

Copyright 2016 Thomas Sahajdack

THREE ESSAYS ON MATCHING OVER CHARACTERISTICS

BY

THOMAS SAHAJDACK

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Steven R. Williams, Chair
Professor Dan Bernhardt
Professor Nolan Miller
Professor George Deltas

Abstract

This thesis is divided into three chapters. In the first chapter, I study the use of an alternative method of eliciting preferences in a two-sided, one-to-one matching market. In the second chapter, I analyze this method of eliciting preferences in a two-sided, many-to-one matching setting with different primitives on preferences. Finally, in the third chapter, I use survey data to estimate the effects of information exposure on the reported preferences of the participants with the alternative method of eliciting preferences.

Many two-sided matching markets, in practice, are large, complex for the participants and suffer from incomplete information. These features provide challenges for matching theorists, and these challenges are not always well met with traditional matching mechanisms where agents directly submit a ranking of agents on the other side of the market. In Chapter 1, I study one approach to dealing with these challenges that, while used in some real world matching markets, has not received much attention in the literature. I analyze an alternative message space where agents submit some combination of their own characteristics and their preferences over characteristics of other participants instead of directly submitting preference rankings. By doing this, the market designer can elicit preferences without requiring agents to rank each other directly, which is often infeasible in real world applications. Using online dating markets as a motivating example, I study the incentive, mechanism and market design implications of using the alternative message space approach in one-to-one matching markets. I find that for a restricted class of preference rankings, using this approach can decrease the number of binary pieces of information, such as either/or questions, necessary to generate the preference rankings of agents. I quantify a lower bound on the number of such questions that allows for any agent to have his or her true preference ranking generated based on his or her answers. I show that a strategy-proof matching algorithm under direct reporting of preference rankings can extend strategy-proofness into the alternative message space if that message space only asks for either an agent's own characteristics or their desired traits in a partner. I also show how an alternative message space that asks an agent for both aspects can lead to dishonesty, even when used with well-known strategy-proof matching algorithms under direct reporting of preference rankings such as deferred acceptance for the proposing side. I identify the optimal strategy of an agent when the message space cannot generate their true preferences but the mechanism is strategy-proof and discuss the limitations of such a strategy in practice.

In the Chapter 2, I shift the focus on understanding the implications of the alternative message space to a two-sided, many-to-one matching market such as centralized public school choice. I also change the primitive of the agents' preferences such that they have preferences directly over the characteristics of other agents, and not the agents themselves as in the first chapter. I discuss the advantages that the alternative message space approach have in markets where there is incomplete information, such as when a student only discovers the characteristics of a subset of schools and thus can only rank those schools if asked to directly submit a ranking. I then study the effects of switching the agents' reports from a traditional matching framework where agents would submit the ranking directly over the schools they have discovered to one in which they instead submit their preferences through the alternative message space. I consider the implications of such

a switch with two of the most well-known and widely used matching algorithms, deferred acceptance and the Boston mechanism. I find that any individual agent, if that agent is the only agent asked to switch from one report type to the other and if the matching algorithm is deferred acceptance, will prefer the match he receives by reporting preferences through the alternative message space. I find that if the matching algorithm is the Boston mechanism, the agent may still prefer the outcome from directly submitting his ranking over the subset . I also consider asking all agents to switch simultaneously from one method of eliciting preferences to the other, as if the market designer suddenly implemented an alternative message space in the market. I find that with either matching algorithm, there may be an agent who preferred their outcome prior to the switch, showing that such a switch does not always result in a Pareto improved outcome, even if it may in some cases. Finally, I show that under specific conditions, where each seat at each school is full and where students have preferences on characteristics such that they agree on a common ranking of schools, switching from direct reporting of rankings to the alternative message space will always result in an outcome that is not Pareto comparable to the outcome under direct reporting. I discuss the implications of these results for market designers, especially the likelihood of resistance to changing the report type. I find that while there are benefits to switching, such as reduced need for individual agents to spend resources discovering the characteristics of potential matches, market designers are likely to encounter resistance to such a change and will need to carefully consider asking agents to switch methods.

Using an alternative message space where agents submit preferences over characteristics of others allows the market designer to gather information from the agents and use that information on their behalf to generate a preference ranking for them. The designer must decide how much, if any, of this information to share with the agents in the market. In Chapter 3, I collect survey data by randomly assigning respondents to a control group and two treatment groups to test if sharing all or some of the designer's information affects how agents report their preferences over characteristics. In the survey, I use characteristics of potential romantic partners, allowing the first treatment group to see all potential partners and their characteristics, the second to see only some of the potential partners and all of their characteristics and the control group to see no potential partners. The treatment groups are required to rank directly the potential partners they are shown, then asked to give their preferences over the characteristics themselves by answering binary-choice questions about each one and ranking their importance. The control group is only for their preferences over the characteristics. I find that there is some significant evidence, though somewhat weak, that exposing the agents to the information about their potential partners does change how they report their preferences over characteristics. Using an alternative message space that elicits preferences over the characteristics creates a restriction on the set of preference rankings that can be generated for the agents. With my data, I am also able to test if the rankings submitted by the treatment groups fall into the set that satisfies this restriction. I find that overwhelmingly they do not. In fact, only a very small portion of the respondents selected the same most preferred potential partner as was generated by their answers to the questions regarding their preferences over characteristics. While this result is limited in scope by the particular features of the survey,

it does offer some suggestive evidence that the restriction on the set of preference rankings that can be generated may be an important one in some markets.

Acknowledgments

This work has benefited from the contributions and comments of many brilliant people, and I want to thank them very much for their assistance and support. I am deeply grateful to my advisor, Professor Steve Williams, for his time, comments, support and patience. His guidance was essential in finishing this project. I would also like to thank the members of my committee for their valued time and assistance in completing my thesis: Professor Dan Bernhardt for his insightful comments and guidance when I presented the work and in helping me understand and summarize what I had accomplished in my thesis; Professor Nolan Miller for his guidance early on in helping me shape the direction of my research, his comments through many revisions of my job market paper and his suggestions when I presented my papers; and Professor George Deltas for his comments that allowed me to improve my final chapter despite limited data and his professional advice in the later stages of my research. I would also like to thank the Department of Economics University of Illinois at Urbana-Champaign for its continued financial support through my PhD, without which I would not have been able to complete this project.

I would also like to thank my many colleagues whose advice and discussions helped focus my research and improve my work. To my cohort, and friends, who supported each other in the early years of our studies and will do great things in the future, thank you. I give a special thanks to the members of the microeconomic theory group in the Department of Economics whose contributions were numerous and so beneficial in improving this work.

I am deeply grateful to my parents, family, friends and loved ones whose love, support and belief, among so many other things, made this and all aspects of my life possible. Finally, I thank God for the blessings in my life and for the incredible opportunities he has granted me.

Contents

Chapter 1. Can Designers Create Better Dates? Incentives in a Matching over Characteristics	1
1.1. Finding Matches in Practice: OkCupid and eHarmony	4
1.2. Model	7
1.3. Main Results	10
1.4. Advice on Designing the Message Space	21
1.5. Conclusion	23
Chapter 2. Matching over Characteristics with Incomplete Information	25
2.1. Matching Algorithms and Properties	28
2.2. Model	31
2.3. Main Results	33
2.4. Conclusion	38
Chapter 3. Preferences that are Responsive over Characteristics and the Role of Information	40
3.1. Introduction	40
3.2. Survey Design	43
3.3. Results	44
3.4. Conclusion	51
Bibliography	53
Appendix	55

Can Designers Create Better Dates? Incentives in a Matching over Characteristics

Many recent papers in the large market, two-sided matching literature assume that agents can announce their entire preference ordering, which, in large markets, is often unreasonable. For example, it is unlikely that users of a large online dating service will be able to submit lists that individually rank the vast number of other users, as is a basic and common assumption in the standard matching literature. Indeed, a common practice in this type of market is to elicit information about each member's characteristics and their preferences over *characteristics* of other agents instead, through what I will call an alternative message space. In alternative message space mechanisms, composed of an alternative message space and a matching algorithm, an agent's incentives for honest reporting may be quite different from those of the standard approach of a matching algorithm with directly reported preference lists. This paper examines the questions of strategy-proofness and design in such environments.

Throughout the paper, I argue that an alternative message space, if well-designed, can be effective in reducing incentives to lie in non-strategy-proof matching algorithms. However, if poorly designed, it can create incentives for some agents to strategically manipulate preferences even with algorithms that are well-known to be strategy-proof for those agents when allowed to directly report their preference lists. I provide advice for mechanism designers to improve the design of the alternative message space. My paper frames this discussion in the context of online dating, though my results are applicable to many markets that share similar features.

In this work, I will study the mechanisms of two popular dating sites, OkCupid and eHarmony. These sites were selected for several reasons. First, these sites typify the large dating markets I seek to analyze. eHarmony claims more than a million registered users in the New York City area alone, while OkCupid boasts over 400,000 in that market, as of 2015. Second, enough public information about their mechanisms exists to be confident that the approximated mechanisms and model discussed in this paper are similar to the actual mechanisms, giving my work a firm grounding in real-world markets. And third, each provides a distinctly different approach to constructing the message space, eliciting preferences and matching agents. While I cannot model these mechanisms exactly, as there are features of the market I will simplify to make the problem more like a standard matching problem, I attempt to get as close as I can with the information available. This allows me to frame and motivate my theoretical discussion, as well as provide some advice

for practitioners on how to design their message space, linking my theoretical results to the larger matching literature and back to practical application.

I contribute several results in markets with an alternative message space. I first define what makes an alternative message space *sufficiently rich*, that is, if the alternative message space offers each agent a report over characteristics that generates a preference list exactly matching his or her true preference list. I prove a lower bound for the size of an alternative message space that will be sufficiently rich and discuss how this approach limits the class of preference lists that can be represented. I show that a strategy-proof matching algorithm under direct reporting, that takes as inputs the preference lists generated from the characteristic reports in the alternative message space, will give agents incentive to report honestly when the alternative message space is sufficiently rich. I also identify the optimal strategy for an agent when the message space is not sufficiently rich and discuss an agent's likelihood of playing such a strategy.

I also show that a strategy-proof matching algorithm under direct reporting of preference lists is not necessary for obtaining honest reporting as a Nash equilibrium, which offers hope if the matching algorithm itself is not strategy-proof, as is often the case in practice. In fact, the alternative message space can sustain a Nash equilibrium that is not possible under direct reporting. Additionally, I establish a result that even a matching algorithm that is strategy-proof under direct reporting of preferences, such as deferred acceptance for the proposing side, may fail to guarantee honest reporting for any agent if the message space separately elicits preferences over characteristics and desired traits. I also show how these results might be applied practically for a mechanism designer in these markets.

The "large market" distinction, though it may seem obvious for online dating markets, has important implications in the matching literature. In many matching applications it has been shown or suggested that, all else equal, problems of manipulability might be lessened in a larger (even if still finite) market (Kojima and Pathak 2009, Roth and Peranson 1999). In one-to-one matching settings, Immorlica and Mahdian (2005) establish that the proportion of the non-proposing side that is matched to the same agent in all stable matches approaches one as the market grows. This has been shown to be necessary and sufficient for the non-proposing side to report honestly as well in deferred acceptance (Gale and Sotomayor 1985). Azevedo and Budish (2013) classify deferred acceptance as strategy-proof in the large, similarly showing that it is an approximately strategy-proof mechanism in large markets as agents achieve, within a very small bound, their maximal utility by reporting honestly. Because these results rely on large markets, it is important to understand the information and computational burdens placed on agents in these markets by their use. This paper offers an alternative to intensive task of ranking all agents populating the market individually.

A key contribution of the paper is to discuss mechanism design in markets in which agents cannot rank all other agents on the other side of the market, which is a common feature of many in-use matching markets but that has not received researchers' full attentions. This inability to rank other agents occurs in a specific way. It is not simply that agents can only submit a shortened list, which has been studied before, especially in the school choice matching literature (Calsamiglia, Haeringer and Klijn, 2010; Haeringer and Klijn, 2009; Pathak

and Sonmez, 2013a). Instead, agents must submit their own characteristics, preferences about characteristics that other agents have (their desired traits), or both. These mechanisms use questionnaires to elicit this information, altering the message space to ease costs on the individuals who join. Instead of ranking a million users individually, an agent might be asked a number of questions concerning what she likes in a partner or her own characteristics. Based on the answers to these questions, the sites then use an algorithm to generate full preference lists for the agents and match participants together.

My results can apply more broadly to a number of markets that share the inability of agents to rank other agents. In some cases, participants have no way of knowing who populates the other side. This may be due to legal privacy matters as in medical markets, or, as in dorm roommate matching problems, where identities remain anonymous until after the match. In these situations, a market designer might use blood types and health questionnaire responses to create a ranking to use in the match or ask about preferences for cleanliness and sleeping or studying habits. Another type of market that shares this feature is those where the timing may not be the same for both sides. One side of the market may not be established until it is clear that there is enough demand from the other. A mentoring program that matches professionals to undergraduate students might fall in this category. Professionals require more effort from the organizers to participate, and so are recruited after knowing how many students are interested. Additionally, a good mentoring program would likely want information about student preferences even before searching for professionals to ensure the mentors recruited were appropriate. In all of these cases, agents are unable to form a direct ranking of the other side of the market and my results would likely offer insights in these markets as well.

