financial support from Fundação para a Ciência e a Tecnologia under project reference no. PTDC/EGE-ECO/113403/2009. M. Vorsatz gratefully acknowledges financial support from the Spanish Ministry of Education and Science through the project ECO2009-07530.

Electronic supplementary material The online version of this article (doi:10.1007/s10683-012-9329-5) contains supplementary material, which is available to authorized users.

F. Klijn
Institute for Economic Analysis (CSIC) and Barcelona GSE, Campus UAB, 08193 Bellaterra (Barcelona), Spain
e-mail: flip.klijn@iae.csic.es

J. Pais (✉)
ISEG/Technical University of Lisbon and UECE–Research Unit on Complexity and Economics, Rua Miguel Lupi, 20, 1249-078, Lisboa, Portugal
e-mail: jpais@iseg.utl.pt

M. Vorsatz
Departamento de Análisis Económico II, UNED, Paseo Senda del Rey 11, 28040 Madrid, Spain
e-mail: mvorsatz@cee.uned.es

M. Vorsatz
Fundación de Estudios de Economía Aplicada (FEDEA), Calle Jorge Juan 46, 28001 Madrid, Spain

bust to changes in cardinal preferences than the Boston mechanism independently of whether individuals can submit a complete or only a restricted ranking of the schools and (b) subjects with a higher degree of risk aversion are more likely to play "safer" strategies under the Gale–Shapley but not under the Boston mechanism. Both results have important implications for enrollment planning and the possible protection risk averse agents seek.

Keywords School choice · Risk aversion · Preference intensities · Laboratory experiment · Gale–Shapley mechanism · Boston mechanism · Efficiency · Stability · Constrained choice

JEL Classification C78 · C91 · C92 · D78 · I20

1 Introduction

In school choice programs parents can express their preferences regarding the assignment of their children to public schools. Abdulkadiroğlu and Sönmez (2003) showed that prominent assignment mechanisms in the US lacked efficiency, were manipulable, and/or had other serious shortcomings that often led to lawsuits by unsatisfied parents. To overcome these critical issues, Abdulkadiroğlu and Sönmez (2003) took a mechanism design approach and employed matching theory to propose alternative school choice mechanisms. Their seminal paper triggered a rapidly growing literature that has looked into the design and performance of assignment mechanisms. Simultaneously, several economists were invited to meetings with the school district authorities of New York City and Boston to explore possible ways to redesign the assignment procedures. It was decided to adopt variants of the so-called deferred acceptance mechanism due to Gale and Shapley (1962) (aka the Gale–Shapley mechanism) in New York City and Boston as of 2004 and 2006, respectively.[1] Since many other US school districts still use variants of what was baptized the "Boston" mechanism,[2] it is not unlikely that these first redesign decisions will lead to similar adoptions elsewhere.[3]

Chen and Sönmez (2006) turned to controlled laboratory experiments and showed that the Gale–Shapley mechanism outperforms the Boston mechanism in terms of efficiency if subjects are allowed to rank all schools. Since parents are only allowed

[1] Abdulkadiroğlu *et al.* (2005a, 2005b, 2009) reported in more detail on their assistance and the key issues in the redesign for New York City and Boston, respectively.

[2] That is, the mechanism employed in Boston before it was replaced by the Gale–Shapley mechanism.

[3] The literature has also studied other mechanisms. Abdulkadiroğlu and Sönmez (2003) proposed a mechanism based on Gale's top trading cycles algorithm as a second alternative for the Boston mechanism. However, we are not aware of school districts that employ this other alternative. More importantly, since in Boston and New York the Boston mechanism was *replaced* by Gale–Shapley, our study focuses on the ongoing debate on Gale–Shapley *vs.* Boston. For further recent developments on school choice we refer to Al Roth's blog on market design.

to submit a list containing a limited number of schools in many real-life instances, Calsamiglia *et al.* (2009) experimentally analyzed the impact of imposing such a constraint. They find that manipulation is drastically increased and both efficiency and stability of the final allocations are negatively affected. Another important issue concerns the level of information agents hold on the preferences of the others. Pais and Pintér (2008) focused on this comparing environments where subjects, while aware of their own preferences, have no information at all about the preferences of their peers. A different approach was taken in Featherstone and Niederle (2008), where subjects may not know the preferences of the others, but are aware of their underlying distribution. Both papers studied how strategic behavior is affected by the level of information subjects hold. Featherstone and Niederle (2008) found that truth-telling rates of the two mechanisms are very similar. In Pais and Pintér (2008), truth-telling is higher under Gale–Shapley only when information is substantial, so that the Gale–Shapley mechanism outperforms the Boston mechanism only in some informational settings.

The need of reassessing the school choice mechanisms is reinforced by the recent theoretical findings in Abdulkadiroğlu *et al.* (2011) who showed that the Boston mechanism Pareto dominates the Gale–Shapley mechanism in *ex ante* welfare in certain school choice environments. This happens because the Boston mechanism induces participants to reveal their cardinal preferences (*i.e.*, their relative preference intensities), whereas the Gale–Shapley mechanism does not. In view of this and other results, Abdulkadiroğlu *et al.* (2011) cautioned against a hasty rejection of the Boston mechanism in favor of mechanisms such as the Gale–Shapley mechanism.[4]

Theoretically, whereas the Gale–Shapley mechanism is strategy-proof (that is, agents have incentives to report their ordinal preferences truthfully), a student can increase the likelihood of being assigned a given school by ranking it higher under the Boston mechanism. That is, the Boston mechanism is manipulable and therefore sensitive to underlying cardinal preferences and attitudes towards risk. Motivated by these findings, we experimentally investigate how individual behavior in the Gale–Shapley and Boston mechanisms is influenced by preference intensities and risk aversion and whether this affects the performance of the two mechanisms. We opt for a stylized experimental design that has several important advantages. *First*, by letting subjects participate repeatedly in the same market with varying payoffs, we are able to investigate the impact of preference intensities on individual behavior and welfare. *Second*, a special feature of our laboratory experiment is that before subjects participate in the matching markets, they go through a first phase in which they have to make lottery choices. This allows us to see whether subjects with different degrees of risk aversion behave differently in the matching market. *Third*, the complete information and the simple preference structure form an environment that can be thought through by the subjects, so that clear theoretical predictions about how preference intensities and risk aversion should affect behavior can be made.[5] *Finally*, our setup purposely

[4] Miralles (2008) drew a similar conclusion based on his analytical results and simulations.

[5] On the other hand, since the market we consider in the second phase is small, the results may not scale up to very large real-life matching markets.

does not include coarse school priorities in order to avoid possible problems in entangling the causes of observed behavior.[6]

Our main results are as follows. Subjects tend to list a school higher up (lower down) in the submitted ranking if the payoff of that particular school is increased (decreased) everything else equal. Moreover, the Gale–Shapley mechanism is more robust to changes in cardinal preferences than the Boston mechanism (Result 1). This finding has policy appeal as robustness implies predictability, a valuable asset in enrollment planning. We also find that subjects with a higher degree of risk aversion are more likely to play protective strategies[7] under the Gale–Shapley but not under the Boston mechanism (Result 2). Ease in recognizing protective strategies may make risk averse agents more comfortable with the Gale–Shapley mechanism.

The remainder of the paper is organized as follows. The experimental design is explained in Sect. 2. In Sect. 3, we derive hypotheses regarding the effect of relative preference intensities and risk aversion on strategic behavior. In Sect. 4, we analyze the impact of changes in cardinal preferences, how risk aversion affects behavior in the matching market, and the implications of two variables for the welfare properties of the mechanisms. In Sect. 5, we conclude with some possible policy implications.

