THE JOURNAL OF ECONOMIC EDUCATION, *44*(1), 47–57, 2013

ISSN: 0022-0485 print/2152-4068 online
DOI: 10.1080/00220485.2013.740391

ECONOMIC INSTRUCTION

Choosing Partners: A Classroom Experiment

Carl T. Bergstrom, Theodore C. Bergstrom, and Rodney J. Garratt

The authors describe a classroom experiment designed to present the idea of two-sided matching, the concept of a stable assignment, and the Gale-Shapley deferred-acceptance mechanism. Participants need no prior training in economics or game theory, but the exercise will also interest trained economists and game theorists.

Keywords *experiment, matching, stability*

JEL codes *A22, A23, C78*

In standard models of economic choice, consumers select their favorite commodity bundles from among those that they can afford. Nobody asks the chosen commodities how they feel about the consumers who choose them. Students, however, are familiar with real-life decision problems in which the outcomes depend crucially on the preferences of both the choosers and the chosen. The so-called "marriage problem" is a classic example. Consider a population of men and women who seek partners of the opposite sex. Each person has a personal ranking of possible partners. These rankings may be similar or wildly different among individuals. Clearly, it is unlikely that partners can be assigned in such a way that everyone gets their first choice. What can we expect to happen in a "market" where partnerships must be formed by mutual consent?

This problem is posed in a beautiful article by David Gale and Lloyd Shapley, "College Admissions and the Stability of Marriage" (1962). Gale and Shapley define a stable marriage assignment as a matching of partners such that no two persons of opposite sex prefer each other to their assigned partners. This seems a reasonable requirement for stability, because if two individuals discover that they prefer each other to their current partners, we might expect them to abandon their partners and run off together. Gale and Shapley show that if each participant has a strict preference ordering (no ties) over members of the opposite sex, then there always must

Carl T. Bergstrom is a professor of biology at the University of Washington (e-mail: cbergst@u.washington.edu). Theodore C. Bergstrom (e-mail: tedb@econ.ucsb.edu) and Rodney J. Garratt (e-mail: garratt@econ.ucsb.edu) are professors of economics at the University of California, Santa Barbara. Theodore C. Bergstrom is the corresponding author.
Carl T. Bergstrom's research was supported in part by NSF ATB-1038590. Research by Theodore C. Bergstrom and Rodney J. Garratt was supported in part by NSF 0851357.

be at least one stable marriage assignment. They demonstrate this by means of a pair of simply implemented algorithms, known as "deferred-acceptance" mechanisms, which are guaranteed to produce a stable marriage assignment. Remarkably, one of these algorithms, the male-proposal algorithm, produces a stable assignment that is better for all males (and worse for all females) than any other stable marriage assignment, while its counterpart, the female-proposal algorithm, yields a stable assignment that is better for all females (and worse for all males) than any other stable assignment.

We highly recommend reading the original Gale-Shapley article. The article is very accessible and is a remarkable illustration of the way that simple mathematical reasoning can illuminate important aspects of social interaction. The article could be profitably read either before or after the experiment is performed. More detailed discussions of matching theory and its applications can be found in books by Alvin Roth and Marilda Sotomayor (1990) and by Donald Knuth (1991).

THE CLASSROOM EXPERIMENT

This experiment works well with participants who have never taken a course in economics or game theory. It is also interesting for advanced students. We have run it with high school students attending a math camp, with college freshmen who have not taken any economics courses, with undergraduate biology students, with upper-division economics students in a game theory class, with economics graduate students, and with biology graduate students. It works well as a stand-alone module because students readily understand the problem without preparatory study, and it offers an easily grasped lesson on the applicability of economic reasoning to everyday life. In a game theory class, this lesson is an effective introduction to cooperative game theory.

We have conducted this experiment with numbers of participants ranging from 12 to 34. In larger classes, we have run the experiment as a "demonstration" with 18 to 34 active participants at a time. As we move from one treatment to another, we rotate the participants so that all or most students in the class get a chance to participate actively.