I will take the inability to directly rank agents as a feature of the market. I find this approach preferable as, in other markets such as those above, there may be many different reasons why agents cannot rank other agents. Though a convincing argument can be made that in online dating it is too time-intensive and costly for a man to get to know millions of women well enough to rank them, I will not rely on this reason for why they cannot. Instead, I note that even if a man did have a complete ranking of the women, the mechanism is not designed to accept such an input. There is simply no avenue to submit a full preference list on the website. By remaining vague on why agents cannot rank other agents directly, I avoid limiting my results only to one particular market and I can more clearly connect my results to the rich understanding already available in the matching literature.

The rest of the paper is organized as follows. Section 1 discusses the mechanisms used by two of the largest online dating sites. Section 2 presents the model from which I derive my results. Section 3 states the main results of the paper. Section 4 offers advice for designing the message space to minimize incentives to lie. Section 5 concludes.

1.1. Finding Matches in Practice: OkCupid and eHarmony

I analyze the mechanisms currently in use by two popular dating sites. I find two very different approaches, each of which can be valuable in illuminating the possible outcomes of a mechanism.

In order to anchor my analysis in reality, I attempt to approximate the real mechanisms of two popular dating sites, OkCupid and eHarmony. Each site takes information from the users through a questionnaire regarding characteristics of themselves or their prospective partners. The sites then, internally, create an ordering of agents on the other side of the market based on this information. They, at least implicitly, generate an internal preference list, that is used in the site's particular matching algorithm to output acceptable matches for the agents. In reality, the sites offer more than one potential match, though I will simplify by having the approximate mechanisms match each agent to one single agent on the other side. The individual agent never sees this internally generated ranking of other agents. Even if he did, it is a key feature that he cannot manipulate such a list directly, only by changing the answers to his initial questionnaire.

The two mechanisms differ in several important ways. OkCupid has a “dual-aspect” message space, meaning that the website asks for two separate inputs from the agents, information about their own characteristics, and preferences over the characteristics of others. Agents are allowed to choose how many questions they would like to answer. eHarmony has a “single-aspect” approach. It begins with a very long questionnaire, with 436 questions over 29 categories, all of which must be answered. This questionnaire only asks for characteristics about the individual. Based on these answers, eHarmony matches two people whose characteristics are compatible, according to their past research. Additionally, OkCupid is transparent about its matching procedure. Its site includes a page devoted to explaining how exactly how matches are chosen. This page is available to all visitors, even unregistered users. The preference list generation method and matching algorithm for eHarmony are much more closely guarded. Indeed, the site seems to purposely avoid explaining anything specific about how matches are made. However, the information available through its site along with its patent filing offers some insight into the process.

OkCupid advertises itself as “using math to get you dates.” It claims there is no complicated psychology in its mechanism: people know what they want and can articulate it clearly enough to ensure quality matches. An agent tells the site what she wants and who she is, the site combines that information with the responses of other agents, and returns the matches with the highest match-percent. Calculating this match-percent, or as the site describes, “the probability the two would get along together,” is done in the following way. As seen in Figure 1, a user is asked a question. He must provide the answer as it pertains to himself, the answer he would prefer a woman to choose, and the importance that her answer is as you'd like it. The level of importance selected determines how many points this question is worth when calculating the probability the man would get along with the woman. The importance levels (and point values) are irrelevant (0), a little (1), somewhat (10), very (50), and mandatory (250). In the example question from Figure 1, B's answer of '1' does not match A's preferred answer of '2', thus giving B a score of 0 out of 1 for that question, since it

Have you ever cheated in a relationship?

1. yes
2. no

Your answer	2
How you want someone else to answer	2
the question's importance to you	A Little Important
B's answer	1
How B wants someone else to answer	2
The question's importance to B	Somewhat Important

FIGURE 1. A sample question from OkCupid

was only “a little” important. A’s answer of ‘2’ does match B’s preferred answer, giving A a score of 10 of 10 since it was “somewhat” important to B. The probability each gets along with the other is found by adding the score received and the possible score for each question and dividing the first by the second. If Figure 1 was the only question, A would find B 0% compatible, while B would find A 100% compatible.

To find the probability that they get along together, OkCupid multiplies these probabilities and takes the square root. In the above example, the match-percentage would be zero. For better understanding, a nontrivial example might be if A found B 96% compatible, and B found A 78% compatible, the match-percentage would be $\sqrt{96 \cdot 78} = 86.53\%$. To make the site less time-consuming, OkCupid allows members to answer as many or as few questions as they would like. As a result, the match-percentage for any two people is only based on questions they have both taken. From here, the actual matching is largely taken over by individual users. They are shown their top match-percentage agents after choosing filters over age, gender, distance and other, more advanced, options. Each of these filters can also be modeled as another question a user could answer in the alternative message space. For instance, the site could equivalently ask users how old they are and how old they wish their partner to be, and attach a large point value to that question, instead of allowing users to filter out certain ages. For my purposes, as I am focused more on the alternative message space than how users select among several proposed matches anyway, I will simplify this process by considering filters to be questions and by having the matching algorithm be the one that selects some match according to the priorities of the site, returning only one possible match to each agent. This may be an algorithm to find the greatest total match-percentage while stable or the highest number of matches above some threshold. This is a simplifying assumption, as in reality the sites offer several possible matches to each agent and let the agents sort out their exact final match.

eHarmony positions itself in the crowded online dating marketplace by claiming to base their matches on scientific compatibility. In stark contrast to OkCupid, its website states, “Our singles matching models are

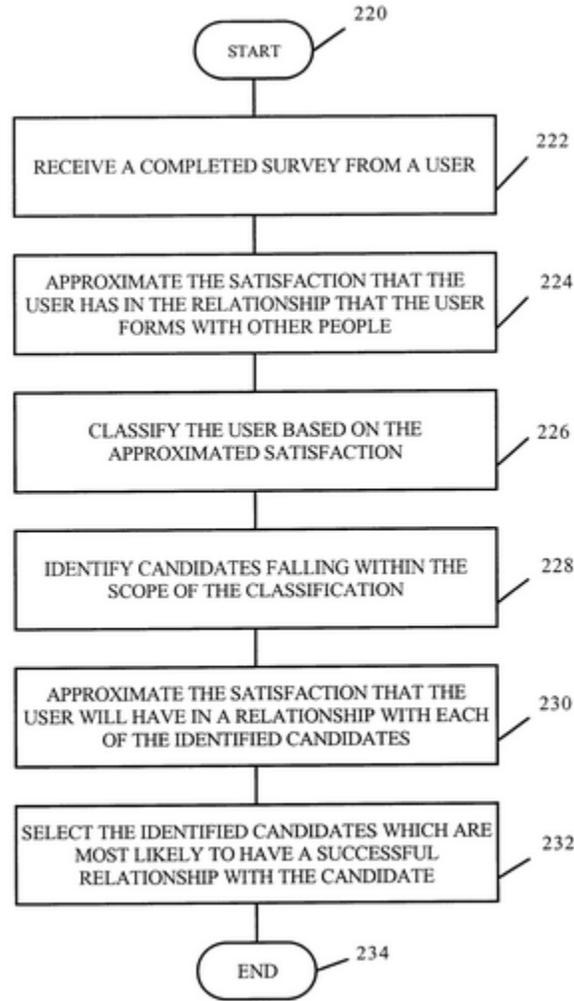


FIGURE 2. eHarmony’s Compatibility Matching System

based on 35 years of clinical experience and rigorous relationship research to determine which commonalities between partners are consistently associated with successful relationships.” In its patent filing, eHarmony describes the steps in its compatibility matching system as in Figure 2 (Buckwalter et al., 2004). The site takes a much more thorough approach to its questionnaire compared with other sites. While OkCupid’s questions are optional, allowing users to self-select their level of effort, eHarmony’s are mandatory. With 436 questions in 29 categories, its intensive questionnaire is designed to screen out uncommitted singles but also to identify agents who are not consistent or may be misrepresenting themselves.

The preference lists are then generated based on this information. The site first estimates the satisfaction an individual gets from being in a relationship in general. This might also be viewed as emotional readiness for a relationship. By sorting agents based on this statistic, and further using the answers to their questionnaires, the site then approximates the compatibility of an agent with all other agents, calling this the satisfaction score for an agent with another agent. Ranking all agents on the other side of the market yields the internal preference list for that agent. Simultaneously, the algorithm does the same for all agents on the other side

of the market. When the site has a satisfaction score for every agent with every other, it can calculate the “couple satisfaction index.” In the notes of the patent, eHarmony suggests this might be done by adding the two individual scores, averaging them, or differencing them. It is worth noting that while OkCupid uses a geometric mean and eHarmony uses an arithmetic mean or sum in calculating the couple score, they both end up with preference lists for the agents that depend on not only how much one agent likes another, but also how much the second agent likes the first.

From the information available, each site’s methodology appears quite similar after individual preferences are solicited. This highlights the importance of the preference elicitation portion of the mechanism, the alternative message space, emphasizing the need to understand how it may affect incentives. This also underscores the importance of the alternative message space in this type of market in the sense that this is how sites differentiate themselves. In the next section I present my results regarding this feature of the market.

1.2. Model

I model a two-sided, one-to-one matching market without transfers and with an alternative message space. As I am motivated by online dating markets, I follow the traditional interpretation of one side as men and the other as women introduced by Gale and Shapley (1962). In this market, there is a set of men $M = \{m_1, \dots, m_M\}$, and a set of women, $W = \{w_1, \dots, w_W\}$. These sets are disjoint and finite, and the cardinality of M and W may be different. I will refer to a nonspecific man as m and a nonspecific woman as w , and generic agents as i and j . Each agent has a true, strict preference ordering over agents on the other side of the market and remaining single, P_i . The set of all possible preference lists is P .

In this paper, I consider the setting in which agents have a full and complete ranking of each other as the primitive of my model. By doing this, I can consider the alternative message space as a particular feature added to an otherwise standard matching problem. Furthermore, by considering my agents’ basic preferences to be a ranking of other agents, I can model the alternative message space as solely an aspect of the design of the mechanism. This allows me to offer meaningful comparisons between incentives when agents can directly report preference lists and when they are forced to restrict themselves to the alternative message space. In another paper, I look more thoroughly at the implications of considering the answers to the questions themselves as the primitives. This approach changes the modeling significantly (see Sahajdack (2016)). I make an effort to distinguish which of my results hold in both information settings.

While agents rank other agents directly in this model, I also assume that every agent can be described by a set of characteristics. For example, a person might be funny, kind, or rude, among many other characteristics. Each agent is asked to report a finite dimensional, real vector of characteristics δ_i belonging to the set of all possible vectors of characteristics, Δ . A particular profile of characteristics vectors for a set of men and women is $\delta = (\delta_i)_{i \in \{M, W\}}$.

Agents may also be asked to report a finite dimensional, real vector of preferences over the characteristics of others that I call their *desired traits*, ϕ_i , with the set of all desired traits called Φ . A particular profile of desired traits vectors is $\phi = (\phi_i)_{i \in \{M, W\}}$. I distinguish between characteristics, which describe a person's own self, and desired traits, which describe the preferred characteristics of others, to avoid confusion in wording. However, both terms are used to describe the features of agents in the marketplace. For instance, "introverted" could be both a characteristic and a desired trait of a particular agent: the first referring to his own personality, the second referring to his potential partner's personality. I make the assumption that the characteristics and desired traits that the mechanism designer asks agents to report are binary. For the characteristic "introverted," select a value of one in the vector if the agent wishes to report that this describes him, or zero if he wishes to report that it does not. For the desired trait "introverted," select a value of one if the agent wishes to report that it describes his preferred partner and zero if not.

In my model, agents cannot directly submit preference lists over other agents, they only submit their own characteristics and desired traits. This is the *alternative message space*. I distinguish between two types of alternative message spaces, single-aspect and dual-aspect.

DEFINITION 1.1. A alternative message space is single-aspect if every agent's report in the alternative message space is only their characteristics, δ_i , or their desired traits, ϕ_i .

A alternative message space is dual-aspect if at least one agent's report in the alternative message space is both their characteristics and desired traits δ_i and ϕ_i .

The mechanism for a particular site is two-part. In the first step, sites ask users to submit their information. I call this step preference elicitation and this portion of the mechanism is the message space. If the agents' preference lists are reported directly, the message space is the same as a standard matching problem. In the alternative message space, however, the information elicited might be both characteristics and desired traits, or only one of the two. If this is the case, then, using the information submitted, an internal preference list for each agent can be generated, $L_i(\delta, \phi)$. Agents may know exactly how this list is generated from the reports, such as with OkCupid, or they may not be aware, as is the case for eHarmony. I will call the profile of all agents' generated preferences $L = (L_{m_1}, \dots, L_{m_M}, L_{w_1}, \dots, L_{w_W})$.

In the second step, agents are actually matched together. Agents are matched by the site using its matching algorithm. A matching is a function $\mu : M \cup W \rightarrow M \cup W$ such that $\mu(m) \in W \cup \{m\}$ and $\mu(w) \in M \cup \{w\}$ and $\mu(m) = w$ if and only if $\mu(w) = m$. A matching joins an agent from one side of the market to another or leaves them single, where $\mu(i) = i$. The set of all possible matchings is \mathcal{M} . A matching algorithm is $\Omega : P \rightarrow \mathcal{M}$ that takes preference lists as inputs and returns a matching. Some examples of matching algorithms include deferred acceptance, the Boston mechanism, top trading cycles and linear optimization algorithms.