2 Experimental design and procedures

Our experimental study comprises four different treatments. Each treatment is divided into two phases.

In the first phase, which is identical for all treatments, we elicit the subjects' degree of risk aversion using the paired lottery choice design introduced by Holt and Laury (2002). Subjects are presented with a list of ten different choices between two lotteries (see Table 10 in Appendix A). Lottery A is less risky than lottery B for the first nine choices, but lottery B first-order stochastically dominates lottery A for the tenth choice. A *rational* individual may choose A at the top of the list, but always chooses B at the bottom, implying some switching point in between. The *switching point*, corresponding to the first time lottery B is chosen, roughly determines the number of safe choices and, in turn, provides a measure of the degree of risk aversion.[8]

In the second phase, subjects face the following stylized school choice problem: There are three teachers, denoted by the natural numbers 1, 2, and 3, and three

[6] Coarse school priorities are a common feature of many school choice environments. Then, in order to apply the assignment mechanisms, random tie-breaking rules are often used. However, the incorporation of such rules in our design would make it very hard to see whether individuals with different degrees of risk aversion behave differently because of strategic uncertainty or because of the random tie-breaking. In other words, we assume that the schools' priority orders are strict in order to study whether the behavioral effect of risk aversion is associated with strategic uncertainty. For the very same reason, we also assume that the induced game is common knowledge even though in practice individuals are likely to have incomplete information regarding the other participants' preferences.

[7] Loosely speaking, a subject plays a protective strategy if she protects herself from the worst eventuality to the extent possible. Consequently, a protective strategy is a maximin strategy.

[8] A rational individual may always choose lottery B, in which case the switching point is equal to 1.

Table 1 Preferences of teachers over schools (*left*) and priority orderings of schools over teachers (*right*)

	Preferences			Priorities		
	Teacher 1	Teacher 2	Teacher 3	School X	School Y	School Z
Best match	X	Y	Z	2	3	1
Second best match	Y	Z	X	3	1	2
Worst match	Z	X	Y	1	2	3

schools, denoted by the capital letters X, Y, and Z, with one open teaching position each.[9] The preferences of the teachers over schools and the priority orderings of schools over teachers, both commonly known to all participants, are presented in Table 1.

It can be seen in Table 1 that the preferences of the teachers form a Condorcet cycle. The priority orderings of the schools form another Condorcet cycle in such a way that every teacher is ranked last in her most preferred school, second in her second most preferred school, and first in her least preferred school. The setup is competitive, so that risk aversion may have a bite, and symmetric to simplify the data analysis.

A 2×2 between-subjects design is obtained from two treatment variables that are known to be empirically relevant in this type of market. The first treatment variable refers to the restrictions on the rankings teachers can submit. We consider the *unconstrained* and one *constrained* setting. In the unconstrained setting (u), teachers have to report a ranking over all three schools. In the constrained setting (c), they have to report the two schools they want to list first and second. The second treatment variable refers to how reported rankings are used by the central clearinghouse to assign teachers to schools. We apply here both Gale–Shapley's teacher-proposing deferred acceptance algorithm (*GS*) and the Boston algorithm (*BOS*). For the particular school choice problem at hand, they are as follows:

Step 1. Each teacher sends an application to the school she listed first.
Step 2. Each school retains the applicant with the highest priority and rejects all other applicants.
Step 3. If a teacher is rejected at a school, she applies to the next highest listed school.
Step 4. (The two algorithms only differ in this step.)

GS: Whenever a school receives new applications, these applications are considered together with the previously retained application (if any). Among the retained and the new applicants, the teacher with the highest priority is retained and all other applicants are rejected.

BOS: Whenever a school receives new applications, all of them are rejected in case the school already retained an application before. If the school did not retain

[9] We "framed" the school choice problem from the point of view of teachers who are looking for jobs because this presentation provides a natural environment that is easy to understand. For example, material payoffs can be directly interpreted as salaries (see Pais and Pintér 2008).

Table 2 Experimental design

Treatment	# of subjects	First phase	Second phase				
			Pref. revelation	Algorithm	Game		
					First	Second	Third
GS_u	54	Holt & Laury	unconstrained	Gale–Shapley	GS_{u20}	GS_{u13}	GS_{u27}
GS_c	54	Holt & Laury	constrained	Gale–Shapley	GS_{c20}	GS_{c13}	GS_{c27}
BOS_u	55	Holt & Laury	unconstrained	Boston	BOS_{u20}	BOS_{u13}	BOS_{u27}
BOS_c	55	Holt & Laury	constrained	Boston	BOS_{c20}	BOS_{c13}	BOS_{c27}

an application so far, it retains among all applicants the one with the highest priority and rejects all other applicants.

Step 5. The procedure described in Steps 3 and 4 is repeated until no more applications can be rejected. Each teacher is finally assigned to the school that retains her application at the end of the process. In case none of a teacher's applications are retained at the end of the process, which can only happen in the constrained mechanisms, she remains unemployed and gets 0 ECU.[10]

Each subject faces one of the four treatments and plays the role of a teacher (schools are not strategic players). The task is to submit a ranking over schools (not necessarily the true preferences) to be used by a central clearinghouse to assign teachers to schools. This is done three times, in three games with payoff structures that differ only in the payoff of the second most preferred school: A subject always receives 30 ECU for her most and 10 ECU for her least preferred schools, but in the first, second, and third games, a subject receives 20 ECU, 13 ECU, and 27 ECU, respectively, if she obtains a job at her second most preferred school.[11] To maintain the notation as simple as possible, we sometimes use 27 ECU to refer to the payoff structure in which the second preferred school is worth 27 ECU. Moreover, GS_{c27} will refer to the game induced by the constrained Gale–Shapley mechanism with the payoff structure 27 ECU. All other situations are indicated accordingly. Table 2 summarizes the experimental design.

The experiment was programmed within the z-Tree toolbox provided by Fischbacher (2007) and carried out in the computer laboratory at a local university. We used the ORSEE registration system by Greiner (2004) to invite students from a wide

[10] If teachers had to list only one school, the two constrained mechanisms would be identical; that is, for all profiles of submitted (degenerate) rankings, the same matching would be obtained under the Gale–Shapley and Boston algorithms.

[11] Since the payoff of the second most preferred school varies for all subjects, subjects face different kinds of opponents in different games. In one alternative design to possibly overcome this drawback the payoff for only one subject (in each group of three subjects) varies. Yet, in this alternative approach, the subjects with fixed preferences would probably believe that the third subject modifies her strategy due to the change in the preference intensities to which they respond by adapting their behavior as well, *etc.* The elicitation of beliefs would certainly provide important information regarding the individual motives but would, at the same time, further complicate the design. Also, if we only changed the preferences of one subject the data to be collected would triple (to a total of 654 subjects).

range of faculties. In total, 218 undergraduates participated in the experiment. We almost obtained a perfectly balanced distribution of participants across treatments even though some students did not show up.[12]

Each session proceeded as follows. At the beginning, each subject only received instructions for the first phase (which included some control questions) together with an official payment receipt. Subjects could study the instructions at their own pace and any doubts were privately clarified. Participants were informed that they would play afterwards a second phase, without providing any information about its structure. Subjects also knew that their decisions in phase 1 would not affect their payoffs in the other phase (to avoid possible hedging across phases) and that they would not receive any information regarding the decisions of any other player until the end of the session (so that they could not condition their actions in the second phase on the behavior of other participants in the first phase). In theory, therefore, the two phases are independent from each other.

