Matching and Chatting:

An Experimental Study of the Impact of Network Communication on School-Matching Mechanisms *

Tingting Ding and Andrew Schotter†

August 7, 2014

Abstract

While, in theory, the school matching problem is a static non-cooperative one-shot game, in reality the "matching game" is played by parents who choose their strategies after consulting or chatting with other parents in their social networks. In this paper we compare the performance of the Boston and the Gale-Shapley mechanisms in the presence of chatting through social networks. Our results indicate that allowing subjects to chat has an important impact on the strategies they choose and is welfare increasing. In addition, chatting appears to enhance the rationality of subjects and the stability of the matching outcomes.

Key Words: School Choice, Matching, Mechanism Design, Networks, Chat

JEL Classification: C78, C91, C72

*Support for this research was funded by the National Science Foundation grant number 1123045. We would like to thank Tayfun Sönmez, Leeat Yariv and Tim Salmon for their valuable comments, and thank the participants of the 2012 International ESA conference, the 2013 North American ESA Conference, the University of Texas at Dallas Economics seminar, and the Santa Barbara Conference on Behavioral Economics for helpful discussion. We would also like to thank the Center for Experimental Social Science for their lab support as well as Anwar Ruff for his programming assistance.

†Ding: Department of Economics, New York University, 19 West 4th St., New York, NY, 100012. (email: tingting.ding@nyu.edu); Schotter: Department of Economics, New York University, 19 West 4th St., New York, NY, 100012. (email: andrew.schotter@nyu.edu)

"One school choice strategy is to find a school you like that is undersubscribed and put it as a top choice, OR, find a school that you like that is popular and put it as a first choice and find a school that is less popular for a 'safe second choice."

Advice from the West Zone Parents Group cited in Pathak (2011) [Boston Mechanism]

"I was a Type 2 subject and I knew I would have priority over Object A (*subject's second best object*) and I thought it would be safe to have Object A as a second choice because even if I didn't get my first choice, and other subjects were temporarily assigned to Object A, I would have priority over them. Then, I had Object B (*subject's worst object*) as my last choice because I don't have priority over it, and who wants to settle for $4 if you can try for $16 or $24?" (Italics added.)

Advice from Subject 159 in session 02082012_1247 [Gale-Shapley Mechanism]

In recent years there has been a great deal of interest in designing matching mechanisms that can be used to match public school students to schools (the student matching problem).[1] The design of matching mechanisms relies on a combination of economic theory and common sense, and these attempts have proven extremely useful in helping organizations solve this complicated problem.[2] The premise of this paper is that when testing mechanisms we must do so in the environment in which they are used in the real world rather than in the environment envisioned by theory. More precisely, in theory the school matching problem is a static one-shot game played by parents of children seeking places in a finite number of schools and played non-cooperatively without any form of communication or commitment between parents. However, in the real world, the school choice program is played out in a different manner. Typically parents choose their strategies after consulting with other parents in their social networks and exchanging advice on both the quality of schools and the proper way they should play the "school-matching game." (See the quotes above.) In addition, parents who have engaged in the matching mechanism in the past may also communicate with parents currently in the match and offer their words of wisdom. We can call these two systems of advice the "horizontal" and the "vertical" advice systems. This paper focuses on horizontal advice networks and a companion paper (Ding and Schotter, 2014) focuses on vertical or inter-generational advice. The question we ask here is whether chat between parents (similar to naive advice defined in Schotter (2003)) affects the strategies they choose, and if so, whether it does so in a welfare-increasing or decreasing manner.[3]

[1]See Balinski and Sönmez (1999); Abdulkadiroğlu and Sönmez (2003); Ergin and Sönmez (2006); Erdic and Ergin (2008); Abdulkadiroğlu et al. (2009); Pathak and Sethuraman (2011); Kesten and Ünver (2013); Abdulkadiroğlu et al. (2011); Haeringer and Klijn (2009) for some of the central theoretical contributions, Chen and Sönmez (2006); Pais and Pintér (2008); Calsamiglia et al. (2010); Featherstone and Niederle (2011); Klijn et al. (2013); Chen and Kesten (2013); Guillen and Hing (2013) for experimental studies, and Abdulkadiroğlu et al. (2005a) and Abdulkadiroğlu et al. (2005b) for the summary of school choice reforms in New York City and Boston.

[2]See Toch and Aldeman (2009) for a news article praising the New York City matching scheme, as well as Herszenhorn (2003) "Revised Admissions for High School," New York Times.

[3]In other work (Schotter and Sopher, 2003, 2007; Nyarko et al., 2006; Iyengar and Schotter, 2008),

More precisely, in this paper we compare the performance of two variants of the Boston (Boston 16 and Boston 10) and the Gale-Shapley mechanisms, two often used school matching algorithms, in the presence of chatting through social networks. In each mechanism subjects must submit rankings over three objects whose values to them are either \$24, \$16, and \$4, in the Boston 16 and Gale-Shapley mechanisms, or \$24, \$10, and \$4 in the Boston 10 mechanism. We allow subjects in our experiments to play the "matching game" twice, once before (Phase 1) and once after chatting (Phase 2).[4] Which outcome, Phase 1 or Phase 2, is payoff relevant is determined randomly, so subjects have incentives to choose those strategies that maximize their payoffs in each phase. Given our design, the behavior in Phase 1 is a test of the static mechanisms, while the difference between Phase 1 and Phase 2 defines the impact of chat.[5]

Among the questions we attempt to answer are:

1. In the absence of chat (Phase 1), is the performance of our baseline Boston 16 and Gale-Shapley mechanisms equivalent? I.e., are the strategies employed across subject types and the welfare of the mechanisms the same?

2. Do subjects who chat change their Phase-1 submitted rankings more than those who are isolated and do not chat?

3. Is chatting welfare increasing?

4. Is chat rationality increasing? I.e., does chatting lead subjects to decrease their use of irrational strategies or to employ strategies in Phase 2 that are best responses to the behavior of other subjects in Phase 1?

5. Is chat stability increasing? I.e., does chat lead to an increase in stable outcomes?

6. If chatting is beneficial, does it matter to whom you chat? I.e., does it matter whether your social network is populated by people like you or different from you?

7. Does the content of chat change as we look across mechanisms and subject types?

8. Does chatting influence welfare differently when subjects have priority rights or not? (In school matching programs some students (subjects) have priority for admission to some schools (objects).)

advice received by chatting has proven to have a very powerful influence on decision makers in the sense that advice tends not only to be followed but typically has a welfare increasing consequence.

[4]For the subjects who are not allowed to chat, we ask them to write down the logic of their strategies by introspection when the others chat.

[5]The only paper we know of that discusses advice in the context of matching is Guillen and Hing (2013). Their design is very different from ours.

While we will not discuss the answers to these questions in detail in this introduction, as a preview we find that chatting has many beneficial characteristics. More precisely, over all, allowing people to chat not only increases their welfare but also their rationality and the stability of the matches created. This is interesting since we find that there is no difference in subject behavior or subject welfare between our baseline Boston 16 and Gale-Shapley mechanisms in the absence of chat in Phase 1, which may lead to the conclusion that these two mechanisms are equivalent.[6] After chatting is introduced, however, behavior and welfare diverge across these mechanisms. The implication, therefore, is that including chat into an experimental design on matching enhances its external validity since in the real world chatting is ubiquitous.

If chatting is beneficial, it might make sense to ask if it matters with whom you chat. We find that, those subjects who communicate with people of their own experimental type (i.e., who have the same induced preferences and priorities) tend to change their submitted preference rankings more between phases than those who communicate with people of different types but their payoffs increase less, indicating that sometimes the beneficial aspect of chatting is to persuade subjects who are already using good strategies not to change.

With respect to welfare, chatting appears to influence welfare differently between subjects with priority rights and those without. Though there is little welfare change among subjects with priority rights between phases, chatting does significantly change the welfare of those without priority rights. If one equates subjects without priority rights with people living in disadvantaged areas (where they may have priority rights but in the least desirable schools), our results indicate that allowing better communication between parents in those areas may improve their welfare. In addition, our results show

[6] Previous experimental papers on school matching present mixed evidence on the performance of various mechanisms. In terms of efficiency, for example, Pais and Pintér (2008) and Calsamiglia et al. (2010), find no difference across all of their treatments. Other papers have mixed results. Chen and Sönmez (2006) find in their "designed environment" that the efficiency of the Gale-Shapley mechanism is greater than that of the Boston mechanism, though in their "random environment" no such difference is found. Featherstone and Niederle (2011) and Klijn et al. (2013) find under some circumstances the efficiency in the Boston mechanism is greater than that in the Gale-Shapely mechanism. The results on strategy differences are also unclear. Klijn et al. (2013) find differences in strategic behavior in all their treatments. Pais and Pintér (2008) find that when subjects have zero information (only know of their own preference rankings) there is no difference in the fraction of subjects submitting truthful preferences while, when subjects have more information about school priorities and others' preferences, they are more likely to tell the truth in the Gale-Shapley mechanism. Featherstone and Niederle (2011) find that while in their "uncorrelated preference environment" there is no difference in truth-telling across the Boston and Gale Shapley mechanisms, in their "aligned preference environment" subjects are more likely to tell the truth in the Gale-Shapley. In contrast, Chen and Sönmez (2006) and Calsamiglia et al. (2010) show in all their treatments the fraction of truth-telling in the Gale-Shapley mechanism is greater than that in the Boston mechanism. Because the designs of these experiments are so different and the number of subjects in a typical session varies so widely, it is hard to compare these results or reach any definite conclusions.

that among subjects who chat, those chatting with others of different types appear to increase their payoffs between phases more than those chatting with others like them. Chat also increases the fraction of stable outcomes in the Gale-Shapley and the Boston 10 mechanisms as well as decreasing the use of irrational strategies.

Finally, while we find no relationship between a subject's cognitive level and her behavior in either Phase 1 or Phase 2, we do find that more risk averse subjects tend to report strategic preferences more often in Phase 1 and are less likely to change their submitted preference rankings in Phase 2.

This paper is the first to study the performance of school matching mechanisms (or, perhaps, any economic mechanism) in the presence of chatting via social networks. As stated above, we feel this is an important step toward reality. It is important to point out that the object of chat in our experiment is the strategies of subjects but not the quality of schools or objects. While in the world outside of the lab parents are very likely to talk about the quality of the schools which they might send their children to or the fit between the schools and their children, in our experiment, because subjects know the values of the objects to them precisely, there is no scope for such discussion. We have done this intentionally to focus the attention of subjects on strategic issues.

The remainder of the paper is organized as follows. In Section 1 we provide a quick overview of the Boston and the Gale-Shapley mechanisms and their theoretical properties. In Section 2 we show our experimental design, while in Section 3 we present our results by answering the questions stated above and several more. Finally, Section 4 offers some conclusions.

1 The School Choice Matching Problem

In this section we will closely follow Pathak (2011) in describing the school choice problem. A school choice model with I students and N schools consists of a triple (P, q, π) where

1. $P = (P_1; \cdots ; P_I)$ are the preferences of students,
2. $q = (q_1; \cdots ; q_n)$ are the school capacities, and
3. $\pi = (\pi_1; \cdots ; \pi_n)$ are the school priorities.

The preferences of the students express their rankings over schools, the school capacities state the number of seats in each school, while the school priorities express information about how applicants are ordered at schools. In this paper we investigate only the school choice problem in which students are free to choose their strategies while schools are constrained to accept students on the basis of their exogenously defined priorities. Hence we will be looking at the "one-sided" version of the matching problem.

The outcome of a school choice problem is a student assignment, or a matching μ :

$I \rightarrow S$, where $\mu(i)$ indicates the school assignment of student i. A matching is **Pareto efficient** if there is no way to improve the allocation of a student without making another student worse off. (Note that only the welfare of students is considered in this definition.) A matching is **stable** if there is no student-school pair $(i; s)$ such that (a) student i prefers school s to her assignment $\mu(i)$, and (b) there is another student j with lower priority than student i assigned to s under μ. Finally, a mechanism is **strategy-proof** if truth-telling is a dominant strategy for all students.

Our experiments are not aimed at investigating which mechanisms are "best" when judged by the criteria listed above. Such work has already been done by Chen and Sönmez (2006), Pais and Pintér (2008), Calsamiglia et al. (2010) and Featherstone and Niederle (2011). Instead, we are interested in the impact of advice and network communication on behavior and mechanism performance. As mentioned above, we are therefore more interested in mechanisms that are *not* strategy-proof like the Boston mechanism, since such mechanisms leave a large amount of room for strategic dissembling and hence are more likely to exhibit strategic diversity across subnetworks when chatting is allowed.