An alternative message space mechanism is $\Psi : \{\delta, \phi\} \rightarrow \mathcal{M}$ that takes reports of users in the alternative message space and delivers a matching. Because the mechanism is two distinct parts, it may also be written as $\Psi(\delta, \phi) = \Omega(L(\delta, \phi))$ where L is determined in the first step then used in a matching algorithm.

Finally, I formalize the notions of stability and strategy-proofness.

DEFINITION 1.2. A matching μ is stable if it is individually rational and there are no blocking pairs.

Individual rationality is formally defined as $\mu(i) \succ_i i$, where any assigned match must be preferable to remaining single. This is another way of stating that participation is voluntary. A blocking pair is a pair (m, w) for which each member prefers the other to their assigned match, i.e. $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. The existence of blocking pairs may cause an assigned match to break apart as agents leave their partners. An algorithm that always generates a stable matching is called a stable algorithm.

DEFINITION 1.3. A matching algorithm, Ω , is strategy-proof if it is a (weakly) dominant strategy for agents to submit their true preference list under direct preference reporting.

DEFINITION 1.4. A report $r_i \in \{\delta_i, \phi_i, (\delta_i, \phi_i)\}$ is true if and only if $L_i(r_i) = P_i$.

DEFINITION 1.5. An alternative message space mechanism, Ψ , is strategy-proof if it is a (weakly) dominant strategy for agents to submit their true report.

I distinguish between matching algorithms, Ω , which may be strategy-proof under direct reporting of preferences, and alternative message space mechanisms, Ψ , a combination of an alternative message space and a matching algorithm, which may or may not be strategy-proof regardless of whether the particular matching algorithm used is strategy-proof in itself.

As an example, deferred acceptance is a well known matching algorithm. In this algorithm, one side of the market is chosen as the proposing side. Suppose women are chosen as the proposers. All men and women submit a list of preferences over agents on the other side. The algorithm proceeds as if each woman on the proposing side makes an offer to her top choice man. Agents on the receiving side tentatively choose the most preferred offer among their proposals to hold to the next round, rejecting all others. All proposing

agents who were rejected make offers to their next best choice in the next round. Again, all agents who received offers examine their new offers from this round and any offer they held from the previous round and tentatively select the best one. All women who were rejected in this round proceed to their third favorite man and make an offer. This process continues until all proposing agents run out of new agents to whom to make offers. The result is a matching. This algorithm has many desirable properties. It always terminates in finite time and it selects the proposing-side-optimal stable match. Most importantly, it is always stable and is strategy-proof for the proposing side.

An alternative message space mechanism in this example adds an alternative message space to this algorithm. Instead of submitting preference lists, agents would answer some questions, the mechanism designer would generate a preference list for each agent, then the designer would run the deferred acceptance algorithm with these preferences. This would result in a matching. The properties of the alternative message space mechanism are the main focus of this paper.

Strategy-proofness for all agents with stability is impossible to achieve in finite markets. Deferred acceptance, for instance, guarantees stability but is only strategy-proof for the proposing side of the market. In order to achieve strategy-proofness for all agents, stability would be sacrificed. As an example, if agents are required to submit a preference list over all other agents, a serial dictatorship run for one side would be strategy-proof for all agents. Such an algorithm is not stable. In large markets, even finite markets, a strategy-proof in the large algorithm (deferred acceptance in matching) is a “useful second-best,” offering stability and some practical improvements in incentives to be honest (Azevedo and Budish, 2013).

1.3. Main Results

I highlight here my main theoretical contributions. I define a sufficiently rich alternative message space and show a lower bound on the number of questions that will result in being sufficiently rich. I show that for a restricted class of preferences, the message space requires fewer binary questions than eliciting a ranking over agents directly and in this way, reduces communication complexity. I also show a sufficient condition for honest reporting in a sufficiently rich alternative message space: a strategy-proof matching algorithm. I establish how the design choices of a message space can have effects on incentives. I argue that the choice of a message space can have very important implications for users and their reporting.

My results focus mainly on one aspect of the model, eliciting preferences through the alternative message space. Because it is impossible for agents to completely rank the other side of the market directly in this model, the method of gathering agents’ preferences plays an essential role in ensuring (or failing) to maintain the preferable properties commonly sought in standard two-sided, one-to-one matching models, stability and strategy-proofness.

It is important to first distinguish between an alternative message space that is sufficiently rich enough and one that is not.

DEFINITION 1.6. A alternative message space is sufficiently rich if for all agents there exists a true report in the alternative message space that generates a preference list in the mechanism that exactly corresponds to his or her true preferences. That is, $\forall i, \exists \{r_i\}$ s.t. $L_i(r_i) = P_i$.

If the alternative message space is sufficiently rich, then the mechanism is a direct revelation mechanism. Agents instead could submit their types directly. In matching applications, one way to think about an agent’s type is to consider it a way for the mechanism designer to recover his preference list for use in the matching algorithm. If the alternative message space is sufficiently rich, it is as if agents submit their preference list, albeit in a different form in practice or via their type in theory. This connects the mechanisms in this paper with the standard matching literature from a theoretical perspective, though there are still practical implications of the alternative message space to consider.

While in theory, any direct revelation mechanism can be reduced to a simple report of one’s type, in practice the details of how this report is collected often matter. I show below that for a class of preference lists, the use of the alternative message space reduces the communication complexity in this sense that fewer binary pieces of information need to be communicated to the mechanism designer by each agent. The class of preferences lists that result from using the alternative message space *responsive over characteristics*.

DEFINITION 1.7. Let agents A and B have vectors of characteristics δ_a and δ_b that are identical in all elements except one, where the value of that characteristics for A is a and the value of that characteristic for B is b . A preference list is responsive over characteristics for agent i if when $a \succ_i b$, then $\delta_a \succ_i \delta_b$.

If the class of preference lists is restricted to responsive over characteristics, it may be both sufficiently rich and require less communication than eliciting a ranking of agents directly.

Determining a Lower Bound and its Relationship to Responsive over Characteristics Preference Lists.

Given that the richness of the message space matters significantly in my findings, determining if an alternative message space is sufficiently rich is central to understanding the applications of the theory this paper offers. Let n be the number of agents needing to be ranked in a market. Let c be the number of choices for each question in the questionnaire. For instance, a “binary” question such as “Do you like a partner to be clean or messy?” would result in a c of two. I structure the questionnaire such that agents answer the questions about which choice they prefer, then are asked to compare each question to the others and report which is more important. This results in a report of the agent’s preferences regarding a particular characteristic, coming from the first type of question that I call type-one or q_1 , along with a ranking of those characteristics in terms of importance, coming from the second type of questions that I call type-two or q_2 .

Continuing the previous example, a second q_1 question might be “Do you like a partner to be introverted or extroverted?” and a q_2 question would ask which of the two type-one questions is more important, such as “Is your partner’s organization or personality more important to you?” These three questions would then allow the mechanism designer to generate rankings of the other agents according to the answers.

THEOREM 1.1. *A lower bound on the number of questions, q , needed to create a sufficiently rich message space is $\frac{\log_c(n)}{2}(1 + \log_c(n))$.*

PROOF. To apply a unique combination of characteristics to every agent to be ranked, it must be that

$$c^{q_1} \geq n.$$

This guarantees that there will be enough combinations of answers to uniquely identify every agent needing to be ranked with one of these combinations. As I am looking for a lower bound, I set this equal to find the minimum. This yields

$$q_1 = \log_c n.$$

But only assigning identifying combinations of characteristics to each agents is not enough to develop a ranking over these combinations. In order to generate a ranking, it is necessary to know the order of importance of the characteristics. This can be done by asking pairwise comparison questions about which characteristic is more important. If two agents on the other side of the market both had one characteristic that matched a man’s desired traits, he would be unable to rank them without this second type of question. In this sense, type-two questions do not require an agent to have actual rankings of the characteristics, it only allows the user to provide additional information to avoid indifferences in his preferences.

Because there are $\log_c n$ questions about the characteristics themselves, comparing these questions to determine importance requires an addition set of questions q_2 . q_2 can be found as follows

$$q_2 = \binom{\log_c(n)}{2} \text{ or } q_2 = \frac{\log_c(n)(\log_c(n) - 1)}{2}.$$

This yields a total number of questions

$$q = q_1 + q_2 = \log_c n + \frac{\log_c(n)(\log_c(n) - 1)}{2}.$$

Simplification results in

$$q = \frac{\log_c(n)}{2}(1 + \log_c(n)).$$

Suppose the mechanism designer needs to generate a ranking of n agents but is using a $q < \frac{\log_c(n)}{2}(1 + \log_c(n))$. I show that this message space cannot be sufficiently rich. There are three cases that result in a q that is too small. Either q_1 is too small, q_2 is too small, or both are too small. Considering each case:

Case (1): $q_1 < \log_c n$. In this case, there are too few combinations of characteristics to assign each agent a unique combination. Specifically, at least $n - \log_c n$ agents must share a combination of characteristics with another agent. When a ranking is formed using the type-two questions, some agents must be indifferent, given that they have the same combinations of characteristics. As the resulting L must contain indifferences while the true preference lists P are strict, L cannot be identical to P . The alternative message space then cannot be sufficiently rich.

Case (2): $q_2 < \frac{\log_c(n)(\log_c(n)-1)}{2}$. In this case, there are not enough questions to sort the agents once they are assigned a unique combination of characteristics. If the designer cannot discern which combinations are preferred, either there will be indifferences in the list L between agents with different combinations of characteristics or some agents simply will not be able to be compared and thus cannot be ranked. In either event, it is again impossible for L to be identical to P , and the alternative message space cannot be sufficiently rich.

Case (3): If both q_1 and q_2 are too small, the result follows from Case (1) and Case (2). □

To avoid problems with generated preferences becoming cyclical, it should be noted that the answers to type-two questions, those about which characteristics are more important, are restricted to be transitive. This is a convenient technical assumption, but does lead to a restriction on the resulting preferences that they are responsive over characteristics. Responsive over characteristics preference lists rule out characteristics being complementary, where the presence of one trait makes another more desirable than it would be otherwise. As an example, if four men each had a unique combination of two characteristics, responsive preferences would rule out the case where the man who matched both characteristics to the desired traits of the woman is ranked first, the man who matched zero is ranked second, and the men who matched one were ranked third and fourth. For a full example, see Example 2 in the appendix.

In the abstract, satisfying the responsive over desired traits preference ranking restriction is not an issue as the mechanism designer can simply choose characteristics that are not complements to use to uniquely identify agents. In reality, unless all agents closely agree on which characteristic pairs would be complementary, this will likely cause the number of questions necessary to balloon as more questions would be needed to represent the preferences of agents who have differing ideas on which combinations are complementary. At the lower bound, the type space of agents whose true preferences can be represented from the alternative message space is limited to those for whom the order of agents on their preference rankings is possible given the responsive nature of the generated preferences. On the other hand, as the number of questions increases to deal with this problem coming from the responsive generated preferences, the type space able to be represented also expands.

As an example to put Theorem 1.1 in context, I consider again the example of binary questions, such as “yes or no” questions or choosing between two alternatives. This type of question is very commonly used by real sites in these markets. Table 1 demonstrates the gains an alternative message space can offer

as compared to asking agents to individually rank the other agents using pairwise questions. This is one representation of the growth rate using pairwise questions about agents’ preferences that has been shown to be $\Omega(n^2)$ (Gonczarowski et al., 2015). Theorem 1.1 shows that the growth rate of the alternative message space using binary questions is $\Omega((\log_2 n)^2)$. Of course, the “questions needed” column shown in Table 1 is the best case scenario in that I show results at the lower bound, as if all the agents’ preference lists are responsive over characteristics. Nonetheless, it suggests that there will be advantages in terms of the time, information and communication burdens placed on agents in these situations if the mechanism designer can even approach this lower bound.

TABLE 1. Minimum Number of Questions for a Sufficiently Rich Alternative Message Space with Binary Questions

Agents to be Ranked	Characteristics Needed	Questions Needed	Pairwise Questions Needed for Ranking
n	$\log_2(n)$	$\log_2(n) + \binom{\log_2(n)}{2}$	$\binom{n}{2}$ or $\frac{n(n-1)}{2}$
3	2	3	3
8	3	6	28
100	7	28	4,950
1,000	10	55	499,500
1,000,000	20	210	499,999,500,000

It is important to note here that the lower bound can be reached for the restricted class of preference lists that are responsive over characteristics. Despite this, representing this class of preference lists using pairwise questions that rank each agent directly still requires the same number of questions as required to represent the full set of preference lists. In this way, it becomes clear that for some markets where preferences lists are responsive, using an alternative message space creates real reductions in the communication necessary to elicit the agents’ preference lists and benefits the agents if this communication is costly. However, representing the full class of preference lists using the alternative message space would require at least as much, if not more, costly communication as any other direct revelation mechanism. This must be the case, as Segal proves that at least $\Omega(n^2)$ Boolean queries are necessary to generate the full class of preference lists (Segal, 2007). This result is also proved by Gonczarowski et al. (2015).

It is also important to make clear that then lower bound presented above represents the best case for the market designer. It does not take into consideration how the designer chooses the characteristics. While I assume above that the designer is able to perfectly select the characteristics to uniquely identify the agents by combination, in practice, this is a daunting, if not impossible task in many markets. The only mitigating factor in this difficulty is that online dating sites how an incredibly large amount of data available to them tho help understand their markets.