After completing the first phase, subjects were anonymously matched into groups of three (within each group, one subject became teacher 1, one subject teacher 2, and one subject teacher 3) and entered the second phase of the experiment, where they faced one of the four treatments. The roles within the groups remained the same throughout the second phase. Subjects were informed that three school choice games would be played sequentially within the same group, but they never knew how the parameters would change. It was also made clear that no information regarding the co-players' decisions, the induced matching, or the resulting payoffs would be revealed at any point in time. No feedback whatsoever was provided. Apart from (very likely) avoiding issues with learning, this prevented subjects from conditioning their decisions on former actions of other group members.[13]

To prevent income effects, either phase 1 or 2 was payoff relevant (one participant determined the payoff relevant phase by throwing a fair coin at the end of the experiment), which was known by the subjects from the beginning. If the first phase was payoff relevant, the computer selected randomly one of the ten decision situations and the uncertainty in the lottery chosen by the subject then resolved in order to determine the final payoff. If the second phase was payoff relevant, the computer randomly selected one of the games. Subjects were then paid according to the matching induced by the submitted rankings. At the end of the experiment, subjects were informed about the payoff relevant situation and their final payoff. Subjects received 4 Euro (40 Eurocents) per ECU in case the first (second) phase was payoff relevant. These numbers were chosen to induce similar expected payoffs. A typical

[12] In each treatment using the Boston algorithm, we had one student left that could not be matched with other participants. These two students took decisions without knowing that they remained unmatched. Finally, we paid them as if they were assigned a place at their most preferred school.

[13] It is well known (Dubins and Freedman 1981 and Roth 1982) that teachers have incentives to report their ordinal preferences truthfully in treatment GS_u, in which case the induced matching would be stable and efficient with respect to the teachers' true preferences. However, to put all treatments at the same level, these incentives were neither directly revealed in the instructions nor were they indirectly taught by going over several examples. Otherwise, a convincing argument in favor of truth-telling in GS_u would render the comparison between GS_u and the other mechanisms rather obvious. Also, explicit advice would only increase the (observed) efficiency and stability gap between GS_u and BOS_u, *i.e.*, strengthen our results.

Table 3 A given strategy is undominated if and only if the corresponding entry is ×

Treatment	Rankings				
	(1, 2, 3)	(1, 3, 2)	(2, 1, 3)	(2, 3, 1)	(3, ×, ×)
Gale–Shapley unconstrained	×				
Gale–Shapley constrained	×	×		×	
Boston unconstrained	×		×		
Boston constrained	×	×	×	×	×

session lasted about 75 minutes and subjects earned on average 12.21 Euro (including a 3 Euro show-up fee) for their participation. The instructions, which are translated from Spanish, can be found in Supplementary Material.

3 Experimental hypotheses

Since the school choice problem is set up symmetrically, the three teachers face exactly the same decision problem and we can simplify the description of the strategy spaces. For instance, in the unconstrained (constrained) setting we make use of the notation $(2, 3, 1)$ for the ranking where a teacher lists her second most preferred school first, her least preferred school second, and her most preferred school last (does not rank her most preferred school). The other five strategies $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(3, 1, 2)$, and $(3, 2, 1)$ have similar interpretations. Also, note that for all four mechanisms the strategies $(3, 1, 2)$ and $(3, 2, 1)$ are strategically equivalent; that is, they always yield a payoff of 10 ECU for sure, independently of the other players' strategies. Although possibly not all subjects were aware of the strategic equivalence of $(3, 1, 2)$ and $(3, 2, 1)$, we nevertheless decided to pool these two strategies in our analysis through the notation $(3, \times, \times)$.

3.1 Preference intensities

The first step in the derivation of our experimental hypotheses is the assumption that rational subjects do not play dominated strategies. Table 3 shows the undominated strategies for each of the four treatments.

The second step is to derive predictions about how variations in the cardinal preference structure affect individual behavior in the matching markets:

Prediction 1 *Subjects no longer list school* 2 *or list school* 2 *further down in their submitted ranking if the payoff of this school decreases from* 20 *ECU to* 13 *ECU. Similarly, subjects no longer exclude school* 2 *from their submitted ranking or list school* 2 *further up in their ranking if the payoff of this school increases to* 27 *ECU.*

The economic intuition behind this prediction is fairly simple. Whenever the payoff of a school decreases everything else equal, its relative attractiveness decreases. Consequently, subjects who originally rank school 2 above some other school(s) may

Table 4 Hypotheses about how preference intensities affect the play of undominated strategies

Treatment	Rankings				
	$(1, 2, 3)$	$(1, 3, 2)$	$(2, 1, 3)$	$(2, 3, 1)$	$(3, \times, \times)$
Gale–Shapley unconstrained					
Change from 20 to 13 ECU	$=$				
Change from 20 to 27 ECU	$=$				
Gale–Shapley constrained					
Change from 20 to 13 ECU	$-$	$+$		$-$	
Change from 20 to 27 ECU	$+$	$-$		$+$	
Boston unconstrained					
Change from 20 to 13 ECU	$+$		$-$		
Change from 20 to 27 ECU	$-$		$+$		
Boston constrained					
Change from 20 to 13 ECU	?	$+$	$-$	$-$	$+$
Change from 20 to 27 ECU	?	$-$	$+$	$+$	$-$

decide to push it further down their ranking or not list it at all. A symmetric argument applies if the payoff of school 2 is increased. Combining Table 3 and Prediction 1 we obtain Hypothesis 1, on how the use of undominated strategies changes due to variations in cardinal preferences.

Hypothesis 1 *Preference intensities affect the play of undominated strategies as described in Table 4.*

We explain Hypothesis 1 for the case in which the payoff of the second school is reduced from 20 ECU to 13 ECU (the argument regarding an increase to 27 ECU is similar). Consider first the Gale–Shapley mechanisms. There should not be any effect in treatment GS_u, simply because truth-telling is the only undominated strategy for this mechanism. In treatment GS_c, only the strategies $(1, 2, 3)$, $(1, 3, 2)$, and $(2, 3, 1)$ are undominated. Subjects who initially played $(1, 3, 2)$ will also do so after the reduction of the payoff of school 2. Also, subjects who initially told the truth may change to play $(1, 3, 2)$ instead. Finally, subjects who initially played $(2, 3, 1)$ could be tempted to play $(3, 2, 1)$ or $(3, 1, 2)$, as suggested by our prediction. However, these strategies are dominated by $(2, 3, 1)$ and $(1, 3, 2)$, respectively. Hence, if a subject who initially played $(2, 3, 1)$ changes her strategy, then we expect her to play $(1, 3, 2)$. So, when the second school pays 13 ECU the strategies $(1, 2, 3)$ and $(2, 3, 1)$ will be played less often and $(1, 3, 2)$ more often compared to the situation where the second schools pays 20 ECU.