In our experiments we run matching mechanisms in an environment with incomplete information, where subjects know only their own preferences but not those of others or the distribution from which their preferences are drawn.

1.1 Boston Mechanism

The Boston mechanism works as follows:

Step 1) Only the first choices of the students are considered. For each school, consider the students who have listed it as their first choice, and assign the seats in the school to these students one at a time following the school's priority order until either there are no seats left or there are no students left who have listed it as their first choice.

In general, at

Step k) Consider the remaining students. Only the kth choices of these students are considered. For each school with still available seats, consider the students who have listed it as their kth choice, and assign the remaining seats to these students one at a time according to priority until either there are no seats left or there are no students left who have listed it as their kth choice.

This mechanism, while widely used, is not strategy-proof.

1.2 Gale-Shapley Deferred Acceptance Mechanism

The Gale-Shapley deferred acceptance mechanism works as follows:

Step 1) Each student proposes to her first choice. Each school tentatively assigns its seats to its proposers one at a time following its priority order until either there are no seats left or there are no students left who have listed it as their first choice.

In general, at

Step k) Each student who was rejected in the previous step proposes to her next choice. Each school considers the students who has been held together with its new proposers, and tentatively assigns its seats to these students one at a time according to priority, until either there are no students left who have proposed or all seats are exhausted. In the latter case, any remaining proposers beyond the capacity are rejected. The algorithm terminates either when there are no new proposals, or when all rejected students have exhausted their preference lists.

The Gale-Shapley mechanism is strategy-proof regardless of information structures. Though it generally does not guarantee Pareto-optimal results, it does determine student-optimal stable matches when students and schools have strict preferences (see Dubins and Freedman (1981), and Roth (1982)), *i.e.*, no other stable assignment Pareto dominates the outcome produced by the Gale-Shapley mechanism, although this outcome might be Pareto dominated by some unstable assignments. However, when schools have coarse priorities, in the sense of not being able to express a strict order over students, the welfare consequences change as the Gale-Shapley mechanism may not always produce student-optimal stable matches. In our experimental design we consider coarse priorities.

2 Experimental Design

All our experiments were conducted in the experimental laboratory of the Center for Experimental Social Science at New York University. Five hundred and ten students were recruited from the general undergraduate population of the university using the CESS recruitment software. The experiment was programmed using the z-tree programming software (Fischbacher (2007)). The typical experiment lasted about an hour with average earnings of \$23.52. Subjects were paid in an experimental currency called Experimental Currency Units (ECU's) and these units were converted into U.S. dollars at a rate specified in the instructions. To standardize the presentation of the instructions, instead of reading the instructions, after the students had looked the instructions over, we showed a pre-recorded video which read them out load and simultaneously projected the written text on a screen in front of the room. The video for one treatment can be downloaded at https://files.nyu.edu/td648/public/SchoolChoice/, while the printed instructions are available in Appendix A.

2.1 Structure of Experiment

In the experiment subjects participated in three distinct decision tasks, and the monetary payoffs they received were the sum of their payoffs in each task. The first task was the matching experiment to be described below, while the second task was one play of the $2/3^{rd}$'s Beauty Contest and the third was the Holt-Laury risk aversion task. We used the beauty contest game as a diagnostic tool to evaluate their strategic sophistication, and we had subjects perform the Holt-Laury task for the obvious reason of eliciting their attitudes of risk aversion.

2.2 The Matching Problem

Our experiments are designed with an eye toward integrating the school choice mechanisms with social networks. To evaluate the impact of chat we ran a static school choice matching experiment twice, once before (Phase 1) and once after (Phase 2) allowing chatting via networks. Subjects were paid for either Phase 1 or Phase 2, but not for both. At the end of the experiment, the computer randomly determined which payoff, Phase 1 or Phase 2, would be paid. Subjects did not receive any feedback about their decisions or the decisions of other participants until the very end of the experiment, so the results of Phase 1 were not known until after the experiment. We used neutral language so schools were called "objects" and students were designated "subjects."

2.2.1 Phase 1

In both phases of the experiment there are 20 subjects. At the beginning of Phase 1 we randomly assign subjects to types in the sense that among these 20 subjects four are designated Type 1, four Type 2, four Type 3, four Type 4 and four Type 5. In addition to these 20 subjects there are also a set of 20 objects grouped into three types which are called Object A, Object B, and Object C. In total there are 8 units of Object A, 8 units of Object B and 4 units of Object C.[7]

Table 1 presents the full preference matrix of our subject types. These preferences are the same for all our experimental sessions, although the cardinal utilities associated with these objects vary across treatments.

In terms of payoffs, in six of our nine treatments that use either what we call the Boston 16 or the Gale-Shapley Mechanism, if a subject is matched to her first-best she will receive 24 ECU, if she is matched to her second-best she will receive 16 ECU, and

[7]The numbers of subjects in the isolated network treatment, defined in Section 2.2.3, vary from 10 to 20. In each session the number is a multiple of 5, and then the number of subjects for each type and objects change accordingly. For example, when there are 15 subjects in one session, we have 3 subjects for each type, 6 units of Object A, 6 units of Object B and 3 units of Object C.

Table 1: Preferences over Schools

	Student Preferences				
Type	1	2	3	4	5
1^{st} choice	C	C	C	A	A
2^{nd}choice	A	A	B	B	C
3^{rd}choice	B	B	A	C	B

if she is matched to her third-best she will receive only 4 ECU. In the remaining three treatments using what we call the Boston 10 Mechanism, we lower the value of the second-best object to each subject from 16 ECU to 10 ECU (see Section 2.2.3 for details).

In all treatments, it is common knowledge that Types 1 and 2 are given priority for Object A, while Type 3 is given priority for Object B and Types 4 and 5 are not given priority for any objects. When the number of subjects of equal priority applying for an object is greater than the number of objects available, the algorithm employs a lottery to break ties.[8] Note that Types 1 and 2 are identical with respect to both their preferences and priorities.

Subjects are told their own types and matching payoffs for each object as well as a priority table, but they do not know the types or the object payoffs of any subject other than themselves. They are then required to state their rankings over objects. Based on the information subjects provide, one of the matching algorithms (either the Boston or the Gale-Shapley mechanism) determines the allocation outcome. Each subject is matched to one and only one object.

2.2.2 Phase 2

To measure the impact of chat on matching, we run Phase 2 which allows subjects to submit their preference rankings after they communicate with each other via chat boxes. By comparing the submitted preference rankings of subjects in Phase 2 to those of the same subjects in Phase 1, we are able to observe the impact of chatting on subject behavior and welfare.

At Phase 2, subjects face the same matching problem as at Phase 1. Their types and matching payoffs for each object do not change. However, at Phase 2, before subjects enter their rankings, they are assigned to some subnetwork and allowed to talk with other subjects in their subnetwork for five minutes via chat boxes. The size of subnetworks range from 1 to 5 subjects. If a subject is assigned to a subnetwork with only one subject, she can not talk to anyone but is asked to enter into the chat box what factors influence her decision in Phase 2. In other words, we ask her to write in the chat box what she thinks

[8]In our experiment we used a single lottery instead of object specific lotteries to break ties.

of as she contemplates her ranking.[9] All communication goes via chat boxes. After five minutes all chat boxes become inactive and subjects have to submit their rankings again, i.e., they have to enter into the computers which objects they rank first, second and third, just as they did in Phase 1. The exact types of subnetworks used will be described in Section 2.2.3.

Payoffs in Phase 2 are determined in the same way as they were in Phase 1. At the end of the experiment, the computer randomly determines which payoff, Phase 1 or Phase 2, will be paid to subjects.

Several points about our design are worth noticing. First, the preferences and priorities are chosen to maximize competition among subjects, and hence subjects are likely to think strategically and exchange their thoughts with others when chatting. For example, in our design no subject has priority over her first-best object. In fact those with priorities always have priority for their second-best objects. Hence when a subject with priority states her second-best object first, it is a strategic move. In addition, note there are only 4 units of Object C available but all subjects with priorities, 12 subjects in total, have it as their first-best. Therefore Object C is in short supply and strategizing is more effective. Furthermore, for subjects without priority rights, their first-best object is Object A, which both Types 1 and 2 have priority over. Hence, their ability to obtain their first-best is determined not only by their own behavior but also by the behavior of Types 1 and 2. Without having a priority school to use as a target, Types 4 and 5 are more uncertain about their "best" strategies and therefore might be more likely to be influenced by chat.

Second, our matching game has comparatively large groups of subjects in order to mimic the real world where participants have no ability to coordinate. Since our experiment allows subjects to chat, we were afraid that, if we used only a small number of subjects in each group and allowed them all to chat with each other, they might collude. To avoid such collusion we recruited 20 subjects in each session and allowed them to chat via subnetworks of at most five subjects. This made it impossible (or at least extremely hard) for subjects to coordinate.[10]

Third, our design has an incomplete information structure to make the experimental environment closer to the real world. Note, however, as subjects know neither the preferences of others nor the distribution from which their preferences are drawn, when the Boston mechanism is used, there is no way to calculate the equilibrium. It is not a problem for us, because we are primarily interested in the changes of submitted preference rankings with the impact of chat and these changes can be easily measured without the

[9] See Cooper and Kagel (2012) for a similar "self advice" feature.

[10] As there is no collusion concern in the isolated treatment, sometimes we have less subjects per session.

knowledge of equilibrium.

Finally, some readers have suggested that it might make sense to tell subjects using the Gale-Shapley mechanism that it is strategy-proof and that they should therefore simply submit their truthful preferences. While this is certainly an option for school boards which they avail themselves of, the fact of the matter is that such advice is rarely distributed broadly and many times not believed. In addition, the credibility of the school board is many times suspected and the advice they offer is substandard. For example, when the Boston mechanism was first proposed parents were many times advised that it was strategy-proof despite the fact that it is not. In addition, less sophisticated parents or parents from under-privileged backgrounds are very likely not to read any literature about the matching program and hence unlikely to get the message, while those who are sophisticated tend to doubt the public advice given and think that they can still strategize. For these reasons we decided not to add a recommendation to our subjects that they should truthfully reveal their preferences when using the Gale-Shapley mechanism.

2.2.3 Treatments

Using the experimental procedures described above, we run a set of 9 treatments. These treatments differ by changing the matching mechanisms used while holding preferences and priorities constant, or changing the preference intensities while holding the mechanisms and priorities constant, or changing the network structures while holding all else constant. Below we describe these treatments one by one.

Preferences and Matching Mechanisms In terms of preference intensities we have two treatments, the Boston 16 and the Boston 10 treatments, which differ by the payoff received when subjects are allocated to their second-best objects. In the Boston 10 treatment, subjects are paid 24 ECU if they receive their first-best objects, 10 ECU if they receive their second-best objects, and 4 ECU if they receive their third-best objects, while for the Boston 16 treatment the payoffs are 24, 16 and 4 ECU's, respectively. (In the Gale-Shapley mechanism, the values are also 24, 16 and 4 ECU's, respectively.) The decrease for the second-best object from 16 ECU to 10 ECU was instituted to make subjects using the Boston 10 mechanism have more at stake in getting their first-best objects, since their second-best objects are less desirable. We believed this would increase competition and thereby increase the incidence of strategizing.

Because, in the Gale-Shapley mechanism, subjects have a dominant strategy while the equilibrium of the Boston mechanism is indeterminate, we suspected that the change of preference intensities would have a greater impact on the Boston mechanism. Therefore, we only reduced the value of the second-best object to 10 ECU in the Boston mechanism, and have the Boston 16 mechanism and the Gale-Shapley mechanism (with a value of 16 ECU for the second-best object) as our dual baselines.

Networks In investigating whether chatting impacts the behavior of subjects, it is important to ask with whom a subject chats. Does she talk to someone who has similar preferences and priorities, or someone quite different? While the subjects in our experiment, when being placed into subnetworks, do not know the types of others with whom they are networked, this information could be revealed during the chat (and it indeed is). We designed three different types of networks which we call Networks 1, 2, and 3, and these three networks are illustrated in Figure 1.

Figure 1 here.