However, even so, an online dating mechanism that uses an appropriate number of questions given a certain set of characteristics may not, in fact, function very well, even if the characteristics chosen are mostly well selected. Suppose a designer asks about ten characteristics, but is not aware of an important eleventh

characteristic that is important to the agents when ranking other agents. Even the existence of this one characteristic could hinder the usefulness of that mechanism severely. Thus the ability of a market designer to take advantage of the gains provided by the alternative message space depends heavily on their ability to select a sufficiently rich message space, not merely that it exists.

Dominant Strategy when the Alternative Message Space is Sufficiently Rich.

Having established a lower bound and recognized its limitations, I now analyze how the agents in the market respond to the alternative message space if the space is sufficiently rich and if it is not. I find a result for a market when it is sufficiently rich in Theorem 1.2.

THEOREM 1.2. *If a matching algorithm is strategy-proof under direct reporting of preference lists, then the alternative message space mechanism is also strategy-proof with any single-aspect message space that is sufficiently rich.*

PROOF. Because the alternative message space is sufficiently rich, there exists a δ_i such that for all $i \in \{M, W\}$, $L_i(\delta_i) = P_i$.

Furthermore, because the matching algorithm, Ω , is strategy-proof under direct reporting, submitting P_i is a dominant strategy. Following convention, let P_{-i} denote the reported preference lists of all agents other than i . Also, let $\Omega_i(\cdot)$ denote the outcome of the matching algorithm Ω for agent i , that is, his or her match. Thus, for any other reported preference list P'_i ,

$$\Omega_i(P_i, P_{-i}) \succeq_i \Omega_i(P'_i, P_{-i}).$$

Using the sufficient richness of the alternative message space,

$$\Omega_i(L_i(\delta_i, \delta_{-i}), L_{-i}(\delta_i, \delta_{-i})) \succeq_i \Omega_i(L_i(\delta'_i, \delta_{-i}), L_{-i}(\delta'_i, \delta_{-i})).$$

By the definition of an alternative message space mechanism,

$$\Psi_i(\delta_i, \delta_{-i}) \succeq_i \Psi_i(\delta'_i, \delta_{-i}).$$

□

Theorem 1.2 states that if it is a dominant strategy to have his true preference list go into the matching algorithm, it will also be the case that an agent will have a dominant strategy to report honestly in the alternative message space if it has certain features. Again, it is only possible for an algorithm to be strictly strategy-proof for all agents if stability is sacrificed. For an algorithm such as deferred acceptance, Theorem 1.2 would only hold for the proposing side.

The intuition of the proof is quite simple. The best way to ensure that an agent’s generated preference list, which is used in the matching, aligns with his or her own true preferences is for that agent to answer the questionnaire honestly. This result holds under other informational settings as well, including markets where agents do not know who exists on the other side of the market. The intuition in that case is similar: in large markets like online dating, it is likely that one or more fully compatible agents exist. As a result, agents will be wary of lying even once, due to the chance of missing out on such an agent by not having them ranked first. This result suggests that matching algorithms that are strategy-proof in large markets will extend this desirable property into the alternative message space.

For further intuition, I consider a top-down approach using a man as an example. Suppose there exists a woman who the man finds fully, 100% compatible. Any lie by the man must lower his own compatibility with that woman. In a single-aspect message space, that means a lie about his own characteristics. In a dual-aspect space, that means a lie about his desired traits in a partner. Any such lie, since it must lower the compatibility of the top woman but may improve the score of the second best woman, will either switch the two or keep the order the same. If it keeps the ranking the same, the man is indifferent between lying and telling the truth. If it switches the ranking, the man is made worse off by lying, as a distorted preference list will be generated by the alternative message space mechanism and used in the matching algorithm portion of the mechanism, which clearly violates the man’s dominant strategy. The same logic applies to the top ranked woman, even if she is not completely compatible. In this case, while lying may improve the top agent and lower the compatibility of the next best, this still results in an unchanged preference list, at best.

As pointed out above, Theorem 1.2 is not guaranteed to hold for a dual-aspect message space. In this case, when the alternative message space mechanism asks users to separately submit their own characteristics and desired traits, I find a negative result regarding incentives for an agent to honestly report his or her own characteristics. Furthermore, I find this result persists even under well-known strategy-proof matching algorithms.

THEOREM 1.3. In a dual-aspect message space, agents may have incentive to misreport their own characteristics, even if the alternative message space mechanism uses a matching algorithm that is strategy-proof for those agents under direct reporting of preference lists.

It is well known that the men-proposing deferred acceptance algorithm is strategy-proof for men ((Roth, 1982)). However, this result is for standard matching markets where agents directly submit preference lists. In these markets, there is no way for a man to affect his ranking in the preference lists of women. Under a dual-aspect message space, when compatibility is found by comparing a woman’s desired traits to a man’s characteristics, his position in her preference list might change with his reported characteristics. Consider the following example to illustrate the theorem.

EXAMPLE 1.1. There are two men and two women. Their preferences are as follows.

$$m_1 : w_1 \succ w_2 \quad w_1 : m_2 \succ m_1$$

$$m_2 : w_1 \succ w_2 \quad w_2 : m_1 \succ m_2$$

Under the men-proposing deferred acceptance algorithm, both m_1 and m_2 make an offer to w_1 in the first round. w_1 chooses m_2 according to her preferences, and m_1 and w_2 are then matched in the following round. In a standard setting, this is both stable and strategy-proof for men. However, suppose m_1 can misreport his own characteristics in such a way that he becomes more attractive to w_1 . Perhaps he claims to be taller, smarter, or in better shape. By doing so, he switches his ranking in her preferences from second to first. Now, in the first round, w_1 accepts his offer, instead of m_2 's, and m_2 and w_2 are matched in the following round. By misreporting, he now gets his top choice. Whether she stays with him after learning the truth when they meet is anyone's guess, but the incentive to lie in the alternative message space mechanism exists nonetheless.

Optimal Strategy for an Agent when the Alternative Message Space is not Sufficiently Rich.

I now consider when the lower bound is not met and the alternative message space is not sufficiently rich. I show an agent's optimal strategy in this case in Theorem 1.4.

DEFINITION 1.8. An agent j is achievable for agent i if there is a stable matching μ with $\mu(i) = j$.

THEOREM 1.4. *Suppose the matching algorithm is stable and strategy-proof for agent i under direct reporting but the message space is not sufficiently rich. Agent i should choose a report, r_i , such that the top ranked achievable match in the resulting preference list $L_i(r_i)$, is preferred under her true preferences to the top ranked achievable match in the resulting preference list from any other report, $L_i'(r_i')$.*

PROOF. If there is no report r_i that yields $L_i(r_i) = P_i$, the agent must choose a report r_i' that results in $L_i'(r_i')$. The best report an agent can make is one such that

$$\Psi_i(r_i', r_{-i}) \succeq_i \Psi_i(r_i'', r_{-i}) \quad \forall r_i.$$

Finding which report to submit requires an agent to understand how her report will affect her outcome.

Because the matching algorithm is stable, the set of agents with whom agent i can be matched is only those agents that are achievable for her. I will call this set of agents A_i . Since agent i is never matched with an agent outside A_i , changing the order of any two elements that do not both belong to A_i will not change

the outcome for agent i as long the order of all other elements are preserved. Thus the order of all elements outside A_i is irrelevant and only the order of elements within A_i is needed to find agent i 's best report.

Let $a, a' \in A_i$.

Let r_i^a be a report that results in a generated preference list ranking agent i 's most preferred achievable match, a , above her other achievable matches. r_i^a must exist, as all agents must be ranked under the alternative message space. Thus, trivially, there must exist a report in which one achievable match is ranked above the others in the generated preference list.

Let $r_i^{a'}$ be any report that results in ranking any a' above a as the most preferred achievable match. Suppose these were the only two reports an agent could make. Agent i searches each resulting preference list for the best ranked achievable match and compares the two based on her true preferences. Since a is preferred to a' , she selects r_i^a .

Suppose agent i reports $r_i^{a'}$. By contradiction, I will show that this cannot be her best report.

First, because a' is achievable and the matching algorithm is strategy-proof for agent i , ranking a' above a must result in $\mu(i) = a'$ and ranking a above a' must result in $\mu(i) = a$. If not, the agent would want to switch the two to gain the outcome she desires, violating strategy-proofness.

Thus,

$$\Psi_i(r_i^{a'}, r_{-i}) = a' \text{ and } \Psi_i(r_i^a, r_{-i}) = a.$$

For $r_i^{a'}$ to be her best report, it must be that

$$\Psi_i(r_i^{a'}, r_{-i}) \succeq_i \Psi_i(r_i^a, r_{-i}).$$

But by substitution, this is equivalent to

$$a' \succeq_i a.$$

Now there is a contradiction, as according to her true preferences,

$$a \succeq_i a'.$$

This proves the theorem, as a' is any arbitrary achievable match other than the true most preferred. Therefore, no report other than r_i^a can be the best report for agent i . □

Theorem 1.4 identifies an agent's best report if there is no generated preference list that matches his or her true preferences. However, the calculations necessary to find this report would not be possible unless the agent had a true preference list in mind and could perfectly predict how his or her report would change the generated preference list, so this result is limited by the information and computation available to an agent in practice. Also, this agent must know his or her set of achievable agents, which is likely a difficult task, especially in a large market setting, as knowing the set of achievable matches requires knowledge of

the preferences of all other agents. Thus while this is an agents optimal strategy if the alternative message space is not sufficiently rich, the likelihood of an agent being able to apply it perfectly is low.

Despite this, Theorem 1.4 can offer some insight into how an agent might behave. For example, in many centralized school-choice or college admission matching programs, it is common for students to rank a school where they are very likely to be accepted as a “safety school,” even if that school is not truly preferred to others. It makes intuitive sense that such a student would rank his most preferred safety school above any other safety school, at least. So while the strategy identified in Theorem 1.4 may be difficult to achieve in practice, it does appear that sophisticated agents heuristically attempt a similar strategy based on their limited information and the message space of those programs.

This result also emphasizes the importance of the mechanism designer carefully choosing the alternative message space. Given that the consequences for agents when the space is not sufficiently rich are likely negative, designers should be thorough in ensuring they understand the market and the preferences of the agents as much as possible.

Results when the Matching Algorithm is not Strategy-Proof.

Theorems 2, 3 and 4 give insights into incentives and optimal behavior for agents when matching algorithms are strategy-proof, but not when these algorithms fail to meet this criterion, such as with the matching algorithms of the sites discussed in this paper. Azevedo and Budish (2013) offer a classification of mechanisms in various mechanism design problems that are known to be *strategy-proof in the large*. In the matching literature, the only such matching algorithm they find is deferred acceptance. Unfortunately, the sites do not use a matching algorithm that resembles deferred acceptance, to the best of my knowledge. Instead, both sites’ emphasis appears to be heavily skewed towards maximizing total compatibility in the market. Though it may be that the outcome of deferred acceptance, the man- or woman-optimal stable match, offers the highest total compatibility score, this is far from guaranteed and would be only coincidental. It is extremely unlikely that deferred acceptance will be chosen as the matching algorithm if maximizing this score is the priority.

This is a worrying result, as if the matching algorithm is not strategy-proof for all agents, Theorem 1.2 does not guarantee that the alternative message space mechanism will be, no matter how it is designed. But I am able to mitigate this issue with the alternative message space. If the matching algorithm is not deferred acceptance or is not strategy-proof, honest reporting may be an unattainable goal. Luckily, the particular features of the market actually positively affect incentives to report honestly, even if it may not be a dominant strategy.

DEFINITION 1.9. A manipulation is an alteration of the preference list, whether that alteration is direct as when agents directly report preference lists, or indirect as in an alternative message space where the actual preference list is generated by the mechanism designer.

THEOREM 1.5. *Any manipulation possible using reports in an alternative message space can be replicated in a direct reporting environment. Not all manipulations in a direct reporting can be replicated in an alternative message space, such as any truncation strategy.*

A proof is provided in the appendix, though much is trivial. As an example, it would be impossible to switch the top and bottom agents in a generated preference list described in this paper through a report in the alternative message space, though this would be simple if preference lists were reported directly. Similarly, under direct reporting an agent could truncate his preference list, leaving some agents off of it. This is not possible in the alternative message space, as all agents are ranked by construction in this case.

COROLLARY 1.1. *There may exist honest reporting Nash equilibria in an alternative message space that cannot be sustained as honest reporting Nash equilibria under direct reporting.*

PROOF. Follows directly, any equilibrium where the manipulation required to gain by deviating is an element of P but not L will be sustained in an alternative message space but not with direct reporting. □

Theorem 1.5 and its corollary demonstrate how a well-designed alternative message space can improve incentives in a market. By ruling out some possible cheating strategies, a mechanism designer can limit an agent's ability to manipulate and gain from deviating. While the obvious manipulation that is taken away is truncation, the agents also are unable to permute their preference orderings freely. Example 3 in the Appendix gives an example of such a situation. It shows how an agent may be unable to change the order of a particular pair of elements without also changing their respective orders in comparison to a third. Introducing this third element would not be a concern under direct reporting, but in an alternative message space it may be unavoidable.