Now consider the Boston mechanisms. According to Table 3, only the strategies $(1, 2, 3)$ and $(2, 1, 3)$ are undominated in BOS_u. Clearly, every individual who told the truth under the original payoffs will still prefer to tell the truth when the payoff of school 2 is reduced. On the other hand, subjects who initially played the strategy $(2, 1, 3)$ may switch to telling the truth. Hence, our hypothesis states that the

Table 5 A given strategy is protective if and only if the corresponding entry is ×

Treatment	Rankings				
	(1, 2, 3)	(1, 3, 2)	(2, 1, 3)	(2, 3, 1)	(3, ×, ×)
Gale–Shapley unconstrained	×				
Gale–Shapley constrained		×		×	
Boston unconstrained	×		×		
Boston constrained					×

change in the payoffs makes subjects report more often the ranking $(1, 2, 3)$ and less often the ranking $(2, 1, 3)$. Finally, we consider BOS_c. Here, every strategy is undominated. Similarly to GS_c, subjects who initially played $(1, 3, 2)$ will also do so after the reduction of the payoff, and subjects who initially told the truth may change to play $(1, 3, 2)$ instead. Individuals who submitted the ranking $(3, \times, \times)$ opted for the school that guarantees access and hence a payoff reduction of school 2 should not affect their choice. However, subjects who initially chose $(2, 3, 1)$ may now submit the riskless strategy $(3, \times, \times)$ so that this strategy could be played more often after the reduction of the payoff. Finally, subjects who initially played $(2, 1, 3)$ could possibly change to $(1, 2, 3)$ or $(1, 3, 2)$. All in all, strategies $(1, 3, 2)$ and $(3, \times, \times)$ will be played more often, and strategies $(2, 1, 3)$ and $(2, 3, 1)$ will be played less often. Since there are two opposite effects regarding strategy $(1, 2, 3)$, we do not make a prediction regarding the change in truth-telling.

3.2 Risk aversion

In the second phase of the experiment, subjects face strategic uncertainty and thus form subjective beliefs about the other group members' strategies. So, for instance they have to ponder the economic benefits from working at their top school against the probability that another subject with a higher priority for that school applies and grabs the slot.

To investigate whether subjects with different attitudes towards risk as obtained in the first phase of the experiment act differently in the second phase, we use the concept of protective strategies introduced by Barberà and Dutta (1982). Loosely speaking, when an agent has no information about the others' submitted preferences, she behaves in a protective way if she plays a strategy so as to protect herself from the worst eventuality to the extent possible.[14] We discuss the formal definition of protective strategies in Appendix B, where we also prove that protective strategies in the second phase are those reported in Table 5.

We can now formally state our prediction regarding the use of protective strategies.

Hypothesis 2 *Subjects who are more risk averse are more likely to play a protective strategy in the matching market.*

[14]Two settings in which protective strategies have been studied are two-sided matching markets (Barberà and Dutta 1995) and, more recently, paired kidney exchange (Nicolò and Rodríguez-Álvarez 2012).

Table 6 Each row gives the probability distribution of submitted rankings in the corresponding game. For each row, the most salient strategies (undominated strategies) are indicated in *boldface* (*underlined*)

Game	Submitted rankings				
	(1, 2, 3)	(1, 3, 2)	(2, 1, 3)	(2, 3, 1)	(3, ×, ×)
Gale–Shapley unconstrained					
20 ECU	<u>**0.50**</u>	0.00	0.41	0.03	0.06
13 ECU	<u>**0.65**</u>	0.04	0.19	0.02	0.10
27 ECU	<u>**0.44**</u>	0.00	0.43	0.07	0.06
Gale–Shapley constrained					
20 ECU	<u>0.24</u>	<u>0.19</u>	0.15	<u>**0.31**</u>	0.11
13 ECU	<u>0.17</u>	<u>**0.31**</u>	0.09	<u>0.28</u>	0.15
27 ECU	<u>0.21</u>	<u>0.13</u>	0.26	<u>**0.31**</u>	0.09
Boston unconstrained					
20 ECU	<u>**0.40**</u>	0.02	<u>**0.40**</u>	0.16	0.02
13 ECU	<u>**0.62**</u>	0.04	<u>0.14</u>	0.07	0.13
27 ECU	<u>0.31</u>	0.00	<u>**0.55**</u>	0.09	0.05
Boston constrained					
20 ECU	<u>**0.27**</u>	<u>0.20</u>	<u>0.15</u>	<u>0.25</u>	<u>0.13</u>
13 ECU	<u>0.18</u>	<u>**0.37**</u>	<u>0.13</u>	<u>0.16</u>	<u>0.16</u>
27 ECU	<u>0.14</u>	<u>0.06</u>	<u>0.27</u>	<u>**0.44**</u>	<u>0.09</u>

4 Results

4.1 Preference intensities

First, we present aggregate data and analyze how the empirical distribution of submitted rankings changes according to the applied cardinal preferences.

It can be seen from Table 6 that the most salient ranking is always an undominated strategy. It follows from inspection of column (1, 2, 3) that for each payoff constellation and among all four mechanisms, the level of truth-telling is highest in GS_u. This is not a surprise because it is the only mechanism for which truth-telling is the unique undominated strategy (Table 3). Still, it falls well short of 100 % in this treatment, as several subjects did not recognize that it is in their best interest to reveal preferences honestly.[15,16]

[15] In Chen and Sönmez (2006), in their "random" and "designed" treatments of GS_u, 56 % and 72 % of the subjects, respectively, submitted their true preferences. The numbers are 58 % and 57 % in Calsamiglia *et al.* (2009). Our numbers seem to be slightly lower but a real comparison is not possible due to the very different environments.

[16] Using χ^2 tests for homogeneity one verifies that for all cardinal payoff constellations, (a) the distribution of submitted rankings in treatment GS_u (BOS_u) is significantly different from the one in treatment GS_c (BOS_c) and (b) the distributions of submitted rankings in treatments GS_u and BOS_u (GS_c and BOS_c) are not significantly different from each other. The second finding might create the impression that subjects perceive the Gale–Shapley matching algorithm in the same way as the Boston algorithm. Results 1 and 2 presented below, however, will reveal that this is not the case.

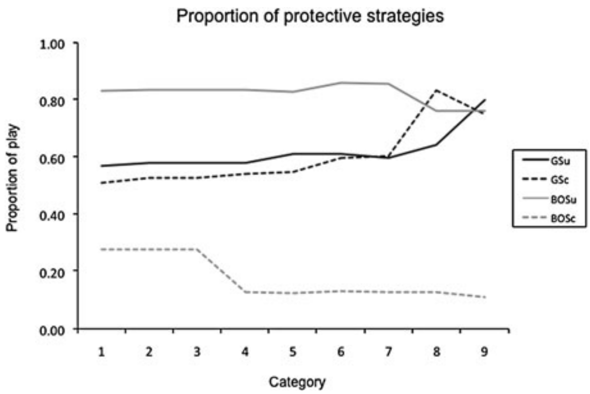
Table 7 Changes in the probability distributions of submitted rankings. A *negative* (*positive*) number indicates that the corresponding ranking is used more (less) often with the payoff structure 20 ECU. We also present the one-sided p-value of the χ^2 test for homogeneity that analyzes whether the empirical distribution depends on the relative preference intensities. A *boldfaced* number indicates that the use of the corresponding ranking changes (one-sided Wilcoxon signed-rank test at the 5 % significance level). Undominated strategies are *underlined*