Both Networks 1 and 2 are composed of four distinct 5-person subnetworks, some of which are complete and some incomplete. More precisely, the first subnetwork on the left in these two treatments is complete in the sense that each subject can send and receive messages from all the others, while the other subnetworks are incomplete. The subnetwork on the far right is "isolated" because no subject can communicate with any other. When chatting is allowed, these subjects can only talk to themselves. Note that subjects in the complete subnetwork of Network 2 have the same preferences and priorities (Types 1 and 2) while those in the incomplete subnetworks are all of different types, while in Network 1 the opposite is true. In Network 3, all subjects are isolated. Our design is summarized in Table 2.

Table 2: Experimental Design

Treatment	Pref. Intensity	Network	Mechanism	# of Sessions	# of Subjects
1	24-16-4	1	Boston	4	80
2	24-16-4	2	Boston	4	80
3	24-16-4	3	Boston	2	40
4	24-16-4	1	Gale-Shapley	4	80
5	24-16-4	2	Gale-Shapley	4	80
6	24-16-4	3	Gale-Shapley	3	45
7	24-10-4	1	Boston	2	40
8	24-10-4	2	Boston	2	40
9	24-10-4	3	Boston	2	25
				Total: 27	Total: 510

3 Results

We will present our results by stating a set of questions that we will attempt to answer using the data generated by our experiment. By comparing the behavior of subjects before and after chat, we are able to investigate whether chatting affects the strategies that subjects use and their matching outcomes. We first consider the baseline comparison between the Boston 16 and the Gale-Shapley mechanisms. We then discuss the Boston 10 mechanism separately in order to investigate the impact of changing preference intensity. Our discussion starts from Phase 1 behavior and then proceeds to Phase 2.

3.1 Phase 1 Results

If the behavior of our subjects and the performance of our dual baseline mechanisms were identical in Phase 1 then one might be tempted to conclude that these mechanisms were equivalent and hence interchangeable. It is our point here, however, that we must test the performance of mechanisms in the context in which they are used and since, in the real world, chatting via networks is the norm, such conclusions could be misleading if chat changed behavior. To properly examine the impact of chatting on mechanism performance and subject behavior, we will first examine how our dual baseline mechanisms performed in Phase 1 where there was no chatting to see if we conclude that they were equivalent. We will then proceed to Phase 2 and investigate the impact of chatting on changes in behavior and welfare in order to make what we consider to be the most telling comparisons.

To investigate Phase 1 we ask the following question:

Question 1 – Baseline Mechanism Equivalence:

Is there a difference, in aggregate or subject type by subject type, in the stated preference rankings of subjects in our dual baseline (the Boston 16 and the Gale-Shapley) mechanisms in Phase 1? In addition, given the stated preferences of subjects in Phase 1, is there a difference between the welfare of subjects across these two mechanisms?

Table 3 presents the percentages of stated preference rankings in Phase 1 aggregated over all subjects and type by type. It is worth pointing out that the definitions we use to classify strategies in Table 3 depart slightly from the conventional ones. In our analysis we classify a subject's strategy on the basis of what she places first in her submitted ranking. For example, we classify a strategy as "truthful" if a subject submits 1-2-3, or 1-3-2, "strategic" if she submits 2-1-3 or 2-3-1, and "irrational" if she submits 3-2-1 or 3-1-2. We use this classification for three reasons. First, the main goal of this paper is to investigate how subjects change their submitted preference rankings as a result of chat. If we use a very disaggregated way to categorize strategies, each subject would have 6 possible strategies and the transition matrix of strategy changes, across Phases 1 and 2, would contain 36 cells. Such transition matrices, given our data, would contain many cells with few observations which would make statistical comparisons difficult. In contrast, under our definitions the transition matrix is 3×3 and is much easier to handle. Such considerations are not present in previous experimental matching studies since those papers do not study the changes in submitted preference rankings and hence do not need to face the type of problems as we do. Second, we believe that this aggregation has a cognitive justification since we consider what a subject states as her first choice (among three objects) as more indicative of her strategy than how she ranks other objects. A subject who ranks her second-best object first is clearly strategizing, while a strategy which places the third-best first is irrational since there is no scenario, in either mechanism, where that strategy can be beneficial. If a subject had to rank more than three objects it might be myopic to look only at her first choice, but, for a three-object world, little is lost since there are not many manipulations one can do after the first choice is fixed. Finally, there are relatively few subjects using strategies such as 1-3-2 and 2-3-1.[11]

Table 3 here.

To ensure our strategy classification does not create a bias in our results, we also do our analysis using the more conventional definition with six strategies in which the truthful strategy is defined as only submitting preference 1-2-3 and the strategic as only submitting 2-1-3, and then rerun our three main regressions concerning strategy changes across Phases 1 and 2. These new regressions are presented in Appendix D where we find no statistical differences between the results using either classification.

Returning to Table 3, we see that there is little difference in submitted preferences between the Boston 16 and the Gale-Shapley mechanisms. While there are 37.00%, 58.00% and 5.00% of subjects submitting truthful, strategic, and irrational preference rankings respectively in the Boston 16 mechanism, in the Gale-Shapley mechanism these percentages are 39.51%, 53.66% and 6.38%, respectively. These percentages are not significantly different (using a set of t-tests) either when we pool subjects across types or when we compare the strategies chosen subject type by subject type except for the use of the irrational strategy by Type 1 subjects.[12]

To compare the welfare of subjects across our dual baseline treatments, we can not simply compare the outcomes determined at the end of the experiment since those outcomes are many times the result of the lottery used to break ties. If the lottery had turned

[11] In our data only a small fraction of subjects use strategies 1-3-2 and 2-3-1. Among all treatments in Phase 1 there are 2.75% of subjects playing 1-3-2 while 8.63% playing 2-3-1. In Phase 2 the respective fractions are 2.16% and 10.00%. The fractions of subjects using irrational strategies are also small, 6.67% in Phase 1 and 4.17% in Phase 2. Therefore, if we use the disaggregative way to categorize strategies, some cells of the transition matrices would contain few observations.

[12] The high frequency use of irrational strategy in the Gale-Shapley, though, is hard to explain given truth-telling is the dominant strategy.

out differently, we would have had a different outcome. To avoid this bias we simulate the expected matches of our subjects by randomly drawing a large number (2,500) of lottery orders while holding the stated preference rankings of subjects constant at their submitted preferences. We will refer to the results of this simulation often especially when we consider our results on welfare.

Table 4 presents the average percentage of times in our simulation that subjects receive their first, second, and third-best objects given their submitted preferences and the randomly drawn lottery orders. Our results show there are no statistical differences in the expected matching outcomes across the dual baseline mechanisms either in aggregate or type by type. In each simulation iteration we run a χ^2 test with the null hypothesis that the outcomes across the two mechanisms are identical. In the last column of Table 4 we report the fraction of times when the null hypothesis is not rejected.

Table 4 here.

With these simulated outcomes, we also calculate the fraction of the first-best surplus that is captured by our subjects. Our results show in the Boston 16 mechanism the subjects are able to capture 84.77% of the potential payoffs available to them, while in the Gale-Shapley mechanism they capture 85.62%. These fractions are not statistically different ($p = 0.6898$).

As stated previously, in our experimental design we have a Boston 10 treatment where the second-best objects of subjects decrease in value from 16 ECU to 10 ECU. By increasing the opportunity cost of losing one's first-best object, we expect that competition for the first-best objects will increase and subjects will behave differently than they do in our dual baseline mechanisms. These considerations yield Question 2.

Question 2 - Preference Intensity and Behavior:

Compared to subjects using the Boston 16 and the Gale-Shapley mechanisms, do subjects using the Boston 10 mechanism submit preferences that differ, in Phase 1, from those submitted in our dual baseline experiments?

The short answer here is yes. Subject behavior in Phase 1 of our Boston 10 mechanism differs from that of both the Boston 16 and Gale-Shapley mechanisms. This can be seen most easily in Figure 2 where we present the use of truth-telling, strategic and irrational strategies across our three mechanisms in Phase 1.

Figure 2 here.

Looking at Figure 2 , notice that while there is little difference between the submitted rankings across the Boston 16 and Gale-Shapley mechanisms, there is a marked increase in the use of truthful strategies in the Boston 10 mechanism as well as an increase in the use of irrational strategies.

Table 3 presents a more disaggregated view of the data. When we compare the stated preference rankings of subjects in the Boston 10 mechanism with those in our dual baseline mechanisms, we find that for each type except Type 5, there is an increase (and in many cases a substantial increase) in the submission of truthful preferences when we compare the Boston 10 to either the Boston 16 or the Gale-Shapley mechanism.[13] This increase is significant for all subjects who have priority rights (Types 1, 2 and 3) and for the pooled set of subjects. Meanwhile we find a decrease in the use of the strategic strategy for all types when comparing the Boston 10 to both the Boston 16 and the Gale-Shapley mechanisms, and the decrease is significant not only for the subjects who have priority rights but also those who do not. Finally, there is no statistical difference in the fraction of subjects using the irrational strategy in the Boston 10 and the Boston 16 mechanisms, while there is when we compare the Boston 10 and the Gale-Shapley mechanisms since fewer subjects without priority use the irrational strategy while more subjects with priority do in the Gale-Shapley mechanism.

Given our answer to Question 2, one question that arises is whether the behavioral differences observed between the Boston 10 and our dual baseline mechanisms translate into differences in achieved welfare. This raises Question 3.

Question 3 – Welfare and Preference Intensity:

Compared to subjects in the Boston 16 and the Gale-Shapley mechanisms, are subjects in the Boston 10 mechanism more or less likely to receive their first, second or third best objects?

As we do in answering Question 1, we draw 2,500 lottery orders and compute the expected matches given the preference rankings submitted in the Boston 10 mechanism. The simulation results are reported in Table 4 along the expected allocation outcomes in our dual baseline mechanisms. In aggregate, subjects using the Boston 10 mechanism are more likely to receive their first-best, less likely to receive their second-best and more likely to receive their third-best. Using a set of χ^2 tests, the null hypothesis that the allocations are identical is rejected in 86.40% of the 2,500 simulation runs when we compare across the Boston 10 and Boston 16 mechanisms, and in 93.88% of the simulation runs when we compare the Boston 10 and the Gale-Shapley mechanisms.

[13] A similar impact of preference intensity is found in Klijn et al. (2013).

3.2 The Impact of Chat: Phase 2–Phase 1 Differences

To investigate the impact of chat, we examine the behavior of our subjects and the performance of our mechanisms in Phase 2 and compare them to Phase 1. In particular, we check whether the changes in behavior and matching outcomes of the non-isolated subjects are different from those of the isolated subjects. In addition, we are interested in how chat influences the changes in submitted preference rankings when we condition on whether subjects have priority or not for any object, on the preferences they state in Phase 1, and finally on whom they chat with, i.e., whether they are connected to a heterogeneous or homogeneous subnetwork.

Note that our focus is on the changes in submitted preferences and outcomes across Phases 1 and 2 and between subjects who chat and who do not chat, but not on the aggregate distribution of preferences or outcomes in these two phases. Comparing the distribution of submitted preferences across Phases 1 and 2, but not their change, may mask what is really going on since preference changes may cancel each other out and leave the impression that chatting has no impact. For example, say that in Phase 1 half the subjects submitted truthful preference rankings and half submitted strategic rankings. After chat, say all those who told the truth in Phase 1 submit strategic rankings in Phase 2, while all those who submitted strategic rankings in Phase 1 tell the truth in Phase 2. Furthermore, say for the subjects who are isolated, half of them chose to tell the truth and the other half entered strategic preferences in Phase 1, and no one changes in Phase 2. Then, while the Phase 1 and Phase 2 distributions would remain 50-50 for both the isolated and the non-isolated subjects, we could not claim that chatting has no impact, since it leads all non-isolated subjects to change but simply in a counter balancing direction. The lesson here is that one has to check the changes in submitted rankings and the distribution of these changes but not the absolute distributions themselves.

3.2.1 Changes in Submitted Rankings

Our first question concerns the frequency of changes in submitted preference rankings across our two baseline mechanisms among subjects who chat (i.e., the non-isolated subjects) as compared to isolated subjects who do not. Later we will investigate the direction of those changes, and will condition on the subjects' priority rights, whether their subnetworks are homogeneous or heterogeneous, and what strategies they used in Phase 1. To start, we examine unconditional aggregated changes.