Preliminary Instructions

Get at least two sheets of construction paper of each of a variety of colors.[1] Make cards by cutting out (at least) one circle and one square of each color, about 6 inches wide. On the back of each card, write a ranking of colors as shown in figure 1. (Suggestions of ways to assign the rankings are found in later sections below.) Give each participating student either a colored circle or a colored square. Students who are given circles will seek a match with someone who has a square, and those who have squares will seek a match with someone who has a circle. The higher the number that one's card assigns to a color, the more its holder would like a partner with a card of that color.

First Treatment: Free-Form Matching

Tell the students that they will play a game in which circles must match with squares and squares must match with circles. A player's score for the game is the value that the player's own card

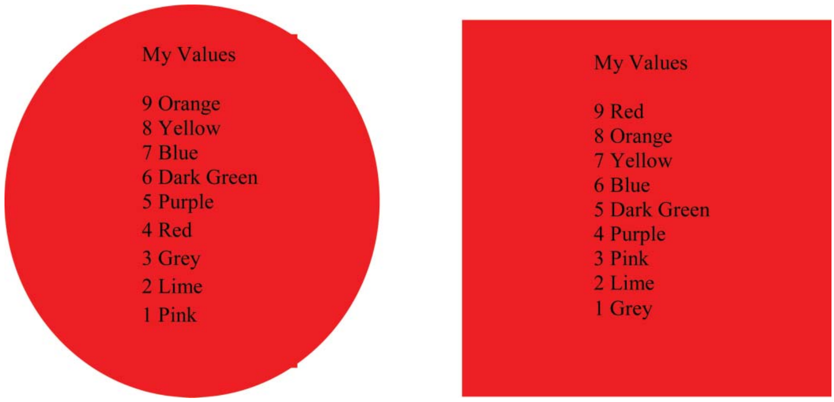

FIGURE 1 Sample game cards (color figure available online).

assigns to the color of the card with whom they match. Tell them that they have four minutes to find a partner and that as soon as they do so, the newly-attached partners should report to the instructor, who will record the color pairs of reported matches on the blackboard along with the score each player receives from the match. Give a warning ten seconds before the deadline. If any students have not chosen partners at the end of the time period, they will remain unmatched.

Perhaps because the situation seems familiar, students find these instructions easy to understand. Almost all of them participate in animated fashion. When the matching is completed, you will have recorded on the blackboard a table of matched pairs and scores similar to that shown in table 1.

The next step is to check whether this assignment is stable. To do so, start at the top of the list. Ask: "Do any squares prefer the green circle to your current match? If so, please hold up your colored squares." Then ask the green circle: "Do you prefer any of these squares to your current match?" If the answer is "Yes," then point out to the class that the matching arrangement shown in table 1 is not stable, because you have found two people of opposite types who prefer each other to their assigned partner. (To motivate the notion of stable assignment, this might be a good time to ask the class what they think would happen in real life if two people preferred each other to their current partners.) If the answer is "No," then proceed down the list, asking the squares: "Do

TABLE 1
Unstructured Matching: Round 1

Circle	Square
Green, 9	Purple, 4
Purple, 6	Blue, 7
Blue, 6	Orange, 7
Orange, 7	Green, 6
...	...
...	...
...	...

any of you squares prefer (the next colored circle on the list) to your current partner?” Usually as you move down the list, you will find one or more instances of unstable matches, and thus can tell the students that the current assignment is unstable. You might then ask them to speculate about whether they think there is any assignment that would be stable.

If you did not reach a stable assignment on the first effort and if time seems to be available, have the students repeat the matching process. This time, the matching will go faster, as people know more about their prospects. Those who learned that they had a better potential partner than their previous match are likely to try to improve their outcomes. The number of unstable matches is likely to diminish, possibly to zero. Leave the resulting table of matches up on the blackboard for possible later discussion.

Second Treatment: The Circles-Propose Process

Ask all of the students with squares to line up at the front of the room, holding their colored squares in front of them.

Stage 1

Ask the circles to line up in front of their favorite squares. Then ask the squares to say “maybe” to their favorite of the circles that are lined up in front of them and “no” to each of the others. Circles who were told “no” must then step to the back of the room.

Stage 2

Ask the circles at the back of the room to line up in front of their second favorite square (whether or not there is anybody already standing in front of this square.) Then ask the squares to consider the circles currently lined up in front of them and to say “maybe” to their favorite among them (this might or might not be the circle they said “maybe” to on the previous round) and “no” to the rest of them. The circles who are rejected must step to the back of the room.