Treatment	Rankings					p-value
	$(1,2,3)$	$(1,3,2)$	$(2,1,3)$	$(2,3,1)$	$(3,\times,\times)$	
Gale–Shapley unconstrained						
Change from 20 to 13 ECU	$\underline{\mathbf{0.15}}$	0.04	**−0.22**	−0.02	0.06	*0.0300*
Change from 20 to 27 ECU	$\underline{-0.06}$	0.00	0.02	0.04	0.00	*0.4650*
Gale–Shapley constrained						
Change from 20 to 13 ECU	$\underline{-0.07}$	$\underline{\mathbf{0.13}}$	−0.06	$\underline{-0.04}$	0.04	*0.2300*
Change from 20 to 27 ECU	$\underline{-0.04}$	$\underline{-0.06}$	**0.11**	$\underline{0.00}$	−0.02	*0.3300*
Boston unconstrained						
Change from 20 to 13 ECU	$\underline{\mathbf{0.22}}$	0.02	$\underline{\mathbf{-0.25}}$	**−0.09**	**0.11**	*0.0002*
Change from 20 to 27 ECU	$\underline{\mathbf{-0.09}}$	−0.02	$\underline{\mathbf{0.15}}$	−0.07	0.04	*0.1400*
Boston constrained						
Change from 20 to 13 ECU	$\underline{-0.09}$	$\underline{\mathbf{0.16}}$	$\underline{-0.02}$	$\underline{-0.09}$	$\underline{0.04}$	*0.1450*
Change from 20 to 27 ECU	$\underline{\mathbf{-0.13}}$	$\underline{\mathbf{-0.15}}$	$\underline{\mathbf{0.13}}$	$\underline{\mathbf{0.18}}$	$\underline{-0.04}$	*0.0100*

Next, we study the impact of cardinal preferences on individual behavior. The relevant data is provided in Table 7, which shows the differences in the probability distribution of submitted rankings when the payoff of the second best school is decreased (increased) from 20 ECU to 13 ECU (27 ECU). For the sake of completeness, we also present the one-sided p-values of the χ^2 tests for homogeneity that analyze whether the respective distributions differ.

We see that a reduction of the payoff of school 2 from 20 to 13 ECU changes the distribution of submitted rankings in the unconstrained but not in the constrained setting, while raising its payoff from 20 to 27 ECU only affects the distributions in BOS_c. To analyze these findings in more detail, we run Wilcoxon signed-rank tests as they allow us to see whether the use of a particular ranking changes. The boldfaced numbers in Table 7 indicate which rankings are used significantly more often or less often. Since Hypothesis 1 only deals with undominated strategies, we simply have to check whether the sign of each boldfaced number that is underlined in Table 7 coincides with the corresponding sign in Table 4. One finds that almost all significant changes related to undominated strategies are in line with the hypothesis, the only exception is that the strategy $(1,2,3)$ is used significantly more often in GS_u when the payoff of the second best school is reduced from 20 to 13 ECU (Hypothesis 1 suggested no change). Moreover, for both the constrained and the unconstrained settings, all significant changes that take place under the Gale–Shapley mechanism also occur under the corresponding Boston mechanism. Consequently, we can summarize our findings as follows.

Fig. 1 Proportion of protective strategies played in each of the four treatments (averages over the three games) as the subjects with lowest degree of risk aversion (*i.e.*, lowest switching point) are eliminated step-by-step from the subject pool



Result 1 (Cardinal preferences) *Hypothesis* 1 *cannot be rejected. Moreover, for both the constrained and the unconstrained settings, the Gale–Shapley mechanism is more robust to changes in cardinal preferences than the corresponding Boston mechanism.*

4.2 Risk aversion

To test Hypothesis 2, we look at the proportion of protective strategies played in each treatment as we eliminate step-by-step the subjects with the lowest degree of risk aversion from the subject pool.

The data obtained from this process is presented in Fig. 1. The horizontal axis indicates which subjects are being considered; on the vertical axis, we plot the proportion with which the considered subjects play a protective strategy. In category 1, the subject pool consists of all subjects with switching point 1 or higher, *i.e.*, the whole pool of rational subjects;[17] in category 2, the reduced subject pool consists of all rational subjects with switching point 2 or higher; and so forth. Consequently, as we move from the left to the right in the graph, the subjects with the lowest risk aversion among all those still considered are being discarded. This procedure has the potential drawback that the distributions of rankings for high switching points are likely to be determined by only a few subjects. It turns out that this is true only in the last step of elimination, when we solely consider subjects with switching point 10 (a total of five in our subject pool), which is why we decided not to include this as a separate category in Fig. 1. The numbers for all the other switching points are based on a considerable amount of data. For instance, in each treatment, approximately half the subjects choose lottery B for the first time in the seventh decision situation or later (GS_u: 25 out of 48 subjects; GS_c: 21 out of 48 subjects; BOS_u: 30 out of 50 subjects; and BOS_c: 25 out of 47 subjects).

Intuitively, the figure should be looked at in the following way: If a curve is flat, then the use of protective strategies does not depend on the degree of risk aversion. On the other hand, if a curve is increasing (decreasing), protective strategies are more (less) likely to be used by the subjects with a higher degree of risk aversion.

[17] We only considered data from subjects who behaved rationally in the first phase of the experiment, omitting those that switch from lottery B to lottery A.

Table 8 Tobit ML estimation results on how risk aversion affects behavior in each treatment. Standard errors are *in parentheses*. Errors are robust to heteroscedasticity

Variable	Treatment			
	GS_u	GS_c	BOS_u	BOS_c
Constant (β_0)	−0.2581	−0.3367	1.0957*	−0.1658
	(0.4349)	(0.3461)	(0.6409)	(0.8481)
Switching point (β_1)	0.1376**	0.1454**	0.0620	−0.0646
	(0.0686)	(0.0569)	(0.0960)	(0.1291)

* Significant at the 10-percent level (two-sided). ** Significant at the 5-percent level (two-sided). OLS and Probit ML (with standard errors clustered at the individual level) estimations yield similar results

The figure suggests only for the Gale–Shapley mechanisms a positive dependence between risk aversion and the use of protective strategies. To formally test this, we estimate the parameters of the following linear model. Let $p_i(t)$ be the pooled probability (over all three payoff constellations) that individual i who participates in treatment t plays a protective strategy. Similarly, $s_i(t)$ is the switching point of individual i in treatment t extracted in the first phase of the experiment. We then have that

$$p_i(t) = \beta_0 + \beta_1 s_i(t) + \varepsilon_i(t),$$

where $\varepsilon_i(t)$ is the error of individual i in treatment t. We assume that the errors are i.i.d. across individuals in a given treatment. The parameter estimates of the Tobit Maximum Likelihood estimation procedure are presented in Table 8.

Table 8 fully confirms the intuition from Fig. 1. In the two treatments using the Gale–Shapley algorithm, protective strategies are played more often the more risk averse subjects are. With respect to the two treatments using the Boston algorithm, we find that risk aversion is uncorrelated with the use of the protective strategies.

Result 2 (Risk aversion) *Subjects who are more risk averse are more likely to play a protective strategy under the Gale–Shapley mechanisms but not under the Boston mechanisms.*

4.3 Performance

In this section, we study how preference intensities and risk aversion affect the performance of the mechanisms in terms of efficiency and stability. Efficiency for teachers is the primary welfare goal (school slots are mere objects and are hence not taken into account). Formally, *efficiency* is defined as the expected payoff per teacher. To obtain this number, we first calculate all possible preference profiles. Next, we determine for each profile the induced average payoff. Finally, we calculate the weighted average of the induced average payoffs, where the weight for each profile is obtained from the empirical distribution presented in Table 6.