Question 4 – Chat and Strategy Changes:

When we compare those subjects who chat with those who do not, is there a difference in the fraction of subjects changing their preference rankings across our three *mechanisms? Also, within each mechanism, is there a difference in the fraction of subjects who change their submitted preferences between those who chat and those who do not?*

To quickly summarize the answer to Question 4 we can say that in the absence of chat (i.e., among the isolated subjects) there is no difference in the fraction of subjects changing their submitted preferences across the Boston 16, the Gale-Shapley and the Boston 10 mechanisms. Among those subjects who chat (the non-isolated subjects) the fraction of strategy changes is significantly greater in the Gale-Shapley and the Boston 10 mechanisms than in the Boston 16 mechanism. Except for the Boston 16 mechanism, the non-isolated subjects change their submitted preference rankings more often than those who are isolated.

More precisely, to answer Question 4 we examine the fractions of subjects who change their submitted rankings from Phase 1 to Phase 2 in our three mechanisms, and check whether chat leads non-isolated subjects to change their preference rankings differently from those who are isolated.[14]

The answer to these question can be found in Figure 3 and Table 5.

Figure 3 and Table 5 here.

As Table 5 and Figure 3 indicate, a substantial fraction of subjects change their preference rankings in Phase 2. In the Gale-Shapley mechanism while 40.83% of the non-isolated subjects change their submitted rankings, only 27.06% of the isolated subjects do so. For the Boston 16 mechanism the percentages are 25.83% and 30.00% respectively, and for the Boston 10 mechanism the fractions are 46.67% and 24.44%, respectively. A set of t-tests show that the non-isolated subjects change their strategies more often than those isolated in the Gale-Shapley and the Boston 10 mechanisms (the p-values are 0.0387 and 0.0172, respectively). However, in the Boston 16 mechanism, more isolated subjects change their strategies than do the non-isolated subjects, though the difference is not significant (the p-value is 0.5245). The result of the Boston 16 mechanism may not be surprising if chatting convinces subjects that they have made the right choices in Phase 1 and so they should not change. For example, in Phase 1 a non-isolated subject in the Boston 16 mechanism may have submitted strategic preferences and, during chatting, become convinced that she need not change. We will, in fact, see that this is indeed the case from our chat records in Section 3.5. A later investigation, conditional on the Phase-1 behavior of our subjects, will also confirm our conjecture.

[14] Here we aggregate over subjects who are isolated in Treatments 1 and 2 and those in Treatment 3 where all subjects are isolated.

It is also interesting to note that even when some subjects are not allowed to chat, they still change their submitted preference rankings in Phase 2. As shown in our chat records, these isolated subjects change their strategies possibly because they realize their mistakes in Phase 1 or they believe the subjects who chat will change their strategies in Phase 2 so they should also change as a response. When comparing across the mechanisms, we find that there is no statistical difference in the fraction of subjects changing their submitted preferences when chat is absent. In contrast, when subjects are allowed to chat, the fractions of subjects changing their strategies are significantly greater in the Gale-Shapley and the Boston 10 mechanisms when compared to that in the Boston 16 mechanism (the p-values are 0.0136 and 0.0075, respectively). This indicates that chat has a differential impact on behavior conditional on the mechanism used or the preference intensities within a mechanism.

To summarize our answer to Question 4 more efficiently we present three logit regressions in which we regress a binary {0,1} variable indicating whether a subject changes her strategy between Phase 1 and Phase 2 on a set of dummy variables indicating her type, whether she is allowed to chat, the mechanism used, and the interaction terms across these dummies. The coefficients measure the changes away from the default situation, which is an isolated subject of Type 5 in Specification 1 and an isolated subject of Type 5 using the Boston 16 mechanism in Specifications 2 and 3.

Table 6 here.

As we see in Specifications 1 and 2, where no interaction terms exist, the presence of chat increases the likelihood that subjects change their submitted preference rankings. In Specification 3, where chat is interacted with either the Gale-Shapley or the Boston 10 mechanism, we find highly significant coefficients indicating that when chat is allowed in these mechanisms there are significantly more subjects changing their preference rankings compared to the Boston 16 mechanism. Specification 3 also shows keeping subjects isolated fails to significantly alter the fraction of subjects who change their submitted preferences across the mechanisms, while allowing them to chat in the Boston 10 and the Gale-Shapley mechanisms leads to a significant increase in the likelihood that a subject will change her submitted preferences across phases.

To ensure our results are not sensitive to the way we classify strategies, we rerun the above three regressions using a disaggregated definition, which allows 6 possible strategies of preference rankings and 36 possible changes between Phase 1 and Phase 2. The results, presented in Table 6′ in Appendix D, are qualitatively identical.

The results just stated are unconditional results in the sense that they do not condition on whether a subject has priority rights or not or on what strategy they used in Phase 1.

Since some of our subjects have priority rights for objects (namely Types 1, 2 and 3) and some do not (Types 4 and 5), we would like to investigate whether chat has a differential impact on the fraction of subjects changing their preference rankings between Phase 1 and Phase 2 conditional on whether a subject has priority or not and also on the strategy they used in Phase 1. For brevity of presentation we will relegate the full answer to these questions to an Appendix.

To quickly summarize our answer to the first question, we can say that chatting has a significant impact on the changes in preference rankings across Phases 1 and 2 for all subjects (with or without priority) using the Boston 10 mechanism and for subjects without priority using the Gale-Shapley mechanism. However, it has no impact on strategy changes for any subjects who use the Boston 16 mechanism. (See Appendix B for a full discussion.)

With respect to the impact of chatting conditional on the strategy used in Phase 1, we find that given the submission of strategic preference rankings at Phase 1, the fraction of the non-isolated subjects switching to truth-telling in Phase 2 is significantly greater in the Gale-Shapley and the Boston 10 mechanisms than in the Boston 16 mechanism, while there is no significant difference across these three mechanisms for the isolated subjects. In addition, holding the mechanisms constant and comparing the isolated and the non-isolated subjects, we see that there is a significant difference in the fraction of subjects switching from strategic to truthful preference rankings for the Boston 10 and the Gale-Shapley mechanisms but no significant difference within the Boston 16 mechanism. Finally, only in the Boston 10 mechanism is there a significant impact of chat on the fraction of the subjects switching from truth-telling to strategic preference rankings. (See Appendix C for a full discussion).

One might think that allowing people to chat and exchange ideas might eliminate some of the irrationalities they exhibit in Phase 1. For example, if a subject submitted irrational preferences in Phase 1 and then chatted, is she more or less likely to repeat this mistake afterwards? On the other hand, the opposite may be true. Those subjects who submitted truthful or strategic preferences in Phase 1 may be more likely to switch and enter irrational preferences in Phase 2. (Remember, as we define it, an irrational strategy is dominated.) These issues lead to the following Question.

Question 5– Chat and Irrationality:

Does chat lead to an increase in rational behavior between Phases 1 and 2, i.e., does chatting route out irrationality? Are the preferences submitted by non-isolated subjects in Phase 2 best responses to the preferences submitted by their cohorts in Phase 1?

We can answer this question by looking at the transition matrices between Phase 1 and Phase 2 strategies across our mechanisms and across isolated and non-isolated subjects.

Such matrices are presented in Table 7. Note that if we compare horizontally across either the top, middle, or bottom panels we are holding the mechanisms used constant and comparing the transitions made by the non-isolated and the isolated subjects, while if we compare vertically along the panels of these tables, we are comparing the behavior of subjects with identical degrees of isolation but using different mechanisms. Before we answer our specific questions about the impact of chatting on rationality, we ask whether these transition matrices are different for subjects who chat as compared to those who do not.

Table 7 here.

Using a set of χ^2 tests we find that for isolated subjects there is no significant difference in the transition matrices across mechanisms for any binary comparison.[15] The situation is different when we consider non-isolated subjects, since there is a significant difference in the transition matrices of the non-isolated subjects when we compare the Boston 16 mechanism to either the Boston 10 or the Gale-Shapley mechanism. However, no significant difference exists between the Boston 10 and the Gale-Shapley mechanisms.[16] In other words, chat influences the way subjects change their submitted preference rankings across phases.

By examining within mechanisms and comparing the transition matrices of the isolated and the non-isolated subjects, we find that only in the Gale-Shapley mechanism is the transition matrix of isolated subjects significantly different from that of non-isolated subjects (the p-value is 0.0273).

To answer the question of whether chat routes out irrationality among our subjects, let us first compare the transition probabilities between subjects who were irrational in Phase 1 and remained so in Phase 2 (the bottom right entry along the diagonal in each transition matrix). If these probabilities are lower for subjects who chat than for those who do not, then we can say that chat routes out irrational behavior.

As we can see, this is true for the Boston 16 and Gale-Shapley mechanisms but not so for the Boston 10. More precisely, for subjects who used the Gale-Shapley mechanism only 11.11% of those who submitted irrational strategies in Phase 1 continued to do so in Phase 2 while 40.00% of those who did not chat did so. For the Boston 16 mechanism these percentages are 40.00% and 60.00%, meaning that amongst those who chatted and

[15]The p-values are 0.1971 for the comparison between the Boston 16 and the Gale-Shapley mechanisms, 0.5326 for the Boston 10 and the Gale-Shapley mechanisms, and 0.7278 for the Boston 16 and the Boston 10 mechanisms.

[16]For a comparison between the Boston 16 and the Gale-Shapley the p-value is 0.0771, between the Boston 16 and the Boston 10 the p-value is 0.0239. No significant difference between the Boston 10 and the Gale-Shapley mechanisms (the p-value is 0.1021).

submitted irrational strategies in Phase 1, 40.00% persisted in doing so in Phase 2 while for those who did not chat, this percentage was 60.00%. For the Boston10 mechanisms 40.00% of those who submitted irrational preferences in Phase 1 continued to do so in Phase 2 whether they chatted or not. It is never the case that subjects who used irrational strategies in Phase 1 persisted more in doing so after chatting than those who did not chat.

Next we ask whether there is relatively more inertia across the mechanisms in the sense of there being less of a tendency to change one's stated Phase 1 rankings. To answer this question we compare the diagonal of each transition matrix, as each cell on the diagonal indicates the fraction of subjects who do not change their submitted strategies across phases. We find that among the isolated subjects there is no difference in the fraction of subjects not changing their Phase 1 preference rankings across mechanisms, and this is true for any stated Phase 1 preference rankings (truthful, strategic or irrational) and any binary comparison of mechanisms. For the non-isolated subjects, however, it appears that among those initially submitting strategic preference rankings in the Boston 16 mechanism, the fraction refusing to change is significantly greater than those in the Boston 10 or the Gale-Shapley mechanisms (the p-values are 0.0240 and 0.0152, respectively). This result again confirms that chat convinces more subjects in the Boston 16 mechanism who enter strategic preference rankings in Phase 1 to stay with that strategy in Phase 2 than in the other two mechanisms.

Finally, there are some anomalies in the results that we need to mention. First, among those who chat and use the Gale-Shapley mechanism, there is a greater movement towards irrational strategies and away from truthful preferences across Phases 1 and 2. For example, in the Gale-Shapley mechanism around 9.09% of subjects who report truthful preferences in Phase 1 switch to submitting irrational strategies in Phase 2 after chatting while for those who do not chat no one switches from truth-telling to irrationality. This is obviously surprising since truth telling is a dominant strategy for the Gale-Shapley mechanism. In addition, subjects using the Boston 10 mechanism are more likely to transit from strategic to irrational preferences across Phases 1 and 2 after chatting than if they did not chat (8.70% versus 0.00%).

It is important to point out, however, that this anomaly is determined by the behavior of non-isolated subjects without priority. In the Gale-Shapley mechanism, for those subjects without priority, after chatting 21.05% of subjects submitting truthful preferences in Phase 1 switch to submitting irrational preference in Phase 2 while no subjects with priority make such a change. In the Boston 10 mechanism, after chatting 25.00% of subjects without priority transit from strategic to irrational preference rankings but none of subjects with priority do so. In fact, the movement toward irrational strategies is almost exclusively done by subjects without priority. More precisely, in no treatment was there ever a subject with priority who migrated from a truthful or strategic preference submission in Phase 1 to an irrational preference submission in Phase 2. The only subjects who ever made such a move were those without priority. The point seems clear, subjects without priority in our experiments seem more lost as to how they should behave and chatting does not seem to provide the clarity it does to subjects with priority.

Another version of Question 5 might be to ask whether the preferences submitted by a non-isolated subject in Phase 2 are best responses to the preferences of the other 19 subjects in Phase 1. In other words, do subjects submit a best response to the constellation of preferences they face in Phase 1?