Repeat the process, with rejected circles lining up in front of the square they like best among those who have not (yet) refused them. Eventually, there will be exactly one circle lined up in front of each square. The circles and squares that are now matched are the final assignment. On the blackboard, make a table showing the assigned matches and the score that each individual receives from this assignment (see table 3 below).

This process will necessarily conclude after a finite number of iterations and the result must be a stable assignment. Ask the class to check whether this assignment is stable, using the same query method used to find whether the free-form assignment was stable.

This may be a good time to offer the students a simple verbal proof that the outcome is a stable assignment.

Proof

Notice that by the rules of the process, if a square ever proposes to a circle, the circle will never have to settle for a square that it likes less well. Also, by the rules of the process, if a square prefers one circle to another, the square will propose to the first of these circles before the second.

TABLE 2
Matching Payoffs for $n = 9$

	Square								
Circle	Green	Purple	Red	Orange	Yellow	Blue	Lime	Grey	Pink
Green	(4, 9)	(9, 4)	(8, 5)	(7, 6)	(6, 7)	(5, 8)	(3, 6)	(1, 7)	(2, 5)
Purple	(5, 8)	(4, 9)	(9, 4)	(8, 5)	(7, 6)	(6, 7)	(1, 5)	(2, 6)	(3, 4)
Red	(6, 7)	(5, 8)	(4, 9)	(9, 4)	(8, 5)	(7, 6)	(2, 4)	(3, 5)	(1, 9)
Orange	(7, 6)	(6, 7)	(5, 8)	(4, 9)	(9, 4)	(8, 5)	(3, 9)	(1, 4)	(2, 8)
Yellow	(8, 5)	(7, 6)	(6, 7)	(5, 8)	(4, 9)	(9, 4)	(1, 8)	(2, 9)	(3, 7)
Blue	(9, 4)	(8, 5)	(7, 6)	(6, 7)	(5, 8)	(4, 9)	(2, 7)	(3, 8)	(1, 6)
Lime	(9, 1)	(1, 2)	(4, 2)	(8, 1)	(3, 2)	(2, 3)	(7, 2)	(5, 3)	(6, 3)
Grey	(4, 2)	(3, 3)	(2, 1)	(1, 2)	(6, 3)	(5, 1)	(9, 1)	(7, 1)	(8, 1)
Pink	(1, 3)	(7, 1)	(8, 3)	(3, 3)	(2, 1)	(9, 2)	(4, 3)	(5, 2)	(6, 2)

Note: The first number of each pair in the matrix gives the points earned by the circle in the match. The second number of each pair in the matrix gives the points earned by the square in the match.

Now suppose that the match reached by our algorithm is unstable. Then there must be two people who would prefer one another to their assigned matches. Suppose (without loss of generality) that these people are Orange Circle and Purple Square. Because Purple Square prefers Orange Circle to his assigned partner, he must have proposed to Orange Circle before proposing to his current partner. But, if Orange Circle prefers Purple Square to her current partner, then she would never have had to settle for her current partner. So it is not possible that both Purple Square and Orange Circle prefer each other to the partners assigned by the circles-propose process.

Third Treatment: The Squares-Propose Process

What happens if we apply the Gale-Shapley mechanism, but reverse the roles of circles and squares? For this treatment, have the circles line up at the front of the room, holding their colored circles in front of them. Ask the squares to line up in front of their favorite circles. Ask the circles to say "maybe" to the square that they like best from among those in front of them and "no" to

TABLE 3
Matches and Payoffs from Circles-Propose Process

Circle	Square
Green, 9	Purple, 4
Purple, 9	Red, 4
Red, 9	Orange, 4
Orange, 9	Yellow, 4
Yellow, 9	Blue, 4
Blue, 9	Green, 4
Lime, 7	Lime, 2
Pink, 6	Pink, 2
Grey, 7	Grey, 1

TABLE 4
Matches and Payoffs from Squares-Propose Process

Circle	Square
Green, 4	Green, 9
Purple, 4	Purple, 9
Red, 4	Red, 9
Orange, 4	Orange, 9
Yellow, 4	Yellow, 9
Blue, 4	Blue, 9
Lime, 7	Lime, 2
Pink, 6	Pink, 2
Grey, 7	Grey, 1

the others. Iterate the process, with rejected squares proposing to their preferred circle among those who have not yet rejected them. When each circle is lined up with one square, record the colors and achieved scores of the assigned partners (see table 4).