We first focus on the data for the whole population. The left-hand side of Table 9 shows that expected payoffs under the Boston mechanisms are not always lower than those under the Gale–Shapley mechanisms. In fact, whereas Gale–Shapley has the

Table 9 Efficiency (*to the left*) and probability of stable matchings (*to the right*) for the whole population (*in boldface on top*), the high risk aversion subjects (*in the middle*), and the low risk aversion subjects (*at the bottom*) in every game

Treatment	Efficiency			Stability		
	20 ECU	13 ECU	27 ECU	20 ECU	13 ECU	27 ECU
Gale–Shapley unconstrained	**21.06**	**19.53**	**26.06**	**0.85**	**0.71**	**0.86**
	21.24	21.42	27.33	0.88	1.00	1.00
	21.49	18.16	25.83	0.76	0.55	0.77
Gale–Shapley constrained	**17.31**	**14.77**	**21.53**	**0.54**	**0.48**	**0.58**
	16.91	15.41	21.58	0.57	0.44	0.66
	17.51	14.98	23.82	0.51	0.46	0.67
Boston unconstrained	**20.63**	**20.09**	**25.36**	**0.65**	**0.43**	**0.67**
	20.29	21.32	25.68	0.68	0.45	0.70
	21.05	19.85	24.55	0.61	0.42	0.54
Boston constrained	**17.99**	**16.22**	**22.89**	**0.33**	**0.30**	**0.60**
	17.80	14.62	24.60	0.42	0.28	0.76
	18.45	17.05	22.23	0.29	0.29	0.53

tendency to create a higher welfare than Boston in the unconstrained case, it turns out that the efficiency is always higher in BOS_c than in GS_c. Using all possible recombinations of submitted rankings, we find with the help of t-tests for equal means that all differences across mechanisms are significant at $p = 0.0001$.

Two elements contribute to the observed differences across mechanisms. First, the mechanisms produce different outcomes for some strategy profiles. This can be accounted for by looking at the efficiency levels when the same distribution of strategy profiles is applied to the Gale–Shapley and Boston mechanisms. Second, even though neither GS_u and BOS_u nor GS_c and BOS_c induce significantly different distributions of submitted rankings (Footnote 16), differences in individual behavior across mechanisms have an impact on efficiency. For instance, when comparing BOS_{u20} and GS_{u20}, the former yields a higher average payoff than the latter independently of the exact (common) distribution of strategy profiles;[18] this strongly suggests that the observed efficiency differences between these two treatments relies exclusively on those small differences in behavior and, in fact, an inspection of Table 6 reveals that the proportion of truth-telling under GS_{u20} is higher than under BOS_{u20}.

To see whether the differences in efficiency are related to the subjects' risk aversion, we divide the subject pool for each treatment into two subgroups. The first group, which we label as the "high risk aversion" subjects, consists of the individuals

[18] The slightly cumbersome calculations are available from the authors upon request.

who selected lottery B for the first time in the seventh decision situation or later. The remaining individuals are labeled "low risk aversion" subjects.[19]

The second and third row of each treatment on the left-hand side of Table 9 present the efficiency for the high and low risk aversion groups, respectively. Note that these numbers are obtained by taking recombinations at the subgroup level. We find for the subjects with a high risk aversion that efficiency is higher in GS_u than in BOS_u. This result is not surprising if one takes into account that for this subgroup, the proportion of truthfully submitted rankings (aggregated over all three games) is 0.60 in treatment GS_u but "only" 0.43 in treatment BOS_u. On the other hand, efficiency for the subjects with a low risk aversion is higher in treatment BOS_u than in treatment GS_u if the payoff of the second school is 13 ECU. Regarding the constrained treatments, we observe that GS_c outperforms BOS_c for the subjects with high (low) risk aversion only if the payoff of the second school is 13 ECU (27 ECU). For both subgroups, all differences across mechanisms are significant at $p = 0.05$.

Result 3 (Efficiency) *For both the low and high risk aversion groups as well as the whole experimental population*, (i) GS_u *tends to outperform* BOS_u *and* (ii) BOS_c *tends to outperform* GS_c.

Finally, we report on stability. Stability of the matchings reached should be met for the assignment procedure to be "successful" (it avoids lawsuits or the appearance of matches that circumvent the mechanism). A matching is blocked if there is a teacher that prefers to be assigned to some school with a slot that is either available or occupied by a teacher with a lower priority. A matching is *stable* if it is not blocked. In our setup, there are three stable matchings labeled teacher optimal, compromise, and school optimal. Under each of these symmetric matchings, every teacher is assigned to its most preferred, second most preferred, and least preferred school, respectively.

Again, we first concentrate on the whole population. The numbers on the right-hand side of Table 9 are the proportions of stable matchings reached for each treatment given all possible recombinations of submitted rankings and taking into account the empirical distribution presented in Table 6. We can see that Gale–Shapley is in general more successful than Boston in producing stable matchings; the only exception regards the constrained mechanisms when the payoff of the second school is 27 ECU (all differences are significant at $p = 0.0001$). This is in line with the findings in Calsamiglia *et al.* (2009).[20] More importantly, when the magnitude of the changes in the proportion of stable matchings is taken into account, it appears to be the case that, very much in resemblance to Result 1, the Gale–Shapley mechanisms are less sensitive to changes in the payoff of school 2 than the Boston mechanisms. In fact, when comparing the percentage of stable matchings reached when school 2 is worth

[19] The common switching point has not been chosen arbitrarily. According to our data, the average switching point is 6.47 in GS_u, 5.98 in GS_c, 6.70 in BOS_u, and 6.55 in treatment BOS_c so that the difference in the group sizes is minimal if the seventh decision situation is taken as the dividing line.

[20] In theory, the two unconstrained mechanisms should yield stable matchings if subjects recognize that telling the truth is weakly dominant (in the case of GS_u) and do not fail to play Nash equilibria (in the case of BOS_u, see Ergin and Sönmez 2006).

13 and 27 ECU, differences in stability reach 0.14 and 0.10 under GS_u and GS_c, respectively, against 0.23 and 0.29 under BOS_u and BOS_c (all differences are significant at $p = 0.0001$). Interestingly, the advantage of Gale–Shapley over Boston is obtained because Gale–Shapley tends to produce far more compromise stable matchings.

A final comment on how stability is affected by the degree of risk aversion. The relevant numbers are again presented in the second and third rows belonging to each mechanism in Table 9. All differences are significant at $p = 0.0001$. In general, the differences in the percentage of stable matchings obtained within each group of subjects follow roughly the same rules as those obtained when the full sample is considered. One point is worth noticing, though: The levels of stability are typically higher among the highly risk averse subjects, reaching even 100 % under GS_u when the second school is worth 13 and 27 ECU.

Result 4 (Stability) *For both the low and high risk aversion groups as well as the whole experimental population, the Gale–Shapley mechanisms are more stable and "stability-robust" to changes in payoffs than the Boston mechanisms.*

5 Conclusion

In this paper, we have seen that cardinal preferences affect individual behavior in a stylized experimental matching market. In particular, the Gale–Shapley mechanism turned out to be more robust to changes in the preference intensities than the Boston mechanism or, to phrase this as in Abdulkadiroğlu *et al.* (2011), the Boston mechanism induces agents to reveal their cardinal preferences more often. Even though robustness is unrelated to efficiency and stability, this result has policy appeal inasmuch as robustness implies predictability, which is crucial in enrollment planning. A second contribution of the present study to the ongoing debate on Gale–Shapley *vs.* Boston is related to risk aversion. It is widely accepted that individual participants in a market try to manage risk in ways that affect the market as a whole. Matching markets are no exception. One reason for this lies in the fact that the Gale–Shapley mechanism fosters the use of "safe" strategies by the highly risk averse. In fact, we observe that there is a clear tendency for highly risk averse agents to resort to protective strategies under this mechanism.