A priori there is no reason to think that they should since subjects chat with at most four other subjects and hence can not get a good sense of the preferences and strategies they face in Phase 1. Still, despite this, chat may be best-response enhancing. Theoretically, subjects in the Gale-Shapley mechanism can always benefit by switching to truth-telling in Phase 2 if they have not done so in Phase 1 independent of what the other subjects do. Of those subjects who chatted when using the Gale-Shapley mechanisms and submitted either strategic or irrational strategies in Phase 1, 34.33% and 22.22% switched to truth-telling in Phase 2. For those who did not chat, the same percentages are 16.28% and 0.00%, respectively. The last percentage is interesting since it means that in the Gale-Shapley treatment while 22.22% of the subjects who chatted and entered irrational preferences in Phase 1 switched to truth-telling in Phase 2, no one who was isolated and used an irrational strategy in Phase 1 did so. This appears to be a dramatic difference. The best responses are not that straightforward for the Boston 16 and Boston 10 mechanisms. Given the preferences submitted in Phase 1, 50.00% of the subjects in the Boston 16 mechanism and 47.62% of the subjects in the Boston 10 mechanisms could have benefitted from switching to some other strategy. Of those who changed in the Boston 16 mechanism, only 30.19% of them chose their best responses while 25.64% of those who changed in the Boston 10 mechanisms did so. There is no statistical difference in the fraction of subjects choosing best responses in these Boston mechanisms for subjects who did and did not chat. Hence, for the Boston 16 and Boston 10 mechanisms, chatting had little impact on the propensity of subjects to best respond.

One of the main objectives of the school matching mechanisms is the determination of stable matches. Matching outcomes that leave people dissatisfied with their matches are bound to face political resistance. This raises the question of whether chatting is stability enhancing. This motivates our next question.

Question 6– Chat and Stability:

Does the percentage of stable matches increase from Phase 1 to Phase 2?

Given the preferences and priorities of our subjects, an outcome will be stable as long as no subjects of Types 1-3 are awarded their third-best objects. This is true because these subjects have priority for their second-best objects and if they were given their third-best they could kick out any subject of Types 4 or 5 who was awarded an object they had priority over. (There are enough Objects A and B to satisfy the demand from all subjects of Types 1 -3 who have priority over them.) Using this criterion to search for stable outcomes we find that in the Boston 16 mechanism 56.48% of the Phase 1 matches are stable while in Phase 2 56.99% are stable. The corresponding numbers for the Gale-Shapley mechanisms are 68.75% and 85.45% respectively (significantly different at p = 0.0000), while for the Boston 10 mechanisms they are 17.16% and 49.65% respectively (significantly different at p = 0.0000).

Several points are of note here. We find no increase in the stability of matches across phases for the Boston 16 mechanism probably because, as we know, there was no significant change in the submitted preferences of subjects across these phases. This was not true for the other mechanisms. Second, notice that for the Gale-Shapley and Boston 10 mechanisms the frequency of stable matches increases considerably across phases. What is interesting is how few stable matches were created in Phase 1 of the Boston 10 mechanism and how dramatically they have increased there. This is probably the result of the fact that subjects in Phase 1 of the Boston 10 mechanisms submitted a large number of truthful preferences and those submissions determined an increase in subjects receiving their third-best objects – exactly the circumstances where instability is likely to occur. Chat appears to have rectified this mistake to some extent.

In our experiment subjects who chat do so with either subjects who are of their own types or of different types, i.e., they are linked via either a homogenous or heterogeneous subnetwork. The natural question that arises is whether subjects make more changes in their submitted preference rankings when they talk to people who are like them or different from them. If so, do they change in a specific direction given whom they are chatting with?

Question 7 – Strategy Change and Network Type:

Do subjects change their submitted preference rankings more when they talk to people who are like them or different from them? If they do change more, do they change in a specific direction given whom they are chatting with?

As we will see below, talking to subjects via heterogeneous subnetworks appears not to increase the incidence of strategy changes above that of isolated subjects. However, subjects communicating in homogeneous subnetworks, tend to change their Phase 1 submitted preference rankings more than isolated subjects. This effect disappears, however, if a subject submits a truthful preference ranking in Phase 1.

Table 8 presents data to answer Question 7. It presents three logit specifications which use aggregate data pooled across all three mechanisms and differ according to their dependent variables. In Specification 4 the dependent variable is a binary $\{0, 1\}$ variable indicating whether a subject makes any change in submitted preference ranking between Phase 1 and Phase 2. Specification 5 examines the subset of subjects who submit strategic preferences in Phase 1 and the dependent variable is a dummy variable indicating whether a subject switches from a strategic to a truthful preference ranking, while in Specification 6 the dependent variable is the indicator of whether a subject switches from a truthful to a strategic preference ranking. In all three specifications, the independent variables are a subject's type and whether she communicates through a homogenous or heterogenous network.

Table 8 here.

As Table 8 shows, compared to the isolated subjects, the non-isolated subjects who communicate with subjects of their own types tend to change their submitted preference rankings in a statistically significant manner. Conditioning on the Phase 1 strategies used, the chat effect is not significant when subjects submit truthful preferences in Phase 1 although it remains significant if subjects initially submit strategic preference rankings.[17]

Finally, as we did above, we check whether the above results are sensitive to our classification of submitted preference rankings. The results are presented in Table 8′ in Appendix D, and we again find no qualitative differences in these results.

If the results about network types are robust and replicable, it would indicate that speaking to people like oneself has a bigger impact on one's willingness to change one's submitted preference ranking.

3.3 Risk Aversion, Cognitive Level, Submitted Preferences and Preference Change

At the end of our experiment we ran one round of the $2/3^{rd}$'s beauty contest game and the Holt-Laury risk aversion test. We did this to investigate how a subject's cognitive level, measured by her choice in the $2/3^{rd}$'s beauty contest and her level risk of aversion, affect the preferences submitted in Phase 1 as well as the likelihood of changing her preference submission in Phase 2. Since a subject's choice in the beauty contest is never significant in our specifications, we only present our results concerning risk aversion.

[17]Since these regressions use aggregate data they do not condition on the mechanism used. Our transition matrices in Table 7, however, suggest that some of our result may be driven by behavior in the Boston 10 mechanism where there is an excess of inertia for those who are isolated and more than average transitions for those who chat.

Question 8 – Risk and Cognitive Level:

Does a subject's level of risk aversion influence the likelihood that she submits truthful preferences in Phase 1 or the likelihood that she changes her submitted preference ranking across Phases 1 and 2?

Our data indicates that the more risk-averse a subject is the less likely she is to tell the truth in Phase 1 or to change her submitted preferences in Phase 2. Our results are presented in Table 9 where we first examine how the level of risk aversion affects the submitted preference rankings of subjects in Phase 1. In particular, we ask whether a higher level of risk aversion leads to less truth-telling in Phase 1. We run two logit regressions where the dependent variables are indicator variables recording whether a subject reports truthfully in Phase 1.

Table 9 here.

In Specification 7 the independent variables are a dummy variable indicating a subject's type and his switching point in the Holt-Laury task. The higher a subject's switching point is the higher is her level of risk aversion. The results show that the coefficient associated with the switching point is both negative and significant indicating that the more risk averse a subject is the less likely she is to report her preferences truthfully in Phase 1. This makes sense in that subjects may consider reporting truthfully to be risky.

In Specification 8, we include two more independent variables, namely the Gale-Shapley and the Boston 10 mechanisms. We then find that not only does the coefficient associated with the switching point continue to be significant and negative, but also that the coefficients in front of both mechanism indicators are positive and, in the case of the Boston 10 mechanism, significant. This result is consistent with our findings in section 3.1 which states that subjects in the Gale-Shapley and the Boston 10 mechanisms are more likely to submit truthful preference rankings.

We next investigate the effect of risk aversion on the likelihood of changing one's Phase 1 stated preference in Phase 2, conditional on what strategy a subject used in Phase 1. We again employ two logit regressions conditioning on the stated preference in Phase 1. In Specification 9 we examine the set of subjects who report the truth in Phase 1 where the dependent variable is a dummy indicating whether a subject switches to the strategic preference in Phase 2, while in Specification 10 the dependent variable is a dummy indicating whether a subject changes from the strategic to the truthful preference.

Table 10 here.

As shown in Table 10, one immediate result is that a subject's risk aversion has no significant impact on whether she changes her submitted preference ranking if she states her ranking strategically in Phase 1. Instead, it is chat that urges subjects to change from strategic to truthful preference submission, and this finding is consistent with our findings in Appendix C where we look at strategy change as a function of Phase 1 submitted preferences. In contrast, when we condition on the subjects who report truthfully in Phase 1, we find that risk aversion makes subjects less likely to switch to strategic preference rankings. This result seems surprising given our previous finding that more risk averse subjects are less likely to submit their preferences truthfully. A possible explanation is that risk averse subjects may consider a change, any change, as risky and hence resist doing so. (There may also be a factor of anticipated regret if the change turns out badly.) This is, of course, a conjecture since the Holt-Laury measure is intended to correlate with the curvature of the utility function and not other aspects of what we might consider to be risk averse behavior. Finally, note that the level of risk aversion is the only variable that is significant in Specification 9, and it indicates that a switch from truth-telling to strategic preferences is not influenced by the type of a subject, the mechanism used, or even whether she is allowed to chat.

3.4 Welfare Changes

A standard welfare measure for experimental markets (of which matching is one example) calculates the fraction of the available surplus (i.e., the first-best payoff sum) that is captured by subjects in an experimental session. In our experiment, however, we care about the impact of chat on welfare and hence would like to compare the welfare of subjects who chat to those who do not, as well as the change in welfare across phases. This is not easily accomplished using the surplus measure, since that measure is an aggregate market measure encompassing both the isolated and the non-isolated subjects and there is no way to impute the changes in aggregate market welfare across phases separately to each group. We can, however, compare how the payoffs of the isolated and the non-isolated subjects change across phases, which we will do below.

To do this, we employ two different payoff measures and run regressions that investigate the impact of chatting on these payoff measures. The first measure, Payoff Difference, is simply the mean payoff difference of subjects between Phase 1 and Phase 2, while the second, Payoff Difference Dummy, is a discrete measure which assigns a -1 for any subject whose payoff decreases from Phase 1 to Phase 2 , a 0 for any subject whose payoff stays the same, and a $+1$ for any subject whose payoff increases. These differences are calculated for each subject using the mean payoff determined by our simulation, which takes account of the possibility that some outcomes are determined randomly by lottery draws.

The impact of chat on welfare may be influenced by a number of factors, such as whom one chats with (heterogeneous vs. homogeneous subnetwork), whether the subjects chatting has a priority, and the type of matching mechanism used. Our next two questions concern the welfare impact of chatting under all of these circumstances.

Question 9 asks about the impact of the type of subnetwork a subject uses to communicate, while Question 10 examines both the type of mechanism used and whether the subjects have priority or not.

Question 9 – Heterogeneity, Chat, and Payoff Increases:

Does allowing subjects to chat increase their payoffs across phases and do payoffs increase more when subjects chat with people like themselves or unlike themselves, i.e., when they chat via heterogeneous or homogeneous networks.

What we will substantiate below is that allowing subjects to chat either increases their payoffs across Phases 1 and 2 or, as was true in the Boston 10 mechanism, when payoffs decrease, they decrease less among those who chat. In addition when we condition on whom the subjects chat with, we find that only when subjects chat with subjects unlike themselves, i.e., in heterogenous networks, do their payoffs increase.

To support this conclusion let us start by pointing out that across our three mechanisms since while the payoffs among those who chat increase by 0.7229 and 0.3128 in moving from Phase 1 to Phase 2 in the Boston 16 and the Gale-Shapley mechanisms, they decrease for those who do not (the payoff differences are -0.2097 and -0.6392, respectively). In the Boston 10 mechanism, payoffs decrease for both subjects who chat and who do not by -0.4974 and -0.7376, respectively. Note however, that while the payoffs for subjects who chat decrease in the Boston 10 mechanism, the decrease is greater for those who do not chat. These mean payoff differences across subjects who chat and those who do not are significant for subjects using the Boston 16 and the Gale-Shapley mechanism (p-values are 0.0159 and 0.0289,) but insignificant for the Boston 10 mechanism (p-value is 0.4140).

To investigate this result more formally, we use our two payoff measures (Payoff Difference and Payoff Difference Dummy) as dependent variables and employ our usual set of right hand dummy variables such as a subject's type, a dummy designating whether she is allowed to chat, and the mechanisms used. In Specifications 13 and 16, we substitute whether a subject is in a homogeneous or heterogeneous subnetwork in place of the chat variable. For the payoff difference regression we use an OLS regression while for the discrete payoff difference dummy variable we use an ordered logit regression.