Again, check whether this assignment is stable, using the procedure applied in the previous treatments. Typically, you will have found two distinct stable outcomes from the two Gale-Shapley processes. If your final free-form outcome was stable, then it is likely to be distinct from either of the two Gale-Shapley outcomes.

Compare the tables for the circles-propose process and the squares-propose process. If you obtained a stable match using the free-form process, then you can refer to this table as well. Ask class members to comment on which type, circles or squares, does better under each of the processes. It should be apparent that the circles-propose process produces the best match for circles and the worst match for squares. Likewise the squares-propose process produces the best match for squares and the worst match for circles. Inform students that these are general properties of the deferred-acceptance algorithm that hold for any specification of preferences of different colored circles over different colored squares and different colored squares over different colored circles.[2]

ASSIGNING PARTNER VALUES

Preassigned Values

We have conducted the experiment several times with rankings assigned as in table 2. With this configuration, table 3 shows the matched pairs and resulting payoffs from the circle-propose version of the deferred-choice mechanism and table 4 shows the matches and payoffs from the squares-propose process. Comparing tables 3 and 4, we see that Grey Circle and Square, Pink Circle and Square, and Lime Circle and Square are matched together whether circles or squares propose and that circles of all other colors get higher payoffs when circles propose than when squares propose. When the circles propose, the circles of each of the "rainbow colors" green, purple, red, orange, yellow, and blue get their most preferred partner. When the squares propose, squares of each of the rainbow colors get their most preferred partner. These are not the only stable matching assignments. In fact, if your free-form treatment reached a stable assignment, it is most likely to be better for some of the squares than the circles-propose outcome and better for some of the circles than the squares-propose treatments.[3]

The configuration reported here is for exactly 18 students. If there are fewer students, the simplest modification is to leave out some or all of the colors of grey, pink, and lime. If there are more students, one can make duplicates of some or all of the color pairs.[4]

Self-Provided Color Rankings

Another more free-form way to provide rankings is to let students rank their "favorite colors." At the start of class, hand out n colored circles and n colored squares with the rankings left blank, as shown in figure 2 for the case of $n = 10$. Tell the students to place unique integer values from 1 to 10 next to the colors shown in accordance with their preferences. Remind them to give the score of 10 to their favorite color, 9 to their second favorite color, and so on. This method of ranking seems to work well and is consistent with the conventional notion that an ideal mate is a perfect 10. There are a few advantages to this free-form approach. First, the students will not suspect that the payoffs are rigged. Second, students have direct ownership of their rankings; they are not imposed upon them. Finally, preferences over colors overlap in much the same way preferences over members of the opposite sex do, creating realistic competition for matches. We have had success running the free-form method and recommend it.

ITEMS FOR CLASS DISCUSSION

After the experiment, students will be primed for a discussion of how the results relate to real world courtship rituals. Students will be quick to draw comparison of the matching of squares and circles to that of men and women. They are likely to notice that the Gale-Shapley procedure with one shape proposing is similar to traditional courtship roles in which men ask women and women respond only if asked. Biologists and social scientists sometimes assume, when they see

FIGURE 2 Blank sample game cards (color figure available online).

females rejecting males rather than males rejecting females, that female choice is dominating and hence females determine the mating structure of the population to their advantage.[5] Thus, if one observes a males-propose deferred-acceptance procedure in nature, it might seem that this is a "female choice" system. Females have the power to accept or reject males, and even to keep males on hold, stringing them along only to be later rejected in favor of preferable suitors. One might think that this form of courtship empowers females relative to males. In fact, we have seen that the opposite is true. The "males-propose, females-choose" Gale-Shapley mechanism generates the matching that is best for males and worst for females among all stable matchings.