If there are no restrictions on how agents submit preference lists, then agents can directly manipulate their submissions by truncating their list or swapping two or more people. By asking agents to submit information about characteristics, it implicitly forces them to rank all agents on the other side of the market. They cannot truncate. This turns out to be very important for incentives. Roth and Rothblum (1999) show that under deferred acceptance when agents have little information about the preferences of other agents, truncation is the optimal cheating strategy. They further state that simply switching two agents will not benefit an agent. Ehlers (2008) extends this result to certain linear programming algorithms, similar to the

mechanisms that seek to optimize the total compatibility score as might be used in online dating. Provided the mechanism satisfies several fairly weak properties, truncating stochastically dominates all other cheating strategies. Teo, Sethuraman and Tan (2001) show that when agents must submit complete preference lists, strategic choices for these agents are limited and thus there are reduced incentives to lie and agents are less likely to be able to gain from deviating.

Since altering the message space doesn't allow truncation and requires a complete preference list by construction, this creates greater incentive to be truthful. As a result, if the algorithm cannot be changed, altering the message space may improve user incentives to be honest. This result holds for either single- or dual-aspect message spaces.

1.4. Advice on Designing the Message Space

In this section, I offer advice to mechanism designers that I hope will improve their design choices as they seek to meet their goals. I use my theoretical results to connect my matching problem to the larger literature on matching, which allows me to give advice using a broad theoretical base of understanding. Because the message space can have significant impacts on the agents' incentives to be honest, I focus my advice on that part of the alternative message space mechanism. I advise care in designing the alternative message space, even with matching algorithms that are strategy-proof under direct reporting of preferences. I recommend designers choose a sufficiently rich single-aspect message space with a strategy-proof matching algorithm if honest reporting is a priority. This advice ties my theoretical findings back to real-world applications.

REMARK 1.1. With less information, incentives to lie are likely lessened, whether single- or dual-aspect.

Since online dating markets are large, and likely thick considering the high number of users and successful marriages (Oyer, 2014), agents are likely to have a very compatible partner who exists in the market. With perfect information, they can manipulate, if necessary, in such a way that they can be sure not to hurt their chances of being matched with such a partner too much. However, if information in the market is not available, whether this is lack of knowledge of who exactly exists or lack of clarity on how that mechanism works, every lie a man tells risks dropping that highly compatible woman further and further down his list. Conversely, if he lies about himself, he misses the chance to find a woman who may actually love him for exactly who he is. Following Roth (2008) and Azevedo and Budish (2013), telling the truth in these situations is a safe strategy, as lying may gain you a small upside but risks a potentially large downside. Pittel (1989; 1992) places the average rank of a woman in the man-optimal deferred acceptance algorithm at around $\ln(n)$ or $\ln^2(n)$, as $n \rightarrow \infty$ with high probability, where n is the market size. While this does require agents to have independent, uniformly drawn preferences, it also suggests that in a market size of two million, agents are very likely to be ranked to another agent within their top ten. It would be very risky,

based on this result, to lie much, as an agent would almost surely be lowering their top choices and putting less desirable agents above them.

REMARK 1.2. If honesty is the main priority in designing the mechanism, the designer should use a sufficiently rich, single-aspect message space with a matching algorithm that is strategy-proof under direct reporting, even though this must come at the cost of stability. If honest is a priority but stability is also important, a strategy-proof in the large matching algorithm is a good choice.

At least according to the theory, eHarmony's mechanism is superior in terms of giving users incentives to tell the truth. Its questionnaire is single-aspect, only asking users about themselves, then it uses its own hidden algorithm to calculate compatibility. If the site could commit to a strategy-proof or approximately strategy-proof matching algorithm, such as deferred acceptance in large markets, then Theorem 1.2 would hold (or at least nearly). As personal characteristics are the only input, reporting these honestly would be an approximately dominant strategy with deferred acceptance, as it is strategy-proof in the large. By its nature, deferred acceptance will favor one side of the market at the cost of the other, but depending on the balance of men and women, it may be that men-proposing or women-proposing deferred acceptance is not unreasonable. eHarmony's thorough questionnaire also improves incentives. It can catch agents lying directly, by checking consistency of answers over several hundred questions, but also a detailed questionnaire helps more finely sort agents and mitigates issues due to indifferences.

If the designer believes that the tie-breaker accurately represents how users sort agents, then these indifferences may not be a major concern. For instance, if users truly tend to like dating other users who live closer to them, which seems likely, a designer could hope the indifferences are sorted out using distance as a tie-breaker. However, under very minimal questioning such as only answering a couple questions on OkCupid, it is conceivable to have hundreds of thousands of indifferent agents in very large markets. This will likely cause imperfect matches under true preferences. It may be better to add the tie-breaker in as an additional question or characteristic, thus expanding the message space and dealing with indifferences in that, more controlled, way. Finally, by keeping large parts of their mechanism private and unknown, eHarmony makes it difficult for agents to know exactly how their manipulations might affect their outcomes. This may improve incentives for honesty, as discussed above.

It should be noted that honesty may not, in fact, be the main goal of a practitioner. As I have focused on honesty throughout much of the paper, and though use of an alternative message space will affect a designers ability to achieve other goals, the extent and pathways of these effects remain an open question.

1.5. Conclusion

In this paper I study a real-world solution to a common practical problem in the matching literature. It is often assumed that agents can directly submit full preference lists of all other agents in the market. This can quickly become unreasonable as markets grow large or information scarce. Taking online dating markets as a motivating example, I formally model a common practice of these websites: altering the message space to having agents submit their own characteristics, and sometimes also their desired traits in a partner, instead.

Incentives for agents in the market to be honest can be greatly affected by the introduction of this type of alternative message space. I find that while a strategy-proof matching algorithm under direct reporting is sufficient for users to report honestly in a single-aspect alternative message space, it is not necessary for honest reporting to be a Nash equilibrium. This finding suggests that even with a matching algorithm that gives agents incentives to lie under direct reporting of preference lists, a well-designed alternative message space can limit their ability to do so and encourage honesty. Unfortunately, if a message space is poorly designed, I find that agents may have incentive to lie even when they would not under direct reporting. For instance, if the alternative message space is dual-aspect, separately asking users to submit their characteristics and desired traits, men may have incentive to lie even under the men-proposing deferred acceptance algorithm, well-known to be strategy-proof for men under direct reporting.

Because the design of the message space is so important for user incentives, I seek to provide advice to designers. I encourage designers to consider implementing an alternative message space to improve incentives for honesty when the matching algorithm itself cannot be changed. This occurs because, for any given possible manipulations of preference lists in a market, a structured message space cannot ever create more methods to manipulate and will often improve the likely of honest reporting by eliminating some deviating strategies, such as truncation. I also suggest that limiting information available to users may also limit their ability to lie, as they become unable to perfectly predict the results of their manipulations and more cautious about lying. Lastly, I advise that if honest reporting is a goal, a sufficiently rich, single-aspect message space, paired with a strategy-proof matching algorithm will provide the best results.

This work offers many areas for extension. First, I can adjust the informational primitives of the agents to only having preferences over characteristics, instead of other agents themselves. This extension allows me to say something about markets with incomplete information, though without addressing that problem directly. I discuss this information setting in another paper ((Sahajdack, 2016)). Second, an alternative message space is just one answer to a very practical question: how can designers deal with matching problems where agents cannot reasonably be expected to rank every alternative? In future work, I hope to more thoroughly analyze how this solution compares with another common practice in these markets, asking agents to submit a shortened list of their preferences. Third, I also hope to extend this methodology to other markets that share this feature, such as roommate matching in a one-sided setting or many-to-one matching problems such as centralized college admissions markets or school choice. Finally, future research in this area might

include data and information from the sites themselves, in order to see more clearly how these markets work in practice. This type of information would allow me to observe and simulate, creating stronger and more concrete recommendations on how to practically design an alternative message space.

Matching over Characteristics with Incomplete Information

Many two-sided markets, especially those in which prices are not or cannot be used as a medium of exchange, rely on matching to organize the participants. Two well-known and successful examples are the National Residency Matching Program (NRMP), which matches future doctors to hospitals, and school choice, which matches students to schools. These markets use a centralized algorithm to arrange the matches based on the rank-order preference lists submitted by the participants. Two such algorithms are the *deferred acceptance* algorithm, famously applied to marriage markets by Gale and Shapley (1962), and the *Boston mechanism*, named for its use in the Boston area school choice program (Abdulkadiroglu and Sönmez, 2003). Each of these algorithms has important properties and considerable effort has been spent by researchers to understand their implications in real world markets. One important aspect of these algorithms that, while known, has not been fully accounted for in considering their effectiveness is their outcomes when agents do not have full information. For example, it is easy to imagine mismatches forming if students are not aware of all schools and schools are not aware of all students, even under an algorithm that works well when all parties are aware of all others. In this paper, I study an alternative method of eliciting preferences that has some advantages when agents lack information about their potential matches. In my model, agents report preferences over the characteristics of other agents, instead of reporting a ranking of the agents directly. I show that the feasibility of transitioning to this type of reporting depends on the current mechanism in use and the details of the market.

The NRMP and many school choice markets use the *student optimal deferred acceptance* algorithm (SOSM) to arrange matches. In this algorithm, matches are formed as if agents on one side of the market, students, request to be matched with one member of the other side, in this case, schools. In the first round, every student requests his or her top choice school according to their submitted preferences. Schools tentatively accept as many students as they have capacity, often called *seats* at that school, according to their preferences over students, rejecting all others who applied. In the next round, any unmatched students make an offer to their next most preferred school. Schools look over the new pool of applicants, including those they held as acceptable from the previous round, and again select their most preferred students. The following rounds continue in the same fashion, with unmatched students applying to their next most preferred school until all students are either matched or have exhausted their acceptable schools. The resulting match has been shown to have several desirable properties under complete information and unrestricted direct preference list reporting. These include *stability*, in that no two agents prefer each other to their assigned match and

no agent prefers being single to his or her match, and *student optimality*, meaning that the match is most preferred by students among all stable matches. This student optimality also results in the SOSM being *strategy-proof* for students, meaning it is a (weakly) dominant strategy for students to report their true preference lists (Roth and Sotomayor, 1992). However, it is also known that these properties may not hold if agents are restricted to a shortened preference list or if they have incomplete information about agents on the other side of the market (Pathak and Sönmez, 2013b).

Other school choice markets use the *Boston mechanism*. This algorithm is similar to deferred acceptance in that in each round any unmatched agent requests a seat at his highest ranked remaining school, but has one important difference. This difference is that as agents request their choices in each round, they are immediately matched and removed from the algorithm if the agent on the other side of the market also requests to be matched with them. For instance, if a student lists a school first and that school finds that student acceptable among its applicants and has capacity for that student in that round, the student is admitted and that seat at the school is removed. In this sense, the Boston mechanism can be thought of as an immediate acceptance mechanism. Intuitively, it is clear that there is room for strategic manipulation in this case, as students may not want to waste the top rankings in their reported preference lists on schools at which they are unlikely to be accepted, only to find that schools for which they may have been more qualified are already filled when they request admittance in later rounds. Indeed, previous works in the matching literature have shown that the Boston mechanism is not strategy-proof and that many agents do manipulate in this way (Chen and Sönmez, 2006). This algorithm is also not stable; there may be students and schools who prefer each other to their assigned matches. Nonetheless, immediate acceptance mechanisms are a popular class of mechanisms in practice. One explanation is that some agents are more able to benefit from the manipulability of the mechanism. These strategic agents are often called *sophisticated* and would likely oppose a change in algorithm. Other explanations for the persistence of the Boston mechanism and other immediate acceptance algorithms cite specific settings where the Boston mechanism outperforms or is not demonstrably worse than the SOSM (Chen, 2014; Kumano, 2013). Due to these factors, and because the setting of this paper is nonstandard, I investigate the effects of eliciting preferences through preferences over characteristics under both mechanisms. As a practical consideration, this allows me to give some insight on the feasibility of moving from a traditional preference list mechanism to a preferences over characteristics mechanism.

Though there are many similarities between the NRMP and centralized school choice programs, there is one key distinction. In the NRMP, hospitals can act strategically, meaning they can submit any preference list they choose, including any manipulation of their preference lists. In school choice, schools often have their preferences set for them by law or by ordinance, such that they do not act strategically. In this case, their preferences are called *priorities* to distinguish that they can not act strategically in regard to the order of students they prefer. If hospitals are assumed to act non-strategically, as is sometimes the case, the two applications become equivalent problems. Eliminating the strategic behavior of schools or hospitals results in

a simpler problem for the researcher, as he or she needs only consider one side of the market misrepresenting their preferences. In this case, the usual notion of stability is replaced by *no justified envy*, which means that no student prefers a school where they have a higher priority than another student matched to that school. Stability and justified envy are very similar notions, both aimed at preventing agents from desiring to leave their assigned match for each other after the match is announced. Furthermore, as the SOSM is stable, at least according to the reported preferences, and the Boston mechanism is not when both sides can act strategically, so too does the SOSM result in no justified envy while the Boston Mechanism may have justified envy (Abdulkadiroglu and Sönmez, 2003).

There are many matching markets in which it is reasonable to believe that agents are missing some information about agents on the other side of the market. In the NRMP and school choice, it may be that the students do not know that a particular hospital or school exists. Even if every characteristic of that school was publicly available, if they are not aware of the school itself, they cannot rank it in a traditional matching framework. In other markets, agents may be aware of the existence of other agents, but not know any information about them beyond their name. A dorm roommate matching program may have this problem. According to its University housing website, at the University of Illinois, students choose a dorm first, then have the option of choosing a roommate. In choosing, students can see a first name and a few voluntary answers to questions about living habits. While a student may know a list of names of others in their dorm and even basic information for some others, the school cannot share important information about those other students for legal or privacy reasons. Based only on a name, the student may not be able to accurately rank potential roommates. In either case, the lack of information limits the agents' ability to accurately create the same ranking of other agents that they would under full information.