Table 11 demonstrates several interesting facts. First, the chat variable is significantly positive in all regressions where it appears. That is, chat increases the average payoffs across Phase 1 and Phase 2, and also the likelihood that one's payoff will increase. Second, when we break the chat variable down by checking whom subjects chat with, it appears that only those who chat with people different from themselves increase their payoffs significantly. Recall, however, when answering Question 7 we find that subjects in the homogeneous subnetworks are more likely to change their Phase 1 strategies. Hence, while speaking to people of one's own type may lead a subject to change her preference ranking, such changes do not necessarily lead to payoff increases. This might imply that there are some circumstances where not changing one's submitted preferences is advantageous and listening to people of a similar type may be damaging.

Table 11 here.

While Question 9 concerns the impact of chat on welfare and, more specifically, the impact of whom one chats with, it is possible that chat affects welfare differently conditional on whether one has priority or not and the matching mechanism used. Question 10 asks about these relationships.

Question 10 - Welfare Changes, Mechanisms, Priority Rights and Chat: *Does chatting influence subject payoffs differently depending on whether they have priority rights or not?*

We answer Question 10 in two steps. First, using a simple OLS regression, we examine the payoff difference across Phases 1 and 2 and across our three mechanisms for subjects with and without priority rights. We do this by running three separate regressions, one for each mechanism, where the dependent variable is the average payoff differences of subjects across Phase 1 and Phase 2, and the independent variables are dummies indicating whether a subject has priority or not, as well as whether they chat or not. In these regressions the default is a subject who does not chat and has priority.

Table 12 here.

As Table 12 indicates, the impact of priority on the payoff differences varies across mechanisms. While the subjects without priority using the Boston 16 mechanism appear to experience a greater increase in their payoffs across Phases 1 and 2 than those with priority, in the Gale-Shapley and the Boson 10 mechanisms, the subjects without priority experience a decrease in their payoffs across phases compared to those with priority (the difference is only significant for the Boston10 mechanism). Note, however, in all mechanisms the subjects who chat do better than those who do not. Finally, subjects using the Gale-Shapley mechanism who are both isolated and have no priority rights experience a decrease in their payoffs across Phases 1 and 2. This result may be of interest for policy makers since if we consider subjects without priority (or with priority in sub-standard schools) to basically be those who inhabit lower income areas, it would appear that, if the city they lived in used the Gale-Shapley mechanism, they would benefit from being included in a better social network with others who could discuss the strategic elements of the mechanism they were using or a consumer education program that could advise them on what to do.

3.5 The Content of Chat

In this section of the paper we describe the chat that transpired between Phases 1 and 2. In addition, we examine how the content of chat varies as we change the mechanism used, or condition on whether the chat is between subjects with (or without) priority, or between subjects connected by a homogeneous or heterogenous subnetwork.

To do our analysis, we recorded all the messages exchanged among subjects during chat periods, and had two independent research assistants code the records. There were nine different categories that the chat text could be classified by. More precisely, a statement was coded with a T (Truth) if it mentioned revealing one's truthful preference and a $T-$ if it urged against doing so. Likewise, a message was coded as S (Strategic) if it encouraged listing one's second-best object first, and $S-$ if the statement urged against being strategic. An I (Irrational) statement suggested an irrational submission, such as placing one's third-best first, and an $I-$ statement warned against doing so. If subjects mentioned the lottery the statement was coded as L, and if they mentioned the scarcity of the objects in comparison to the possible demand the statement was coded as SC. Finally, a statement was coded with a C if subjects simply mentioned that chat might have an impact on the results. We allowed chat statements to be categorized by as many of the nine categories as was appropriate. When a subject repeated the same opinion several times, we counted every statement.

After our two research assistants coded each statement with as many labels as they saw fit, their codings were given to a third assistant, the judge, who compared them and, when they were in conflict, the judge decided which codings were best. There are surprisingly few conflicts between the coders.

Table 13 here.

Table 13 presents the percentage of each message type sent by those subjects who chat and those who are isolated across mechanisms. We find that in the Boston 16 mechanism, subjects discuss strategic preference submission almost twice as often as truthful preference rankings. For example, chat statements coded as S and $T-$ constitute 45.8% of the meaningful chat statements while those coded as T and $S-$ constitute only 26.0%. (Here we lump S and $T-$ and T and $S-$ together, since an $S-$ statement warning against being strategic is almost equivalent to suggesting that a subject report truthfully and the same is true for $T-$ and S.) For the isolated subjects these percentages are 41.6% and 22.0%, respectively. This may explain why subjects in the Boston 16 mechanism chose not to change their strategies especially when they listed their second-best objects first in Phase 1, since chat served to reinforce the efficacy of submitting strategically in Phase 1. In contrast, subjects who chat in the Gale-Shapley mechanism talk about truthful and strategic preference submission with almost equal frequency (31.8% and 35.9%, respectively), while, in the Boston 10 mechanism, these percentages are 29.6% and 26.0%, respectively.

Finally, subjects send more chat messages when using the Boston 10 than other mechanisms. This might be as expected since, in that mechanism, neither truth-telling nor submitting strategic preferences is obviously salient and, perhaps as a result, subjects tend to contemplate other courses of action.

What a subject chats about may, of course, depend on whether she has priority or not and with whom she chats. Looking at priority first, we find that for each of our three mechanisms there is a significant difference in the distribution of chat messages sent by subjects with and without priority. These differences vary across mechanisms as we see in Figures 4(a)-(c).[18]

Figure 4 here.

A few more aspects of these distributions are of interest. First, note that subjects appear more concerned about the possibility of facing the lottery when the Gale-Shapley and Boston 10 mechanisms are used than when the Boston 16 mechanism is used. This is particularly true in the Boston 10 mechanism, where about 21.3% (20.6%) of the messages of subjects who chatted concerned the lottery for subjects with (without) priority.[19] Also, when scarcity is mentioned it is more likely to be mentioned by subjects without priority, since they are the ones who cannot guarantee themselves an object by acting strategically and hence have to worry about being frozen out of a preferred item. The dramatic difference across priority and non-priority subjects in their reference to irrational strategies, i.e., submitting one's least preferred object first, is puzzling but consistently higher for subjects without priority than those with it. For example, while in the Boston

[18]The p-values associated with the χ^2 tests under the null hypothesis that there are no differences in the distribution of chat statements across our 9 categories for subjects with and without priority are 0.023, 0.000 and 0.000 for the Boston 16, Gale-Shapley, and Boston 10 mechanisms, respectively.

[19]Remember that there are 12 subjects who have object C as their first choice and only 4 such objects available. Hence, if many subjects were to report truthfully, there would be a great scarcity of object C and many subjects would face a lottery since no subjects have priority for that object. This finding may explain why the fraction of subjects telling the truth significantly decreases after chat in the Boston 10 mechanism.

16 mechanism 7.7% of the messages of non-priority subjects mentioned irrational preferences, only 3.1% of the messages of subject with priority did so. For the Gale-Shapley and the Boston 10 mechanisms, the percentages are 1.9% vs 14.6% and 4.0% vs 23.5%, respectively. It is hard to explain the use of irrational strategies and the differential reference to them by subjects with and without priority. One explanation is that since subjects without priority could not guarantee themselves their second-best objects by being strategic, they entertained exotic (and irrational) strategies.

We next examine the differences in chat between subjects who communicate through homogeneous, heterogeneous and isolated subnetworks. To start, we notice that subjects in the Boston 16 mechanism are more likely to talk about strategic preferences while those using the Gale-Shapley mechanism who chat with others like themselves, tend to talk more about submitting truthful preferences. These differences are accentuated when subjects talk to others like themselves via homogeneous subnetworks. For example, when subjects using the Gale-Shapley mechanism talk through homogeneous subnetworks they discuss truth-telling (strategic preferences) 34.2% (26.7%) of the time while, when they talk to others via a heterogeneous subnetwork, they discuss truth-telling (strategic preferences) 23.3% (37.9%) of the time. It is not clear, however, which type of chat is beneficial, i.e., are subjects better off talking to people like themselves or to others.

Finally, it is worthy of note that even when we pool all chat massages across all mechanisms and all types of subjects, we find that subjects who chat, whether in a homogeneous or heterogeneous subnetwork, chat in a different manner than those who are isolated ($p = 0.000$ for the comparison between both the heterogenous and homogeneous subnetworks and the isolated subnetworks).

4 Conclusion

In this paper we have investigated the impact of communication on the performance of two frequently used school matching mechanisms, the Boston and the Gale-Shapley mechanisms. We study chat since, in the real world, parents are embedded in a set of social networks through which they communicate and discuss their preference submission strategies. By allowing chat in the lab we feel we are testing these mechanisms in a context closer to the way they are used.

What we find is interesting and we hope informative. To begin, we find little difference in the behavior of subjects or the performance of our baseline Boston 16 and Gale-Shapley mechanisms when these mechanisms are run without chat (as shown in the behavior in Phase 1). However, when we lower the preference intensity of the second-best objects for our subjects from 16 ECU to 10 ECU in the Boston mechanism, we find that subjects tend to report their preferences more truthfully.

The impact of chatting, however, may be dramatic. First, chatting increases the likelihood that subjects will change their submitted preferences compared to those subjects who are isolated and do not chat. More interesting, however, is the fact that a subject is more likely to change her submitted preference if, during chat, she talks to subjects who are like her in type.[20]Second, chatting seems to increase the rationality of subjects by reducing the number of subjects who persist in using irrational strategies after chat. Third, chatting leads to an increase in the fraction of stable outcomes which is important from a policy point of view. With respect to welfare, chatting appears to influence welfare differently for subjects with and without priority rights. Though there is little welfare change among subjects with priority rights between phases, chatting does significantly change the welfare of those without priority rights. Furthermore, our results show that among subjects who chat, those chatting with others of a different type appear to increase their payoffs between phases more than those chatting with others like themselves. Combined with the result that subjects that chat with others like themselves tend to change their submitted preference rankings more between phases, this suggests that some of those preference changes are not beneficial. (Sometimes chat is beneficial if it persuades people not to change their already beneficial strategies.)

Finally, while we find no impact of a subject's cognitive level (as measured by her choice in a $2/3^{rd}$'s beauty contest game) on her behavior in either Phase 1 or Phase 2, we do find that more risk averse subjects tend to report strategic preferences more in Phase 1 and are less likely to change their submitted preferences in Phase 2.

In terms of policy, our results indicate that the advice received through chatting via social networks can influence a subject's behavior and also the performance of the mechanism used. It is of more interest that chat seems to hurt isolated subjects without priority in the Gale-Shapley mechanism in the sense that their average payoffs decrease. This is unfortunate since, as stated above, in the real world, subjects without priority or subjects who have priority but in substandard schools, may be students of low socioeconomic status, the very students these mechanisms are aimed at helping. Therefore, a better consumer education program may help parents navigate a complex matching mechanism and improve the overall efficiency.

[20]One caveat for this result is, however, if subjects submitted truthful preferences in Phase 1, then no matter to whom they chat, the fraction of subjects changing their preference rankings is not significantly different from those who do not chat. On the other hand, when using the Gale-Shapley mechanism those who submit strategic preferences in Phase 1 are more likely to switch to truth-telling in Phase 2 than those using the Boston 16 mechanism.

References

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2011). Resolving conflicting preferences in school choice: The boston mechanism reconsidered. *American Ecnomic Review 101(1)*, 399–410.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005a). The new york city high school match. *American Economic Review, Papers and Proceedings 95*, 364–367.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *American Economic Review 99(5)*, 195478.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005b). The boston public school match. *American Economic Review, Papers and Proceedings 95*, 368–371.

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

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

Calsamiglia, C., G. Haeringer, and F. Klijn (2010). Constrained school choice: An experimental study. *American Economic Review 100(4)*, 186074.

Chen, Y. and O. Kesten (2013). From boston to chinese parallel to deferred acceptance: Theory and experiments on a family of school choice mechanisms. *Working Paper*.

Chen, Y. and T. Sönmez (2006). School choice: An experimental studay. *Jouranl of Economic Theory 127(1)*, 202–231.

Cooper, D. C. and J. H. Kagel (2012). Failure to communicate: An experimental investigation of the effects of advice on strategic play. *Working Paper*.

Ding, T. and A. Schotter (2014). Learning versus inter-generational advice in school matching mechanisms: An experimental study. *Memo in Progress*.