A thought-provoking line of inquiry is to ask students whether they believe that the Gale-Shapley males-propose mechanism is a realistic model of actual courtship. This is likely to provoke discussion of how common it is today for females to do the asking and males the responding. It is interesting to note that neither males-propose nor the females-propose deferred-acceptance mechanism accounts for fear of rejection. If people are afraid of being rejected and if they believe that their favorites of the opposite sex are likely to be very popular, they may not venture to ask their top choices, even when such offers might get positive responses. One of the co-authors recalls that "We used to complain about this in seventh grade. As a guy, you have to ask a girl to the dance, which feels like a bad deal because your feelings go out on the line. Girls complained about this too, because many guys were too damn chicken to ask and so lots of ideal matchings never got made. If girls got to ask, they figured, at least they could get the job done!"

It may be disturbing to romantics to realize that even if the current assignment of partners is stable, it does not follow that partners were "meant for each other" in the sense that neither could do better in some other stable arrangement of matches. This is dramatically illustrated by the stable assignments shown in table 3 for the circles-propose and table 4 for the squares-propose process. When the circles propose, the resulting assignment is stable and green circle gets her first choice (Purple Square) as a match. But, although this assignment is stable, Purple Square may be less than ecstatic, because Green Circle is his sixth choice. Purple Square may well lament that had things only turned out a bit differently, there is another stable assignment, shown in table 4, in which Purple Square gets his first choice, namely Purple Circle.

There is yet worse news for romantics. Knuth (1991) pointed out that for any two individuals paired under a stable matching, if there exists another stable matching in which they are not together, then one mate must prefer the matching that keeps them together, and one must prefer the matching that separates them. Therefore, in a stable matching, if you prefer your current mate to another attainable mate (i.e., someone you could be paired with in a stable matching), your mate necessarily prefers some other attainable mate to you!

The Gale-Shapley deferred-acceptance mechanism will generate a stable matching if all players play according to their true preferences. But could it be in the interest of some to practice deception? For example, in the males-propose version, would it ever benefit a female with multiple suitors to reject the one that she likes best of those who have proposed so far? Would it ever be in the interest of a male to propose to a female who is not the one he likes best among those who have not rejected him? Roth (1982) showed that in the males-propose deferred-acceptance mechanism, it is never in any male's interest to deceive about his preferences. But as Roth, and Bergstrom and Manning (1983) showed, under the males-propose mechanism, it is likely that at least some females can gain by deception. More generally, these authors showed that no mechanism can be found that is guaranteed to produce stable matches without being manipulable by "deceptive" practices.

TABLE 5
Preference Chart

Person	First choice	Second choice	Third choice
Al	Alice	Betsy	Clara
Bill	Betsy	Alice	Clara
Charlie	Alice	Clara	Betsy
Alice	Bill	Al	Charlie
Betsy	Al	Bill	Charlie
Clara	Al	Charlie	Bill

Lest students get the idea that the existence of a stable match is obvious, you may want to present the "roommate problem," in which a population of individuals must be matched into pairs without regard to sex. Gale and Shapley show, by means of a simple example, that this problem does not necessarily have a stable match.

As the title of their article suggests, Gale and Shapley show that the deferred-choice mechanism can also be applied to many-to-one matching problems such as the matching of students to colleges. An algorithm similar to the Gale-Shapley deferred-acceptance mechanism has long been used for the assignment of interns to hospitals. In the early 1950s, without the benefit of reading the Gale-Shapley article, hospital administrators successfully implemented a version of a many-to-one deferred-acceptance algorithm for designating interns to hospitals. This mechanism, which is still in place in slightly modified form, is known as the National Resident Matching Program (NRMP). A nice discussion of the history of the NRMP and recent changes can be found in Roth and Peranson (1999). Another example of a many-to-one matching is the assignment of members to fraternities and sororities. Susan Mongell and Al Roth (1991) wrote an interesting article on the matching process used by college sororities.

HOMEWORK PROBLEMS

You may wish to assign some homework that will help students to develop and expand their understanding of the deferred-acceptance mechanism.

(1) Consider a matching problem with equal numbers of men and women. Suppose that all of the men rank the women in the same order.
   (a) Show that if all of the women also rank the men in the same order, then there is exactly one stable matching. Describe this matching.
   (b) Show that if all of the men rank the women in the same order, but different women have different rankings of the men, then there is still exactly one stable matching. Describe this matching.