All this serves as a word of caution for experimentalists (when considering new designs and when bringing ordinal models to the laboratory) and theorists (when constructing new models) both alike, but perhaps more importantly, it should be taken into account by market designers as our results unveil additional dimensions in which the Gale–Shapley and Boston mechanisms can be compared. The Gale–Shapley mechanism is more efficient and more stable than the Boston mechanism in the unconstrained setting, almost independently of the subject pool and the preference intensities.[21] One could conclude from this that the Gale–Shapley mechanism is to be preferred for "small" markets where it is both allowed and no burden for the

[21] The only two exceptions are found in the efficiency levels for the full subject pool and the low risk aversion group when school 2 has a value of 13 ECU.

participants to submit complete full rankings. Our message is different if the market is "large," in the sense that it is unfeasible for the participants to rank all schools and for which policy-makers decide to implement a constrained mechanism.[22] In that case the Boston mechanism performs better in terms of efficiency not only for the whole subject pool (for all preference intensities) but also within the more homogeneous subgroups (for most preference intensities). The Gale–Shapley mechanism is still more stable and, therefore, the ultimate decision of which mechanism to choose in the constrained setting would depend on whether efficiency or stability is considered more desirable.

Appendix A: Holt and Laury (2002)

Table 10 The Holt and Laury (2002) paired lottery choice design. For each of the ten decision situations, we also indicate the expected payoff difference between the two lotteries. Since we did not want to induce a focal point, subjects were not informed about the expected payoff difference during the experiment

Situation	Lottery *A*	Lottery *B*	Difference
1	(1/10 of 2.00 ECU, 9/10 of 1.60 ECU)	(1/10 of 3.85 ECU, 9/10 of 0.10 ECU)	1.17 ECU
2	(2/10 of 2.00 ECU, 8/10 of 1.60 ECU)	(2/10 of 3.85 ECU, 8/10 of 0.10 ECU)	0.83 ECU
3	(3/10 of 2.00 ECU, 7/10 of 1.60 ECU)	(3/10 of 3.85 ECU, 7/10 of 0.10 ECU)	0.50 ECU
4	(4/10 of 2.00 ECU, 6/10 of 1.60 ECU)	(4/10 of 3.85 ECU, 6/10 of 0.10 ECU)	0.16 ECU
5	(5/10 of 2.00 ECU, 5/10 of 1.60 ECU)	(5/10 of 3.85 ECU, 5/10 of 0.10 ECU)	−0.18 ECU
6	(6/10 of 2.00 ECU, 4/10 of 1.60 ECU)	(6/10 of 3.85 ECU, 4/10 of 0.10 ECU)	−0.51 ECU
7	(7/10 of 2.00 ECU, 3/10 of 1.60 ECU)	(7/10 of 3.85 ECU, 3/10 of 0.10 ECU)	−0.85 ECU
8	(8/10 of 2.00 ECU, 2/10 of 1.60 ECU)	(8/10 of 3.85 ECU, 2/10 of 0.10 ECU)	−1.18 ECU
9	(9/10 of 2.00 ECU, 1/10 of 1.60 ECU)	(9/10 of 3.85 ECU, 1/10 of 0.10 ECU)	−1.52 ECU
10	(10/10 of 2.00 ECU, 0/10 of 1.60 ECU)	(10/10 of 3.85 ECU, 0/10 of 0.10 ECU)	−1.85 ECU

[22] Note that the analysis performed in this paper may still apply to a "large" market where agents do not have complete information, but may still have a good idea of how preference distributions look like.

Appendix B: Protective strategies

Consider the game $G = [I, \mathbb{A}, S, g, u]$, where $I = \{1, 2, \ldots, n\}$ is the set of players, $\mathbb{A}$ is the set of outcomes, $S = S_1 \times \cdots \times S_n$ and S_i is the set of strategies of player i, $g : S \to \mathbb{A}$ is an outcome function, and $u = (u_1, \ldots, u_n)$ denotes a vector of utility functions $u_i : \mathbb{A} \to \mathbb{R}$, where $i = 1, 2, \ldots, n$. Take any number $k \in \mathbb{R}$, any $i \in I$, and $s_i \in S_i$. Let $c(k, s_i) = \{s_{-i} \in S_{-i} : u_i(g(s_i, s_{-i})) = k\}$.

Definition 1 (Barberà and Dutta 1995) For any $i \in I$ and $s_i, s_i' \in S_i$, s_i *protectively dominates* s_i', if there exists $k \in \mathbb{R}$ such that

P1. $c(r, s_i) \cap c(r', s_i') = \emptyset$ for all $r \leq k$ and $r < r'$, and
P2. $c(k, s_i) \subset c(k, s_i')$.

It follows from the definition that if s_i protectively dominates s_i', then s_i' does not protectively dominate s_i.

Definition 2 A *protective strategy* is a strategy that is not protectively dominated.

Let us now apply the above definition to our school choice problem. Take, for instance the mechanism BOS_u and the payoff structure 20 ECU. They define a game $G = [I, \mathbb{A}, S, BOS_u, u]$, where $I = \{1, 2, 3\}$ is the set of teachers; $\mathbb{A}$ is the set of matchings; $S = S_1 \times S_2 \times S_3$, where $S_i = \{(X, Y, Z), (X, Z, Y), (Y, X, Z), (Y, Z, X), (Z, X, Y), (Z, Y, X)\}$ is the set of rankings over schools of teacher i, $i \in I$; and $u = (u_1, u_2, u_3)$ is a vector of utility functions. To define player i's utility function u_i, note that i is indifferent between matchings that deliver the same partner, but has strict preferences over matchings that deliver different partners; four situations have to be considered: i may end up unmatched and receive a level of utility of 0, matched to the school ranked third in her preferences and receive a utility of 10, matched to the school ranked second and receive 20, and matched to the school ranked first, receiving a utility of 30.

Now let us consider teacher 1's problem. The other teachers' problems are similar. Note that every strategy guarantees that teacher 1 is matched, so that $c(k, (\times, \times, \times)) = \emptyset$ for all $k < 10$, implying that P2 is never satisfied for k in this range. Therefore, let us compute for each strategy of teacher 1 the set of complementary strategy profiles that match teacher 1 with school Z, with a corresponding utility of 10:

$$\begin{aligned} c\big(10, (X, Y, Z)\big) = \big\{ &\big((X, Y, Z), (X, Y, Z)\big), \big((X, Y, Z), (Y, X, Z)\big), \\ &\big((X, Y, Z), (Y, Z, X)\big), \big((X, Z, Y), (X, Y, Z)\big), \\ &\big((X, Z, Y), (Y, X, Z)\big), \big((X, Z, Y), (Y, Z, X)\big), \\ &\big((Y, X, Z), (X, Y, Z)\big), \big((Y, X, Z), (X, Z, Y)\big), \\ &\big((Y, Z, X), (X, Y, Z)\big), \big((Y, Z, X), (X, Z, Y)\big)\big\}, \end{aligned}$$

$$
\begin{aligned}
c\big(10,(X,Z,Y)\big) = \big\{&\big((X,Y,Z),(X,Y,Z)\big),\big((X,Y,Z),(X,Z,Y)\big),\\
&\big((X,Y,Z),(Y,X,Z)\big),\big((X,Y,Z),(Y,Z,X)\big),\\
&\big((X,Z,Y),(X,Y,Z)\big),\big((X,Z,Y),(X,Z,Y)\big),\\
&\big((X,Z,Y),(Y,X,Z)\big),\big((X,Z,Y),(Y,Z,X)\big),\\
&\big((Y,X,Z),(X,Y,Z)\big),\big((Y,X,Z),(X,Z,Y)\big),\\
&\big((Y,Z,X),(X,Y,Z)\big),\big((Y,Z,X),(X,Z,Y)\big)\big\},\\
c\big(10,(Y,X,Z)\big) = \big\{&\big((X,Y,Z),(Y,X,Z)\big),\big((X,Y,Z),(Y,Z,X)\big),\\
&\big((X,Z,Y),(Y,X,Z)\big),\big((X,Z,Y),(Y,Z,X)\big),\\
&\big((Y,X,Z),(Y,X,Z)\big),\big((Y,X,Z),(Y,Z,X)\big)\big\},\\
c\big(10,(Y,Z,X)\big) = \big\{&\big((X,Y,Z),(Y,X,Z)\big),\big((X,Y,Z),(Y,Z,X)\big),\\
&\big((X,Z,Y),(Y,X,Z)\big),\big((X,Z,Y),(Y,Z,X)\big),\\
&\big((Y,X,Z),(Y,X,Z)\big),\big((Y,X,Z),(Y,Z,X)\big),\\
&\big((Y,Z,X),(Y,X,Z)\big),\big((Y,Z,X),(Y,Z,X)\big)\big\},\\
c\big(10,(Z,\times,\times)\big) = S_2 &\times S_3.
\end{aligned}
$$