Though I believe that there are many matching markets that share the feature that agents may have incomplete information about the other side of the market, in this paper I focus on many-to-one markets, using the NRMP and school choice as examples for several reasons. First, agents in these markets have many observable or verifiable (at least to the mechanism designer) characteristics. For students, this may be characteristics like test scores, grades or address. For schools and hospitals, it might be class size, placements or facilities. This allows me to focus on the agents submitting preferences only on the characteristics of others, and not requiring them to also submit their own characteristics. In another paper, I show that when asked to submit both, agents have strong incentives to lie about their own characteristics (Sahajdack (2015)). Second, these markets are highly visible and well-known examples of matching in practice that, despite being studied extensively in the matching literature, have many issues of incomplete information in real world applications. This issues have not been fully addressed and need to be considered in giving design advice to practitioners. Last, from a practical design perspective, there is a developed literature on which characteristics matter for student outcomes in school choice (Betts, 1995; Hanushek, 1986; Card and Krueger, 1992). This suggests that mechanism designers have a solid foundation in choosing which characteristics to include when asking for participants' preferences.

In this paper, I study the implications of asking participants to give preferences over characteristics of the agents on the other side of the market, instead of ranking them directly. This method has several advantages, especially in terms of the information required by the participants. Each side of the market only needs to know their preferences about features of the other and not which of the other agents have those features. It also elicits a full preference list. Based on the answers to the questions, assuming the questions are thorough enough, the market designer knows the entire order of that agent’s preferences, instead of only a few, as it might from a reduced list. I present this method as an alternative to the current method that I believe can be represented by agents knowing about only a subset of the other agents and submitting a ranking over that subset. Asking agents to report preferences over characteristics means that agents do not need to expend time and energy to increase the number of schools or hospitals of which they are aware.

In Theorem 2.1, I show that under the SOSM, an individual student weakly prefers to be asked to report his or her preferences over characteristics. This holds as long as the order of schools he reports when asked to rank schools directly is maintained in the list generated based on his preferences over characteristics. Because the tentative acceptance aspect of the SOSM does not penalize students for ranking schools to which they are not accepted, at worst, the student will maintain the seat he had under reporting his preference ranking directly over the subset of schools he knows about. Theorem 2.2 shows that the same is not true for the Boston mechanism: an individual student may prefer to report her ranking over the subset of schools. This is because reporting preferences over characteristics expands the preference list used as an input to the Boston mechanism. This larger list may include schools that the student likes but has a low priority at, thus wasting a round by request admittance to an unlikely school.

Theorem 2.3 considers changing the report types of all students simultaneously from directly reported rankings of the subset of known schools to preferences over characteristics under the SOSM and the Boston Mechanism. I find that in either case, there may be a student who is hurt by the switch. Thus I cannot say that reporting preferences over characteristics always results in a Pareto improved outcome under either algorithm. These theorems provide insight to the challenges designers will likely face in attempting to implement this type of preference elicitation structure. As long as some student believes he or she may be hurt by the switch, the designer will face resistance in implementing a change in report type. The remaining sections of the paper formalize the model and results.

2.1. Matching Algorithms and Properties

In this paper, I examine two of the most used and most studied matching algorithms, the Student Optimal Deferred Acceptance algorithm and the Boston Mechanism. These algorithms, or variations derived from them, are used in many centralized matching applications, including by school districts and the NRMP. I have chosen these two particular algorithms due to the combination of the wide practical implications and the wealth of knowledge from the literature concerning their properties. For clarity and consistency, I will

use the language of school choice, though students and schools can be replaced with doctors and hospitals, respectively.

Matching algorithms are judged on many properties, depending on the goals of the market designer and the features of the market. Three properties that are generally considered desirable are *strategy proofness*, *no justified envy* and *Pareto efficiency*. *Strategy proofness*, in the context of school choice, states that it is a dominant strategy, at least weakly, for a student to submit his or her true preferences. This property is often considered desirable as a measure of fairness to all participants. If it is satisfied, no agent can be disadvantaged by failing to discern and report some sort of optimal cheating strategy due to lacking information or ability. It also implies that students do not need to form beliefs about other students' preferences, information or actions (Azevedo and Budish, 2013). *No justified envy* means no student prefers a school at which he has a higher priority than a student matched to that school (and also that no student or school is given a match that they find unacceptable). If schools are allowed to set their own preferences and act strategically on these preferences, this property would be replaced with stability. Satisfying no justified envy ensures that the participants have no incentive to leave their assigned match once it is announced, at least according to their reported preferences. A matching algorithm is *Pareto efficient* if it results in a match such that there is no other match that makes every student weakly better off and at least one student strictly better off. A Pareto inefficient match is often seen as a failure of the market designer to optimize the assignment in the market.

The Student Optimal Deferred Acceptance Algorithm.

The student optimal deferred acceptance algorithm (SOSM) originated with Gale and Shapley in 1962. Since then, it has been adopted by numerous school choice authorities and is the basis for the algorithm currently in use in the NRMP. Pathak (2014) provides a list of several school districts that use this algorithm, including Boston, Denver, New York and many districts in the UK.

In the traditional matching framework, the SOSM runs as follows when applied to school choice:

Step 1:: In the first round, each student applies to their most preferred school. Schools view all applications they receive. Schools tentatively accept students in order according to their priorities until they are full, rejecting all others.

Step k :: In any subsequent round k , any unmatched student from round $k - 1$ applies to his k^{th} favorite school. Schools view all applications, including both new applications and those held from round $k - 1$. Schools tentatively accept students according to their priorities until they are full, rejecting all others.

The algorithm proceeds in this manner until all unmatched students have exhausted their list of acceptable schools and no school rejects any application in that round. At this point, the tentatively accepted students are assigned a seat at the school to which they were matched.

The resulting match has been shown to be strategy proof for students and to result in no justified envy. Also, while it may not be Pareto efficient among all algorithms, it does Pareto dominate any other algorithm that eliminates justified envy (Abdulkadiroglu and Sönmez, 2003). Thus if eliminating justified envy is an important design goal, as it often is in practice, the SOSM is a favorable choice.

The Boston mechanism.

The Boston mechanism derives its name from the Boston school district where it was used to assign students to schools from 1999 to 2005. Though many school authorities have switched to other algorithms, it remains widely used, including in cities such Seattle, Minneapolis and Tampa (Pathak and Sonmez, 2013a; Pathak, 2014).

In the traditional matching framework, the Boston mechanism runs as follows:

Step 1:: In the first round, each student applies to their most preferred school. Schools view all applications they receive. Schools **immediately** accept students in order according to their priorities until they are full, rejecting all others. Matched students and the seats they fill are removed from the algorithm as assigned matches.

Step k :: In any subsequent round k , any unmatched student from round $k - 1$ applies to his k^{th} favorite school. Schools view all new applications, which includes only those from students that rank the school k^{th} . Schools immediately accept students according to their priorities until they are full, rejecting all others.

The algorithm proceeds in this manner until all students are matched or all unmatched students have exhausted their list of acceptable schools.

The Boston mechanism is not strategy proof and may result in justified envy. As a result, the Boston mechanism has been replaced in many cases. However, it is Pareto efficient according to the reported preferences, so despite having some undesirable properties, it has its own proponents. Most often, it is the sophisticated parents who are able to take advantage of the algorithm's manipulability to improve their outcome. These parents argue that the ability to strategically manipulate their preferences is not a negative, but simply a reflection of the different priorities among parents. They also argue that allowing them to manipulate is the only thing parents can control in the process (Kojima and Unver, 2010). Because the outcome is Pareto efficient, changing to any other outcome will result in some students being worse off, resulting in resistance to changing the algorithm.

2.2. Model

I model a two-sided, many-to-one matching problem with incomplete information. Following the literature on school choice, I call one side of the market students and the other side schools and assume that schools, or more generally agents of the 'one' side of the market, do not act strategically. Let agents only have preferences about the characteristics of agents on the other side of the market. Let $\mathcal{S} = \{s_1, \dots, s_S\}$ be the set of all students and let $\mathcal{K} = \{k_1, \dots, k_K\}$ be the set of all schools. I make no assumption on the number of students, number of schools or the schools' capacities other than that they are finite.

In this paper, I model agents as having characteristics. Let there be a finite number of characteristics and let every characteristic have a finite number of possible states, such as grades being A through F or facilities being high, medium or low. Then I can represent each agent's characteristics as a finite-dimensional real Euclidean vector, where each state is assigned an integer value, with the initial point at the origin. For each student $i \in [1, S]$, I call such a vector δ_i^s . For each school $i \in [1, K]$, I call it δ_i^k . The set of all possible characteristic vectors for students is Δ^s and the set of all possible characteristic vectors for schools is Δ^k . A particular profile of characteristic vectors for a set of students and schools is $\delta = ((\delta_i^s)_{i \in \mathcal{S}}, (\delta_j^k)_{j \in \mathcal{K}})$.

I take preferences over characteristics as the primitive in agents' preferences. Students have preferences over the states of the characteristics of schools, which I call *desired traits*, represented by a finite dimensional real Euclidean vector ϕ_i^s with the initial point at the origin. Each value in this vector represents the most preferred state of that characteristic. Schools have *priorities* over the characteristics of students, represented by a finite dimensional real Euclidean vector ϕ_i^k with initial point at the origin. The set of all possible desired traits vectors for students is Φ^s . The set of all possible desired traits vectors for schools is Φ^k . A particular profile of desired traits and priority vectors for a set of students and schools is $\phi = ((\phi_i^s)_{i \in \mathcal{S}}, (\phi_j^k)_{j \in \mathcal{K}})$. For instance, a student only has preferences over the schools' class sizes, facilities, distance, etc., and not the schools themselves. Schools only have priorities over students' test scores, distance away, siblings, etc., and not the students themselves. Thus, students and schools are fully represented by the combination of characteristics they possess. This closely matches, in many cases, the way in which priorities are determined in practice. For example, in Boston, a student's priority is determined by two characteristics, having a sibling at the school and being in the school's walk zone. A student with both receives a top priority, followed by a student with a sibling, a student in a walk zone and a student with neither, with a lottery ordering students when more than one has the same priority (Chen and Sönmez, 2006).

I introduce this type of preference to the student side of the market as well. A student only prefers one school to another if she prefers the combination of characteristics the first school possesses to the combination of the second school. I assume that a tie-breaker always exists so that preferences and priorities are strict. Thus, though an agent can form a rank-order preference list, P_i^s or P_j^k belonging to the set of all possible preference lists P^s or P^k respectively, any ranking of schools or students must follow from the preferences

over characteristics. I assume that being matched is better for all agents than being unmatched, so that all schools are acceptable for each student and all students are acceptable for each school.

Let the characteristics of agents on both sides of the market and the priorities of schools be observable, or at least verifiable, by the mechanism designer. In this way, students and schools cannot lie about their own characteristics, nor can schools act strategically in reporting their priorities. This assumption serves to reduce the strategic complexity of the problem, but it also seems a very reasonable one in this context. A student may attempt to lie about their entrance exam score, but it is easy to imagine a school authority could observe this directly or verify it.

A School Choice with Characteristics mechanism (SCCM) is two-part. In the first part, the market designer elicits the agents' preferences over characteristics and generates a ranking of agents on the other side for each agent based on their responses. I call this ranking $L_i(\delta, \phi)$. Let $L(\delta, \phi)$ be the profile of all generated preference lists. The mechanism designer may elicit preferences over all characteristics, or some subset of them. Thus a student reports, in this case, preferences over the characteristics selected by the mechanism designer, answering the questions with which trait is most desired for each characteristic. If preferences over characteristics are the primitive, then this change, in effect, shifts the burden of forming the rank order list from the agents to the market designer. For simplicity, I assume the designer accurately generates preferences in the same way an agent would. This assumption can be considered a design objective instead, in which case the number and choice of desired traits for which the designer asks becomes a part of the designers' task as well, as is the case in practice. If this were so, the designer could draw on the large literature about school and student outcomes for assistance. In another paper, I provide some insight into the number of characteristics that could be needed (Sahajdack, 2015).

In the second part, the generated preference list L_i is used as an input to a traditional matching algorithm such as the Boston mechanism or SOSM, resulting in a matching. A *matching* is a function $\mu : \mathcal{S} \rightarrow K \cup \{\emptyset\}$, where the null element signifies that the agent is unmatched. The set of all possible matchings is \mathcal{M} . A student cannot occupy more than one seat and the set of students matched to a school must be less than or equal to its capacity. A matching algorithm is $\Omega : \mathcal{P} \rightarrow \mathcal{M}$ that takes preference lists as inputs and returns a matching. A School Choice with Characteristics mechanism is $\Psi : \{\delta, \phi\} \rightarrow \mathcal{M}$ that takes the characteristics, priorities and reported desired traits and delivers a matching. Because the mechanism has two distinct parts, it may also be written as $\Psi(\delta, \phi) = \Omega(L(\delta, \phi))$ where L is determined in the first step then used as an input in a matching algorithm.