(2) Three men, Al, Bob, and Charlie, and three women, Alice, Betsy, and Clara, seek partners of the opposite sex. Their preferences over members of the opposite sex are shown in table 5.
   (a) What pairs are matched by the men-propose deferred-acceptance mechanism?
   (b) What pairs are matched by the women-propose deferred-acceptance mechanism?

(c) Suppose that partners are assigned by the men-propose deferred-acceptance mechanism. Alice has learned the preferences of all other persons and she believes that they will all play the game truthfully. Show how Alice can get a partner that she likes better than she would get if she played "honestly" by "pretending" that her preferences are different from her true ranking.

(3) Prove that with the assignment of rankings shown in table 2, every stable assignment of partners matches Grey Square with Grey Circle, Pink Square with Pink Circle, and Lime Square with Lime Circle.

(4) You have probably heard the expression "I wouldn't marry you if you were the last man (woman) on earth." (Though we hope this was not directed to you, personally.) Suppose that men and women have preference orderings over all members of the opposite sex, but it is possible that some people rate some members of the opposite sex as "worse than remaining single." Suggest a reasonable definition of a stable assignment of partners in this case. Find a modified version of the Gale-Shapley deferred-choice mechanism that would generate a stable assignment according to your definition.

EPILOGUE

Just as this article went to press, we learned that this year's Nobel Prize in economics was awarded jointly to Alvin Roth and Lloyd Shapley "for the theory of stable allocations and the practice of market design." This is one of the rare occasions where a work that led to a Nobel Prize is profound, yet can be readily explained to a broad range of students. This experiment, which is based on the path-breaking article by David Gale and Lloyd Shapley, is designed to introduce the assignment problem and the Gale-Shapley solution by allowing students to experience a two-sided matching environment first-hand. Students may be motivated to follow up by reading some of the contributions of Alvin Roth and his colleagues in applying this theory to a rich variety of practical problems.

NOTES

1. If you plan to have *n* participants without duplicate colors, you will need *n/2* colors. The detailed instructions below include one design for 18 players and 9 colors. For a larger number of players, you may want to duplicate some color pairs, as explained below.
2. The proof of these claims, which uses a relatively simple induction argument, can be found in Gale and Shapley (1962).
3. In any stable assignment, Grey will be matched with Grey, Pink with Pink, and Lime with Lime. Given the Latin Square arrangement of preferences of the "rainbow colors", it can be shown (See e.g., Knuth 1991, pp. 3–4) that there are exactly 6 stable ways that the rainbow colors could be matched. With this configuration, in any stable match, all squares will get their nth most preferred match where $n = 1, 2, 3, 4, 5,$ or 6.
4. When there are duplicate types, the deferred-choice algorithms may force someone to say "maybe" to one of two equally desirable individuals and no to the other. Those who have been refused by one person of a given color may propose to another of the same color, but may not propose to any individual who has previously refused them.
5. Bergstrom and Real (2000) explore the application of matching theory to nonhuman animals.

REFERENCES

Bergstrom, C. T., and L. A. Real. 2000. Towards a theory of mutual mate choice: Lessons from two-sided matching. *Evolutionary Ecology Research* 2: 493–508.

Bergstrom, T. C., and R. Manning. 1983. Can courtship be cheatproof? Working paper. Santa Barbara, CA: University of California, Santa Barbara, Economics Department. http://repositories.cdlib.org/ucsbecon/bergstrom/1983C (accessed September 4, 2012).

Gale, D., and L. S. Shapley. 1962. College admissions and the stability of marriage. *American Mathematical Monthly* 69: 9–14.

Knuth, D. E. 1991. *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*. Providence, RI: American Mathematical Society.

Mongell, S., and A. E. Roth. 1991. Sorority rush as a two-sided matching mechanism. *American Economic Review* 81: 441–64.

Roth, A. E. 1982. The economics of matching: Stability and incentives. *Mathematics of Operations Research* 7: 617–28.

Roth, A. E., and E. Peranson. 1999. The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American Economic Review* 89: 748–80.

Roth, A. E., and M. Sotomayor. 1990. *Two-sided matching: A study in game-theoretic modelling and analysis*. Econometric Society Monographs. New York, NY: Cambridge University Press.