Dubins, L. E. and D. A. Freedman (1981). Machiavelli and the gale-shapley algorithm. *American Mathematical Monthly 88*, 485–494.

Erdic, A. and H. Ergin (2008). What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review 98*, 669–689.

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

Featherstone, C. and M. Niederle (2011). School choice mechanisms under incomplete information: An experimental investigation. *Working Paper*.

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

Guillen, P. and A. Hing (2013). Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism. *Working Paper, University of Sydney*.

Haeringer, G. and F. Klijn (2009). Constrained school choice. *Journal of Economic Theory 144(5)*, 1921–47.

Herszenhorn, D. (2003). Revised admissions for high school. *New York Times*.

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

Iyengar, R. and A. Schotter (2008). Learning under supervision: An experimental study. *Experimental Economics 11(2)*, 154–173.

Kesten, O. and M. U. Ünver (2013). A theory of school choice lotteries: Why ties should not be broken randomly. *Working Paper*.

Klijn, F., J. Pais, and M. Vorsatz (2013). Preference intensities and risk aversion in school choice: A laboratory experiment. *Experimental Economics 16(1)*, 1–22.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review 85(5)*, 1313–1326.

Nyarko, Y., A. Schotter, and B. Sopher (2006). On the informational content of advice: A theoretical and experimental study. *Economic Theory 29(2)*, 433–452.

Pais, J. and A. Pintér (2008). School choice and information: An experimental study on matching nechanisms. *Games and Economic Behavior 64*, 303–328.

Pathak, P. A. (2011). The mechanism design approach to student assignment. *Annual Review of Economics 3(c)*.

Pathak, P. A. and J. Sethuraman (2011). Lotteries in student assignment: An equivalence result. *Theoretical Economics*, 1 – 17.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of Operation Research 7*, 617–628.

Schotter, A. (2003). Decision making with naive advice. *American Economic Review 93(2)*, 196–201.

Schotter, A. and B. Sopher (2003). Social learning and coordination conventions in intergenerational games: An experimental study. *Journal of Political Economy 111*, 498–529.

Schotter, A. and B. Sopher (2007). Advice and behavior in intergenerational ultimatum games: An experimental approach. *Game and Economic Behavior 58(2)*, 365–393.

Toch, T. and C. Aldeman (2009). Matchmaking: Enabling mandatory public school choice in new york and boston. *Education Sector*.

Table 3: Stated Preference in All Three Mechanisms: Phase 1 (Percent)

	Truthful				Strategic				Irrational			
	Boston 16	GS	Boston 10	p-value GS v.Boston 16 Boston 16 v.10 GS v.Boston 10	Boston 16	GS	Boston 10	p-value GS v.Boston 16 Boston 16 v.10 GS v.Boston 10	Boston 16	GS	Boston 10	p-value GS v.Boston 16 Boston 16 v.10 GS v.Boston 10
All	37.00	39.51	52.38	0.6040 0.0107 0.0324	58.00	53.66	38.10	0.3803 0.0009 0.0009	5.00	6.83	9.52	0.4361 0.1680 0.4260
Type 1	35.00	31.71	47.62	0.7570 0.3568 0.2416	65.00	58.54	52.38	0.5552 0.3568 0.6537	0.00	9.76	0.00	0.0440 N/A 0.0440
Type 2	40.00	48.78	61.90	0.4328 0.1097 0.3342	60.00	51.22	38.10	0.4328 0.1097 0.3342	0.00	0.00	0.00	N/A N/A N/A
Type 3	47.50	34.15	66.67	0.2268 0.1548 0.0160	52.50	60.98	33.33	0.4478 0.1548 0.0404	0.00	4.88	0.00	0.1598 N/A 0.1598
Type 4	27.50	48.78	57.14	0.0496 0.0305 0.5420	60.00	43.90	14.29	0.1509 0.0001 0.0100	12.50	7.32	28.57	0.4407 0.1687 0.0619
Type 5	35.00	34.15	28.57	0.9366 0.6143 0.6599	52.50	53.66	52.38	0.9181 0.9931 0.9260	12.50	12.20	19.05	0.9673 0.5273 0.5059
Priority	40.83	38.21	58.73	0.6775 0.0218 0.0083	59.17	56.91	41.27	0.7229 0.0218 0.0442	0.00	4.88	0.00	0.0137 N/A 0.0137
Non-Priority	31.25	41.46	42.86	0.1786 0.2169 0.8834	56.25	48.78	33.33	0.3442 0.0150 0.0975	12.50	9.76	23.81	0.5818 0.1425 0.0630

Table 4: Outcome Distributions in All Three Mechanisms: Phase 1 (Percent)

	Boston 16			Gale-Shapley			Boston 10			% H_0 not rejected by a χ^2 test
	First-best	Second-best	Third-best	First-best	Second-best	Third-best	First-best	Second-best	Third-best	Boston 16 vs.GS Boston 16 vs.10 GS vs. Boston 10
All	23.08	63.83	13.09	22.57	65.63	11.79	31.10	48.09	20.80	0.9928 0.1360 0.0612
Type 1	18.93	77.16	3.91	16.84	73.43	9.73	18.25	64.72	17.04	0.9452 0.6028 0.9624
Type 2	19.38	72.80	7.81	27.40	72.60	0.00	25.13	48.61	26.26	0.6348 0.3628 0.0032
Type 3	27.86	71.33	0.81	16.89	80.86	2.25	25.47	74.53	0.00	0.8432 0.9712 0.9140
Type 4	22.90	70.56	6.54	27.86	67.28	4.86	57.15	32.15	10.70	0.9904 0.0044 0.1140
Type 5	26.33	27.29	46.38	23.87	33.99	42.13	29.53	20.45	50.02	0.9732 0.9760 0.9132

Table 5: Fraction of Strategy Changes

	Boston 16	GS	Boston 10	Difference Boston 16 - GS Boston 16 - 10 Boston 10 - GS	p-value
All	0.2750	0.3512	0.3714	−0.0762	0.0985
	(0.4476)	(0.4785)	(0.4855)	−0.0964	0.0922
				−0.0202	0.7278
Isolated	0.3000	0.2706	0.2444	0.0294	0.6782
	(0.4611)	(0.4469)	(0.4346)	0.0556	0.5039
				−0.0262	0.7474
Non-isolated	0.2583	0.4083	0.4667	−0.1500	0.0136
	(0.4396)	(0.4936)	(0.5031)	−0.6694	0.0075
				0.2084	0.4620
Difference	0.0417	−0.1377	−0.2223		
(p-value)	(0.5245)	(0.0387)	(0.0172)		

Table 6: Chat and Strategy Change

Variables	Specification 1 Coef. (Std. Err)	Specification 2 Coef. (Std. Err)	Specification 3 Coef. (Std. Err.)
Type 1	−0.3413	−0.0358	−0.0351
	(0.3091)	(0.3104)	(0.3128)
Type 2	0.2258	0.2285	0.2300
	(0.3044)	(0.3056)	(0.3074)
Type 3	0.1809	0.1832	0.1843
	(0.3058)	(0.3067)	(0.3085)
Type 4	0.3575	0.3605	0.3652
	(0.3000)	(0.3012)	(0.3032)
Chat	0.4274**	0.4395**	−0.1689
		(0.2017)	(0.3244)
GS	–	0.3665*	–
		(0.2170)	
B10	–	0.4614*	–
		(0.2585)	
GS×Chat	–	–	0.6874**
			(0.2800)
B10×Chat	–	–	0.9261***
			(0.3333)
GS×Iso	–	–	−0.1435
			(0.3458)
B10×Iso	–	–	−0.2796
			(0.4248)
Constant	−1.1361***	−1.3937***	−1.0241***
	(0.2524)	(0.2863)	(0.3167)

$*$ = significant at 10%, $**$ = significant at 5%, $***$ =significant at 1%

Table 7: Transition Matrices

	Non-Isolated Subjects			Isolated Subjects			
	Boston 16						
	Truthful	Strategic	Irrational		Truthful	Strategic	Irrational
Truthful	0.6386	0.3404	0.0213	Truthful	0.5185	0.4074	0.0741
Strategic	0.1618	0.8382	0.0000	Strategic	0.1875	0.8125	0.0000
Irrational	0.4000	0.2000	0.4000	Irrational	0.2000	0.2000	0.6000
	Gale -Shapley						
	Truthful	Strategic	Irrational		Truthful	Strategic	Irrational
Truthful	0.5909	0.3182	0.0909	Truthful	0.6757	0.3243	0.0000
Strategic	0.3433	0.6567	0.0000	Strategic	0.1628	0.8140	0.0233
Irrational	0.2222	0.6667	0.1111	Irrational	0.0000	0.6000	0.4000
	Boston 10						
	Truthful	Strategic	Irrational		Truthful	Strategic	Irrational
Truthful	0.5313	0.4375	0.0313	Truthful	0.7391	0.2174	0.0435
Strategic	0.3478	0.5652	0.0870	Strategic	0.1176	0.8824	0.0000
Irrational	0.0000	0.6000	0.4000	Irrational	0.2000	0.4000	0.4000

Table 8: Strategy Change and Network Type

	Specification 4 Any Change	Specification 5 Strategic $\rightarrow$ Truthful	Specification 6 Truthful $\rightarrow$ Strategic
Variable	Coeff. (Std. Err)	Coeff. (Std. Err)	Coeff. (Std. Err.)
Type 1	-0.1468	0.7370	-0.3295
	(0.3204)	(0.5217)	(0.5278)
Type 2	0.2096	1.056	0.1939
	(0.3049)	(0.5221)	(0.4647)
Type 3	0.1646	1.034	0.0005
	(0.3060)	(0.5202)	(0.4742)
Type 4	0.3217	0.5322	-0.5049
	(0.3015)	(0.5606)	(0.5046)
Homo	0.6171^{***}	0.8796^{**}	0.4696
	(0.2410)	(0.3921)	(0.3670)
Hetero	0.2626	0.5407	0.0040
	(0.2342)	(0.3767)	(0.3735)
Constant	-1.109^{***}	-2.396^{***}	-0.6943^{*}
	(0.2529)	(0.4819)	(0.3953)

$*$ = significant at 10%, $**$ = significant at 5%, $***$ =significant at 1%

Table 9: Truth-telling at Phase 1 and Risk Aversion

	Specification 7	Specification 8
Variable	Coeff. (Std. Err)	Coeff. (Std. Err)
Type 1	0.1273	0.1299
	(0.2954)	(0.2973)
Type 2	0.6353**	0.6431**
	(0.2901)	(0.2921)
Type 3	0.5816**	0.5879**
	(0.2913)	(0.2933)
Type 4	0.3967	0.4011
	(0.2916)	(0.2938)
Switching Point	−0.0978**	−0.0948**
	(0.0471)	(0.0473)
GS	-	0.0939
		(0.2069)
B10	-	0.6214**
		(0.2470)
Constant	−0.0920	−0.2842
	0.3562	0.3803

∗ = significant at 10%, ∗∗= significant at 5%, ∗ ∗ ∗ =significant at 1%

Table 10: Strategy Change, Chat, Risk Aversion and Mechanism

	Specification 9 Truthful → Strategic	Specification10 Strategic → Truthful
Variable	Coeff. (Std. Err)	Coeff. (Std. Err)
Type 1	−0.2176	0.8798*
	(0.5245)	(0.5056)
Type 2	0.3356	1.0810**
	(0.4713)	(0.5254)
Type 3	0.1387	1.0421*
	(0.4795)	(0.5245)
Type 4	−0.4634	0.7241
	(0.5084)	(0.5652)
Switching Point	−0.1508**	−0.0903
	(0.0735)	(0.0798)
Chat	0.2705	0.6943**
	(0.3195)	(0.3279)
GS	−0.2448	0.6052*
	(0.3533)	(0.3333)
B10	−0.1216	0.5469
	(0.3829)	(0.4545)
Constant	0.2447	−2.2278***
	(0.6201)	(0.7178)