Let us start by comparing strategies (X, Y, Z) and (X, Z, Y). Since $c(10, (X, Y, Z)) \subset c(10, (X, Z, Y))$, P2 is fulfilled for $k = 10$. Moreover, P1 is fulfilled for $r = 10$. Since $c(r, (X, Y, Z)) = \emptyset$ for all $r < 10$, P1 is also fulfilled for $r < 10$. It follows that strategy (X, Y, Z) protectively dominates (X, Z, Y) (and (X, Z, Y) does not protectively dominate (X, Y, Z)).

On the other hand, $c(10, (Y, X, Z)) \subset c(10, (Y, Z, X))$ and $c(r, (Y, X, Z)) = \emptyset$ for all $r < 10$ guarantee that (Y, X, Z) protectively dominates (Y, Z, X) (and (Y, Z, X) does not protectively dominate (Y, X, Z)). Furthermore, since $c(10, (Z, \times, \times)) = S_2 \times S_3$, the strategies $(Z, \times, \times)$ are protectively dominated by the other four strategies (and do not protectively dominate any of them).

Comparing $c(10, (X, Y, Z))$ and $c(10, (Y, X, Z))$, P2 is not verified for $k = 10$. To make sure none of these strategies protectively dominates the other, we have to check what happens for higher levels of k. Computing $c(20, (Y, X, Z))$, it is easy to show that $c(10, (X, Y, Z)) \cap c(20, (Y, X, Z)) \neq \emptyset$, so that P1 fails to hold for $k > 10$ (with $r = 10$ and $r' = 20$) and (X, Y, Z) does not protectively dominate (Y, X, Z). On the other hand, (Y, X, Z) does not protectively dominate (X, Y, Z) as $c(10, (Y, X, Z)) \cap c(30, (X, Y, Z)) \neq \emptyset$ and P1 fails to hold for $k > 10$ (with $r = 10$ and $r' = 30$).

To ensure (X, Y, Z) is not protectively dominated, we still have to compare it with (Y, Z, X). Note that P2 is not verified for $k = 10$. As for $k > 10$, it can easily be shown that $c(10, (Y, Z, X)) \cap c(30, (X, Y, Z)) \neq \emptyset$, so that P1 fails (with $r = 10$ and $r' = 30$). Similarly, (X, Z, Y) does not protectively dominate (Y, X, Z) as P2 is not verified for $k = 10$ and $c(10, (X, Z, Y)) \cap c(20, (Y, X, Z)) \neq \emptyset$, invalidating P1 for $k > 10$ (with $r = 10$ and $r' = 20$).

Therefore, strategies (X, Y, Z) and (Y, X, Z) are not protectively dominated. The set of protective strategies of teacher 1 in BOS_{u20}—in fact, in any game induced by BOS_u—is $\{(X, Y, Z), (Y, X, Z)\}$.

Protective strategies can readily be calculated for the other mechanisms. In fact, following the informal description of protective strategies in Barberà and Dutta (1995, p. 289), in our school choice problem protective behavior means the following. For *any* distribution over the others' strategy profiles: First, choosing a strategy that guarantees access to a school; second, among these, if possible, one that maximizes the probability of obtaining the best or the second best schools; and finally, within this set of strategies and whenever possible, picking one that maximizes the probability of being matched to the best school.

As such, since under GS_u telling the truth never hurts and, for some strategy profiles of the others, leads to a better school slot, truth-telling is the unique protective strategy under this mechanism.[23] In what constrained mechanisms are concerned, protective behavior ensures in the first place that a subject is not left unassigned for any profile of complementary strategies. This implies using a strategy where the least preferred school is ranked first under BOS_c—the unique protective strategy under this mechanism—and, given that acceptance is deferred in GS_c, ranking the least preferred school first or second in the list under this mechanism. Moreover, given that ranking the least preferred school second increases the chances of being assigned to a better school both (X, Z, Y) and (Y, Z, X) are protective strategies for teacher 1 in GS_c.

References

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

Abdulkadiroğlu, A., Pathak, P., & Roth, A. (2005a). The New York city high school match. *American Economic Review, Papers and Proceedings*, *95*(2), 364–367.

Abdulkadiroğlu, A., Pathak, P., Roth, A., & Sönmez, T. (2005b). The Boston public schools match. *American Economic Review, Papers and Proceedings*, *95*(2), 368–371.

Abdulkadiroğlu, A., Pathak, P., & Roth, A. (2009). Strategy-proofness versus efficiency in matching with indifferences: redesigning the NYC high school match. *American Economic Review*, *99*(5), 1954–1978.

Abdulkadiroğlu, A., Che, Y.-K., & Yasuda, Y. (2011). Resolving conflicting preferences in school choice: the Boston mechanism reconsidered. *American Economic Review*, *101*(1), 399–410.

Barberà, S., & Dutta, B. (1982). Implementability via protective equilibria. *Journal of Mathematical Economics*, *10*(1), 49–65.

Barberà, S., & Dutta, B. (1995). Protective behavior in matching models. *Games and Economic Behavior*, *8*(2), 281–296.

Calsamiglia, C., Haeringer, G., & Klijn, F. (2009). Constrained school choice: an experimental study. *American Economic Review*, *100*(4), 1860–1874.

Chen, Y., & Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic Theory*, *127*(1), 202–231.

Dubins, L., & Freedman, D. (1981). Machiavelli and the Gale–Shapley algorithm. *The American Mathematical Monthly*, *88*(7), 485–494.

[23] Barberà and Dutta (1995) showed that under GS_u truth-telling is the unique protective strategy for all participants on both sides of a *two-sided* matching market.

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

Featherstone, C., & Niederle, M. (2008). *Ex ante efficiency in school choice mechanisms: an experimental investigation*. Working paper, Stanford University.

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

Gale, D., & Shapley, L. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, *69*(1), 9–15.

Greiner, B. (2004). *The online recruitment system ORSEE 2.0—a guide for the organization of experiments in economics*. Working Paper Series in Economics 10, University of Cologne.

Holt, C., & Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644–1655.

Miralles, A. (2008). *School choice: the case for the Boston mechanism*. Working paper, Boston University.

Nicolò, A., & Rodríguez-Álvarez, C. (2012). Transplant quality and patients' preferences in paired kidney exchange. *Games and Economic Behavior*, *74*(1), 299–310.

Pais, J., & Pintér, Á. (2008). School choice and information: an experimental study on matching mechanisms. *Games and Economic Behavior*, *64*(1), 303–328.

Roth, A. (1982). The economics of matching: stability and incentives. *Mathematics of Operations Research*, *7*(4), 617–628.