To model incomplete information in this market, I assume that students are unaware of the existence of some schools. Because preferences are defined over characteristics, this is equivalent for students to know of a school but not of any of its characteristics. Suppose each student s_i can "discover" a fraction $d_i K$, $d_i \in [0, 1]$, of the schools and learn about their characteristics before the match. Call the set of discovered schools \mathcal{D}_i , with cardinality d_i . In this case, then each student s_i has a partial preference order $PP_i(\mathcal{D}_i)$ that ranks all schools he has discovered based on their characteristics, which he has now observed. I leave the method and

details of the discovery problem unspecified, but in practice it could occur in many ways: the schools closest to the student, the schools with a particular feature or the “best” schools if schools are known to be ordered by quality, among others.

2.3. Main Results

My results center on the use of eliciting preferences over characteristics instead of having agents report the ranked preference lists. I show in this section several results that give insight into the feasibility of switching from one method to the other under the SOSM and the Boston Mechanism. However, even before considering feasibility, it is important to identify the benefits that agents or market designers might achieve from such a switch to justify a costly change in mechanism. There are three main benefits to changing from preference lists to preferences over characteristics.

First, eliciting preferences over characteristics can allow for agents to avoid the often cumbersome task of learning about their options. In another paper, I show that when the matching algorithm is strategy proof, such as with the SOSM for students, it is also a weakly dominant strategy for agents to report their true preferences over characteristics (Sahajdack, 2015). In this case, students need no information about the schools to use their optimal strategy. In the matching literature, strategy proofness is considered a major benefit of the SOSM, as it is often seen as leveling the playing field for naive agents, those who cannot or do not manipulate their reported preference lists, and sophisticated agents, those who do (Pathak and Sönmez, 2008). However, this assumes agents have the same information. If sophisticated agents can invest time to discover a greater number of schools, for instance, they once more can benefit at the cost of those who cannot or do not do so as well. By eliciting preferences over characteristics, it is as if all agents were able to submit their true, complete preference list, removing the advantage of those who have more information.

Second, from a market designer’s perspective, it is preferable to have an agent’s full preference list. With the current method of eliciting preferences, this is the full ordered list of schools or students. This is especially true if those preferences are the agents’ true preferences. This allows the researcher to study accurate and complete data to improve the program, analyze counterfactuals and understand the state of the market from a participant’s perspective. Even if the preferences are not the true preferences, by having the complete preference lists, the designer can better attempt to identify the method by which agents manipulate. For example, they may be consistently reporting a particular desired trait that seems counter-intuitive. If the researcher only has preferences over a subset of schools, such as when agents report $PP_i(\mathcal{D}_i)$, it may not be possible to identify this precisely. As an example, suppose school A is both good and far away, while school B is bad but close. If student i reports his partial preference order as B followed by A, the market designer cannot distinguish if this order is a result of a preference for bad schools or close schools or even some other characteristic that makes B preferred to A even if good and far away are that student’s desired traits in a school, all else equal. If school C is good and close, and the student is required to also rank it,

the designer can at least determine, according to stated preferences, whether the student prefers a good or bad school. If the designer learns that students are reporting that they prefer bad schools, this information is important, as it seems counter-intuitive and may be indicative of manipulation and that the designer has left out some other key characteristic. In practice, in many applications of the Boston mechanism, school officials boast that the algorithm delivers the first choice school for a high percentage of students. However, this is often based on stated preferences over shortened lists of schools, and so is likely not as useful to the researcher to properly understand the market (Chen and Sönmez, 2006). By, in effect, forcing all agents to report complete preference lists, even if these lists are not truthful, the research gains information. Eliciting preferences over characteristics is one method to ensure agents report complete preference lists without burdening them unduly with researching every school to discover its characteristics.

Third, there can be many gains to efficiency in markets where there is incomplete information when using preferences over characteristics. The most obvious case, which is true with both the SOSM and Boston mechanism, is when students are unmatched in the mechanism and schools are below capacity due to those students being unaware of those schools. Because preferences over characteristics generates a ranking over all agents and uses it in the matching algorithm, these students and schools would be matched under preferences over characteristics. This is seen clearly in the NRMP, where unemployed doctors and hospitals with unfilled positions enter a secondary market after the initial match. It is reasonable to believe that this is a result of an unawareness of these positions, as, according to the NRMP website, many of the doctors do not list as many hospitals as they are allowed in the initial round, yet our matched in the secondary round. According to Gross, DeArmond and Denise (2015), there is also evidence that this occurs in school choice markets. After Denver and New Orleans school districts adopted a school choice mechanism, they began publishing a guide to the schools, listing their characteristics. As one parent stated, “I think more information is being put out there now about schools. Before it was more word of mouth.” Further, they found that among Denver students who only reported one school, about 30% of the total, 7% were unmatched. Of the 20% percent who ranked 5 schools, the maximum they were allowed to rank, only 3% were unmatched. Gross and coauthors propose that additional detailed information about the schools, even beyond what was presented in the guide, might help mitigate this issue. In addition to matching previously unmatched agents, mismatches resulting from incomplete information can be mitigated under preferences over characteristics. There may be cases under reporting partial preference lists where agents with a lower priority can gain admission to a school that a higher priority student prefers to his match. Though the SOSM would eliminate justified envy in the reported partial preference lists, this does not guarantee there will be no justified envy according to the true, complete preferences.

A major challenge for school authorities considering a change in their school choice mechanism is convincing parents, students and other administrators that the new system will be better. These challenges must be considered when analyzing the feasibility of implementing such a change. I consider changing from reporting partial preference lists to reporting preferences over characteristics when the matching algorithm

is both the student optimal deferred acceptance algorithm and the Boston mechanism. I first examine a change for only one student, with all other students' report unchanged.

DEFINITION 2.1. A report type is the elicited response of the participants, according to the design of the mechanism.

In a traditional matching framework, the standard report type is a rank order list of preferences of individual agents. In my model, there are two types of reports, partial preference ordering, $PP_i(\mathcal{D}_i)$, and desired traits, $\phi_i^{s,k}$. Other examples of report types commonly used in practice are truncated lists of preferences or preference lists over groups of agents.

DEFINITION 2.2. A school k_j is achievable for student s_i if there is a matching μ with $\mu(s_i) = k_j$ in which there is no justified envy.

THEOREM 2.1. *Under student-optimal deferred acceptance, holding the what is reported by other agents constant, any student weakly prefers the outcome when asked to report her desired traits vector ϕ_i^s to the outcome when asked to report her partial preference order $PP_i(\mathcal{D}_i)$.*

PROOF. First, note that because the SOSM is strategy-proof for students, she will report truthfully in either situation. There are now four cases to consider.

Case 1: The student is matched in both situations.

In this case, I will show that the student when reporting ϕ_i^s will receive at worst the same school she received when reporting $PP_i(\mathcal{D}_i)$. Let the school she received when reporting $PP_i(\mathcal{D}_i)$ be k . By the no justified envy property and strategy-proofness of SOSM, k must be achievable for student s_i and no school ranked above k in $PP_i(\mathcal{D}_i)$ can be achievable. Now consider $L_i(\delta, \phi_i^s)$. Because the order of the ranking in $PP_i(\mathcal{D}_i)$ is preserved in $L_i(\delta, \phi_i^s)$, any achievable school that was ranked lower than k in $PP_i(\mathcal{D}_i)$ will still be ranked lower and will never be the outcome. If any new achievable school is above k in $L_i(\delta, \phi_i^s)$, deferred acceptance will select that school as the match for s_i and she will be strictly better off. At worst however, her place at school k will still be available as in Case 2, so she will never receive an outcome worse than k .

Case 2: The student is unmatched when submitting $PP_i(\mathcal{D}_i)$ but matched when submitting ϕ_i^s .

Because any match is preferable to no match, she is strictly better off.

Case 3: The student is matched when submitting $PP_i(\mathcal{D}_i)$ but unmatched when submitting ϕ_i^s .

This cannot happen. Because the reports of other students are held constant, if student s_i was acceptable and had a place at school k_j , that place will still be open and she will still be acceptable. Since k_j is still acceptable to s_i to when submitting ϕ_i^s , at worst, she can always take that place and have an equivalent outcome.

Case 4: The student is unmatched in both situations.

The student receives the same outcome, no school, so is indifferent.

Thus in all cases, the student weakly prefers the outcome when asked to report ϕ_i^s .

□

Theorem 2.1 shows that an individual agent would rather report preferences over characteristics, all else equal. Intuitively, the SOSM does not punish students for expanding their reported preference list, and so the complete ranking generated by reporting desired traits can only help the student. One implication of this theorem is that it suggests one way of improving the outcome for parents and students who are not able to behave with sophistication. The designer could offer the choice between reporting a preference list or reporting desired traits. This theorem offers evidence that if any one student is offered the chance to change their report type, she will do so.

THEOREM 2.2. *Under the Boston Mechanism, holding the report type of other agents constant, there may be a student that prefers the outcome when asked to report her partial preference order $PP_i(\mathcal{D}_i)$ to the outcome when asked to report her preference vector ϕ_i^s .*

PROOF. Consider the following example to illustrate the theorem. There are three students, s_1 , s_2 and s_3 , and two schools, k_1 and k_2 with one seat each. Students do not know how many schools exist in the market. Both schools have the same priority over students, $s_1 > s_2 > s_3$. All students prefer k_1 to k_2 based on their characteristics. Students s_2 and s_3 only discover k_2 , while s_1 discovers both. Under reporting partial preference lists, in the first round, s_1 requests k_1 , while s_2 and s_3 request the seat at k_2 . The resulting match assigns s_1 to k_1 and s_2 to k_2 , due to his higher priority than s_3 .

Now suppose that s_2 is required to report preferences over characteristics. Having only discovered k_2 must decide whether to report honestly or lie. If he reports honestly, the algorithm means he will apply in the first round to k_1 , being rejected in favor of s_1 . He will then apply to k_2 in the second round, only to find that it has already been filled by s_3 , who applied in round one.

If he chooses to report dishonestly, he will never receive a seat at k_1 , since s_1 has priority and still applies. Thus, at best he may still receive k_2 , but there is a chance that he ends up unmatched or at a worse school he does not know about. Thus ex-post, s_2 weakly prefers the outcome under reporting partial preference lists.

□

Submitting $PP_i(\mathcal{D}_i)$ may leave out some schools at which the student did not have a high priority. As a result, the student no longer “wastes” a round making an offer to such a school. This avoids potentially losing a place at a school that is ranked lower but where he has a high priority and would be admitted, provided there is enough capacity. Expanding a student’s preference list is not always weakly better under the Boston Mechanism, thus some students may prefer the outcome from continuing to report a partial preference list.

Theorems 1 and 2 analyze the case when only one student at a time switches report type. Theorem 2.3 considers all students changing their report type simultaneously. This type of change is more realistic when

examining a proposed change in mechanism. Unfortunately, implementing such a change may be difficult due to the resulting consequences for some students.

THEOREM 2.3. *If all students simultaneously switch from being asked to submit $PP_i(\mathcal{D}_i)$ to being asked to submit ϕ_i^s under either Boston Mechanism or SOSM, at least one student may be strictly worse off.*

PROOF. Consider the following example to illustrate the theorem. Student s_i is the only student aware of school k_i , his top choice, which has a capacity of one seat. Suppose there exists a student s_j that has a higher priority at k_i . It is clear that under either mechanism s_i may lose his seat at k_i , making him strictly worse off.

□

Theorem 2.3 addresses whether switching from partial preference lists to preferences over characteristics always results in a Pareto improved outcome. The answer is no, as shown in the above example. Pareto improvement would be ideal from a feasibility of adoption perspective. Because the market designer cannot guarantee this, Theorem 2.3 suggests he may have a difficult time predicting the level of resistance to implementing the change in report type.

A further question in this line of thinking is whether the market designer can more accurately understand the likely outcome if a particular market scenario is considered. Corollary 2.1 imposes additional restrictions on the market. First, I suppose students outnumber schools. More specifically, I also assume that under reporting partial preference lists every school is at capacity. In other words, every seat at every school is filled. Finally, I impose the restriction that schools are ordered by quality for students. Though this restriction is stringent, it is not completely unrealistic. Many school districts find that students tend to prefer the same schools and that schools are often considered better or worse across most parents (Hastings et al., 2005, 2007). For instance, this restriction is very reasonable in a district where academic quality is highly valued among students, while less reasonable in districts where proximity is highly valued.

COROLLARY 2.1. *Suppose there is a simultaneous switch for all students from being asked to submit $PP_i(\mathcal{D}_i)$ to being asked to submit ϕ_i^s . If there are more students than schools, each school has no unfilled seats under reporting partial preference lists, and schools are ordered by quality for students, then the outcome under matching with characteristics will always be Pareto incomparable to the outcome under partial preference lists.*

PROOF. Because each school has no unfilled seats, in order for any student to change his outcome he must take the seat of another student. Thus when one student moves to another school, it must be that another student is displaced. Suppose, in the simplest case, that two students switch seats with each other with all others keeping the same seat. By assumption, all students agree on the desired traits of a school

in such a way that the schools are ordered in the same ranking in the true resulting preference list of all students. Thus it must be that if two students switch, one is made strictly better off and one is made strictly worse off and the two outcomes are Pareto incomparable.