∗ = significant at 10%, ∗∗= significant at 5%, ∗ ∗ ∗ =significant at 1%

Table 11: Chatting and Payoff Differences

Variable	**Specification 11** Difference Coeff. (Std. Err)	**Specification 12** Difference Coeff. (Std. Err)	**Specification 13** Difference Coeff. (Std. Err)	**Specification 14** Difference Dummy Coeff. (Std. Err)	**Specification 15** Difference Dummy Coeff. (Std. Err)	**Specification 16** Difference Dummy Coeff. (Std. Err)
Type 1	−0.0508	−0.5048	−0.4541	−0.6955	−0.0688	0.0438
	(0.5861)	(0.5852)	(0.6047)	(0.2694)	(0.2969)	(0.2789)
Type 2	−0.9576	−0.9591	−0.9516	−0.4053	−0.4030	−0.3855
	(0.5824)	(0.5816)	(0.5825)	(0.2712)	(0.2707)	(0.2709)
Type 3	−0.7440	−0.7455	−0.7380	−0.1229	−0.1182	−0.1108
	(0.5824)	(0.5816)	(0.5825)	(0.2706)	(0.2707)	(0.2706)
Type 4	−2.002***	−2.000***	−1.983***	−0.3961	−0.4107	−0.3731
	(0.5812)	(0.5804)	(0.5829)	(0.2730)	(0.2732)	(0.2740)
Chat	0.7559**	0.7398**	-	0.4432***	0.4398***	–
	(0.3827)	(0.3822)		(0.1692)	(0.1694)	
GS	–	−0.4209	−0.4210	–	−0.3070*	−0.3059*
		(0.4119)	(0.4123)		(0.1830)	(0.1831)
B10	–	−0.9289*	−0.9291*	–	−0.1303	0.1310
		(0.4996)	(0.5000)		(0.2245)	(0.2249)
Homo	–	–	0.6489	–	–	0.2466
			(0.4681)			(0.2090)
Hetero	–	–	0.8168*	–	–	0.6023***
			(0.4454)			(0.1989)
Constant	0.3779	0.7478	0.7355	–	–	–
	(0.4686)	(0.5227)	(0.5244)			

$*$ = significant at 10%, $**$ = significant at 5%, $***$ = significant at 1%

Table 12: Mechanism, Priority Rights Chat, and Payoff Difference

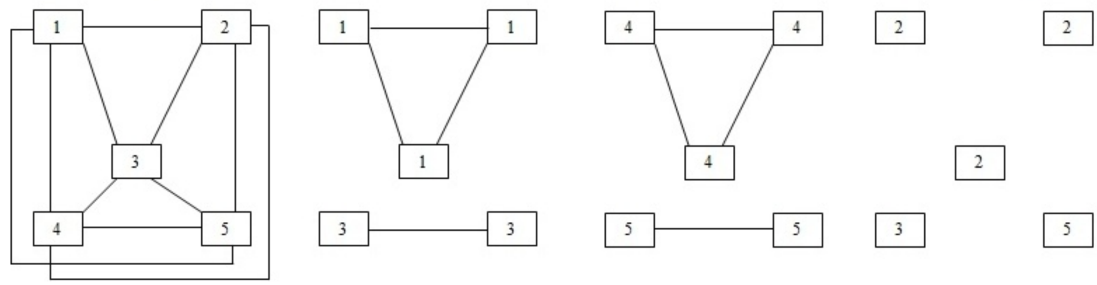
	Boston 16 Specification 17	Gale-Shapley Specification 18	Boston 10 Specification 19
Variable	Coeff. (Std. Err)	Coeff. (Std. Err)	Coeff. (Std. Err)
No Priority	1.3505**	−0.7741	−2.3413**
	(0.5645)	(0.4998)	(1.0957)
Chat	0.9325*	0.9520*	0.2402
	(0.5645)	(0.4970)	(1.0847)
Constant	−0.7499	−0.3300	0.1990
	(0.4921)	(0.4296)	(0.9298)

$*$ = significant at 10%, $**$ = significant at 5%, $***$ =significant at 1%

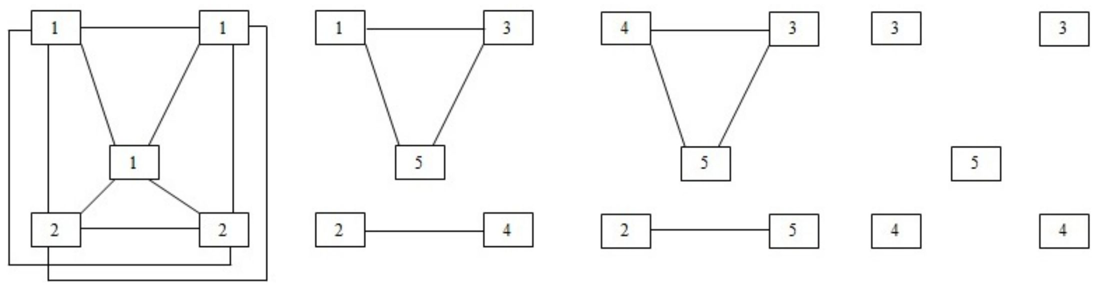
Table 13: Chat Records: Message Type

	Boston 16		**Gale-Shapley**		**Boston 10**	
	Chat	Isolated	Chat	Isolated	Chat	Isolated
T	0.2418	0.1883	0.2915	0.2184	0.2606	0.2091
T-	0.0293	0.0779	0.0404	0.0825	0.0493	0.0455
S	0.4286	0.3377	0.3184	0.2573	0.2113	0.1727
S-	0.0183	0.0325	0.0224	0.0388	0.0352	0.0545
I	0.0440	0.0584	0.0717	0.0631	0.1197	0.1182
I-	0.0110	0.0260	0.0224	0.0728	0.0141	0.0455
SC	0.0769	0.0909	0.0538	0.1359	0.0423	0.2000
L	0.0879	0.1753	0.1525	0.1311	0.2535	0.1545
C	0.0623	0.0130	0.0269	0.0000	0.0141	0.0000

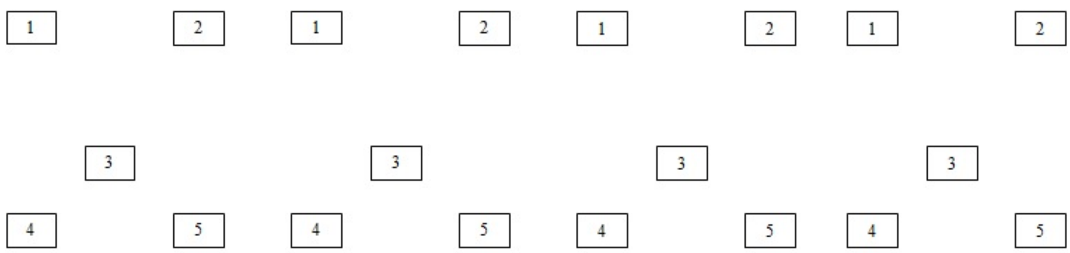
Figure 1: Network Structures



(a) Network 1: Heterogeneous Complete Subnetwork; Homogeneous Incomplete Subnetworks



(b) Network 2: Homogeneous Complete Subnetwork; Heterogeneous Incomplete Subnetworks



(c) Network Treatment 3: All Subjects Are Isolated

Figure 2: Phase 1 Submitted Preference Rankings

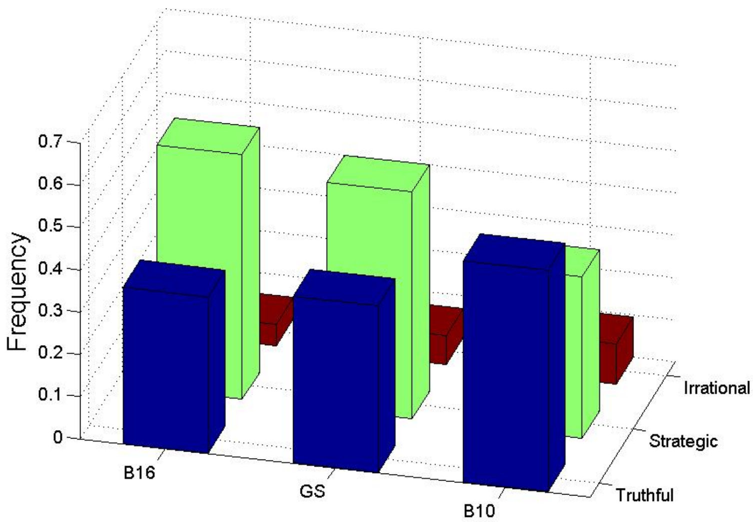


Figure 3: Strategy Changes: Isolated vs. Non-isolated by Mechanism

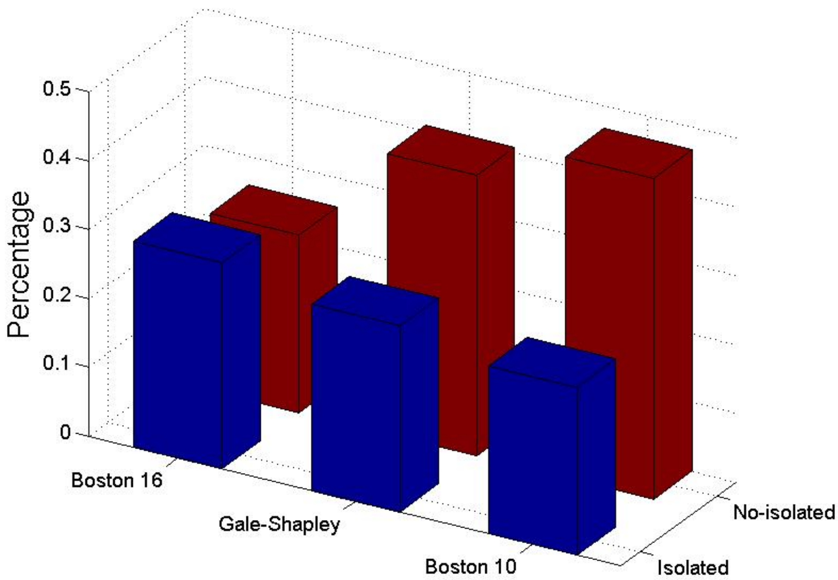
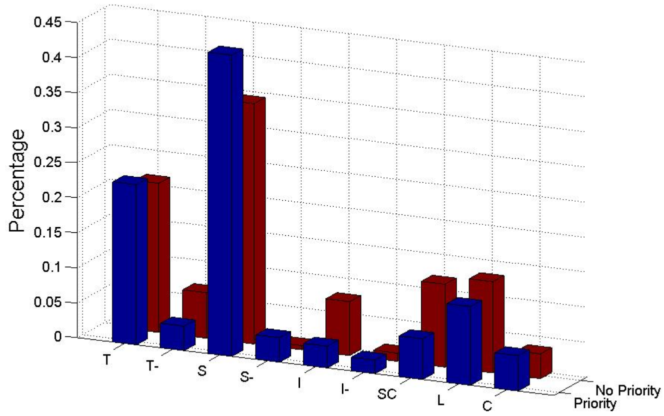
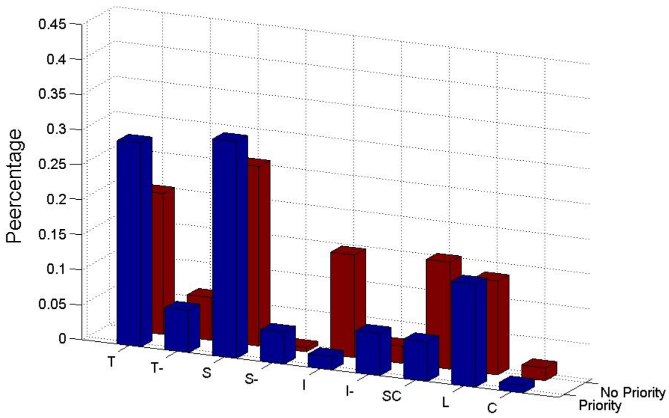


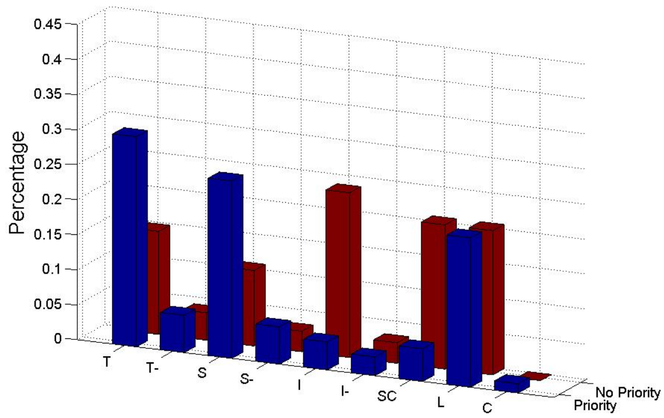
Figure 4: Distribution of Chat: Priority vs. Non-Priority



(a) Boston 16 Mechanism



(b) Gale-Shapley Mechanism



(c) Boston 10 Mechanism