Suppose instead that a chain of students are displaced, with each student improving her outcome when moving to a new school. Because there are a finite number of schools and seats, it must be that, at some point, a student will be displaced and receive a strictly worse outcome, making the outcomes under the two report types Pareto incomparable. In the extreme case, one can imagine a chain of students replacing each other and improving their match until finally a student at the top school is displaced. The displaced student from the top school must always be strictly worse off, so the two outcomes would again be Pareto incomparable. The same argument can be made for a chain of students displacing each other and receiving a worse match. The chain can continue, at its longest, until a student from the worst school is displaced. This student's new match must be strictly better than his previous match, making the outcome Pareto incomparable.

□

Corollary 2.1 demonstrates that when schools are at capacity, no student can improve without making another worse off. Thus, in this case, there is no possibility of a Pareto improvement. This will likely result in much resistance to any proposed change, as there will always be at least one student who is ex-post hurt by the switch. Corollary 2.1 does rule out the case where all students are made weakly worse off with at least one strictly worse off (a Pareto impairment), but this is hardly a strong platform for convincing students and parents to risk a worse outcome as a result of changing to preferences over characteristics.

2.4. Conclusion

The theorems presented in this paper suggest major challenges in implementing a change of preference elicitation method. However, this does not imply that such a switch is a hopeless cause. While a Pareto incomparable outcome will have opponents, it will also have proponents. Indeed, changing from the Boston mechanism to the SOSM results in Pareto incomparable outcomes, where some agents are made worse off, yet many districts have successfully transitioned to the student optimal deferred acceptance algorithm (Pathak, 2014).

As with most policy decisions, the benefits and costs must be weighed in light of the stated goals of the policy maker. In this case, the benefits of adopting a preferences over characteristics mechanism, at least if used with the the deferred acceptance algorithm, are improved market outcomes and a mechanism more closely in line with the stated goals of many school districts. A School Choice Mechanism with Characteristics can ease the informational burden on students and parents by leveraging the information advantage of the market designer. It provides the designer more complete data for continuing to improve the mechanism

in the future. It also improves the efficiency of the resulting outcome but eliminating failed matches and reducing, in some cases, mismatches due to incomplete information.

There is much work to be done in this area, especially investigating the discovery aspect of modeling incomplete information as I have. The particular details of this process and the market likely affect the magnitude of the gains in efficiency and this information is worthwhile to any market designer considering a change in mechanism. For example, suppose that all students discovered a common set of commonly preferred schools. In this case, while students with a high priority will be matched to the best schools, many students will be unmatched under reporting partial preference lists, even if there are many other schools they would be willing to attend. On the other hand, if students and schools match assortatively in an efficient market outcome, and students tend to discover schools near their own quality, the gains to efficiency are likely fairly small. Such details are important and of interest in further work.

Finally, while I have used the centralized school choice setting as a motivating example throughout, a matching mechanism using preferences over characteristics can have much broader impact. It can be used in many to one markets where both sides act strategically, such as college admissions or worker-firm matches. It has use in markets where anonymity is important, which present an impossible task for the traditional matching framework. It can also be used in one-to-one matching settings such as online dating where agents are too numerous to expect any one person to rank all others.

While eliciting preferences over characteristics is not a panacea in matching markets with incomplete information, it does provide market designers with a novel way to address some of the issues that arise in these markets. With careful design, its use can help program authorities to meet their design goals and improve market outcomes.

CHAPTER 3

Preferences that are Responsive over Characteristics and the Role of Information

3.1. Introduction

In this paper, I study the effects of exposing agents to information and how it may change their actions using survey data. The paper is motivated by the two-sided matching literature, in which agents often do not have full information about other agents, despite being asked to rank these agents from most preferred to least preferred. In earlier work, I study an alternative method of eliciting preferences of the agents: asking for their preferences over the characteristics of others, rather than asking them to submit a ranking of the others directly, and then generating the ranking based on the reported preferences. This approach has several advantages and disadvantages. A major advantage of this method is that it allows the market designer to leverage his or her private information, permitting the agents to use this information without having to discover it themselves. There are disadvantages to this approach as well, one of which is that the resulting preferences lists of the agents are restricted to a subset of all possible preference lists that I call *responsive over characteristics*. Given that the preferences over characteristics approach is closely tied to practical application, it is important to understand the effects of the information exposure and the restriction on preference lists. Using survey data with randomly assigned treatment and control groups, I find that there is some evidence that information exposure can affect the reported preferences over characteristics, but that this evidence is somewhat weak. I also find that the restriction on preference lists, at least within the survey design, seems to be meaningful and important.

Many two-sided matching markets use or could use preferences over characteristics in place of asking agents to submit rankings of the other agents directly. Some examples where this approach is currently used include online dating sites, schools in centralized public school choice and medical organ exchange programs ((Buckwalter et al., 2004; Abdulkadiroğlu et al., 2005)). There are further potential applications in other matching markets, such as the National Residency Matching Program, which matches doctors to hospitals, dorm roommate matching programs and students in centralized public school choice. In each of these markets, it is reasonable to believe that individual agents may not have full information about agents on the other side of the market and that the market designer has information about all the agents that any one individual does not.

To study the effect of sharing this information with the market participants, I collected survey data from more than 250 intermediate microeconomics students from the University of Illinois. I randomly assigned

each participant into one of three groups: a control group where participants were only asked questions about their preferences over the characteristics of a potential romantic partner; a treatment group in which the participants were shown the characteristics of sixteen potential partners and asked to rank all of these possible partners, then asked for their preferences over characteristics; and a final treatment group in which participants were shown and asked to rank only some, six out of sixteen, of the possible partners and then asked for their preferences over the characteristics. This design allows me to identify if information exposure plays a role in how the participants report their preferences over characteristics. My findings, though limited in scope, can provide some potentially useful suggestive evidence on the importance of information exposure. For instance, many online dating sites do not allow users to see the details of the profiles of potential dates before reporting their own characteristics and preferences over the characteristics of others. My survey design was inspired by the idea that the sites could instead allow users to view the full profiles of all or some potential dates before having to submit their own information.

I find some evidence that exposure to the characteristics of the potential partners affects the participants' reports of their preferences over characteristics. On two of the four characteristics, there is statistically significant evidence that the treatment groups answer the questions about which characteristic they prefer differently than the control. It is important to note that this evidence is certainly limited by the particular survey design, the choices of the characteristics and the subject pool. However, it is at least suggestive evidence that there are some characteristics for which the knowledge of the population may influence how agents report their preferences about that characteristic, as there is little reason to believe the groups are meaningfully different on a fundamental level.

The survey design also allows me to investigate, at least for the characteristics selected, if the agents in the treatment groups are reporting preferences that responsive over characteristics. Responsive preferences are a common assumption and restriction in two-sided, many-to-one matching problems. In informal language, preferences are considered responsive if a group of agents is preferred to another group with exactly the same makeup except for one changed individual if and only if the original individual is preferred to the changed individual. For instance, suppose a firm is deciding between hiring two groups of workers, based on its preferences for the two groups. Responsive preferences means that if the groups are the same except in the second group Person B is included instead of Person A, then the first group should only be preferred to the second group if Person A is preferred to Person B. In my design, instead of thinking of groups of agents, I define each agent as a group of characteristics. Because I then elicit preferences over the characteristics individually, it is similar to asking for preferences over individual agents when the final match will include a group of them. Thus an agent's preferences over characteristics are *responsive over characteristics* if Person A has the same characteristics as Person B except one, and Person A is only preferred to Person B if Person A's original characteristic is preferred to Person B's one changed characteristic.

By asking about each characteristic individually, the market designer is limited in the preferences he or she can generate for each agent. For example, suppose there are three men to be ranked for some woman.

Each man has two characteristics. The first man fulfills both of the desired characteristics of that woman, the second man fulfills only one, and the third fulfills none. When the market designer generates the preference list for that woman, he can never generate a ranking with the first man most preferred, the third man as the next preferred and the second man as the least preferred using her answers about those two characteristics. The class of rankings that satisfy responsiveness over characteristics is thus a subset of all possible rankings of the agents that a person could submit if they were allowed to submit a ranking directly. This restriction on rankings is important if it seems likely that agents do not, in reality, rank other agents in a responsive manner. In that case, the market designer would have to concede that the mechanism using preferences over characteristics does a poor job of representing the agents' true preferences over others. This is a difficult problem to address, as it is rare for the market designer to have data both on the agents' rankings and also their preferences over characteristics; most in-practice centralized matching programs only ask agents to report one or the other.

In the data collected in the survey, I find that the restriction on preferences has a meaningful impact on the ranking used for each agent. Of the 83 participants in the treatment group who ranked all sixteen potential partners, none submitted a ranking that matched the ranking generated by their submitted preferences over characteristics. More surprisingly, only three submitted a ranking in which the top agent they ranked matched the top agent in the generated preference list. This means that although the participant saw the four characteristics that made up their top choice agent, they did not select those characteristics in their preferences over characteristics. Similarly, among the treatment group that ranked only six potential partners, no participant ranked the six in the same order as they appeared in the generated preference list. There are many possible explanations for this finding, such as the preferences were indeed different due to some unknown factor, that ranking the partners directly by comparing their characteristics was too taxing on the participants or that the characteristics selected were not independent and the combination of two or more changed how the participant felt about the characteristics, violating responsiveness in that way. Nonetheless, the results show that care must be taken if a preferences over characteristics approach is used in practice, as the rankings generated may not be able to represent an agent's true ranking if the market designer is not diligent in designing the questionnaire.

I show using revealed preference data from the rankings submitted by treatment group one some evidence to explain why their likelihood of choosing some characteristics over others is different from the control group. I also use this data to explain why the agents' rankings are not responsive over characteristics. I find that there are complementarities between characteristics in the participants' choices, which, though surprising in the real world, highlights the importance of the market designer understanding his or her market.

The results in this paper offer some suggestive evidence for practitioners who use or are considering using a preferences over characteristics approach. While the responsiveness of the agents' reported preferences over characteristics and rankings are highly dependent on the particular features of the market and the effects

of information exposure certainly depend greatly on the situation, it is at least clear that these items can matter and should be considered and analyzed, when possible, to improve the market design.

3.2. Survey Design

The data gathered for this paper came from a survey offered to students taking intermediate microeconomics at the University of Illinois. The advantages and disadvantages to using student subjects is an extensively discussed topic in the social science experimental literature (see Peterson (2001) for a summary and Levitt and List (2007) for further discussion). In the case of this paper, the ready availability, ability to offer incentives for participation and number of likely respondents outweighed the disadvantages. With the known issues regarding this subject pool however, it is important to mention that any results must be carefully considered for external validity.

The participants were randomly assigned to one of three groups. In the control group, the subjects were asked to answer questions regarding their preferred characteristics of potential romantic partners as they might see on an online dating site. There were four questions, including, “Do you prefer a partner to be messy or clean?”, “Do you prefer a partner to be attractive or very attractive?”, “Do you prefer a partner to be loud or quiet?”, and “Do you prefer a partner to be logical or emotional?”. Each pair of choices can be considered two contrasting options for a characteristic, such as tidiness, attractiveness, extroversion and mental outlook, respectively. The participants were then asked to rank the four characteristics in order of importance, one, most important, to four, least important. An example is included in the appendix as Figure 3.

Using the characteristics shown above, sixteen unique combinations of characteristics can be made. Each combination then defines a potential romantic partner. No other identifying information was given for the hypothetical partners, they were solely and completely defined as a combination of characteristics, such as “messy, attractive, loud, logical.” Potential dates were defined in this way to attempt to minimize outside factors such as bias for or against a name or from seeing a picture. Treatment group one was shown all sixteen potential partners, each one as the combination of their characteristics listed inside a box, and asked to drag the boxes into a ranking of most preferred to least preferred. Participants in this group were then asked to answer the same questions as the control group about which option they preferred for each question and ranking the characteristics by importance. An example is included in the appendix as Figure 4.

The second treatment group was shown only six of the combinations of characteristics and asked to drag those six into a ranking from most preferred to least preferred. The participants were then asked the same questions as the control group about their preferences over characteristics. Along with the information about preferences, start time and time spent on each element of the survey was also collected. Out of 326 enrolled students, 267 responded to the survey invitation, with 89 in the control group, 91 in treatment group one and 87 in treatment group two. After data cleaning for those who did not attempt the survey itself, 83

students remained in each of the groups. Students were only allowed to attempt the survey once and so only appear in one group.

Each student in the class was offered a small incentive to participate. They were not required to complete the survey to receive the incentive, only to declare whether they wished to take part. Those who decided not to take part received the same incentive as those who did, and this was communicated to the subjects as part of the recruitment message and the survey consent. The incentive was extra credit toward the course, amounting to .4% of their course grade. The survey itself was separated from the survey consent so that a participants answers were both anonymous and confidential. Students were allowed ten days to complete the survey.

3.3. Results

Three main questions can be investigated using the data collected in the survey. The first is whether exposure to information about the characteristics of the potential partners creates significant differences in the responses from the three groups regarding their preferences over characteristics. The second is whether participants in the treatment groups submitted rankings that were responsive over characteristics and that matched their responses to the questions regarding preferences over characteristics. The final question is if there is a link between an observable factor and whether the submitted preference rankings were responsive over characteristics or not.

A summary of the responses of the overall sample population are shown in Table 2 and Table 3.

TABLE 2. Summary of Sample Population Proportions

	Proportion choosing first option
Messy or Clean	.0321
Attractive or Very Attractive	.4659
Loud or Quiet	.4458
Logical or Emotional	.8313

TABLE 3. Summary of Responses for Importance of Characteristics

	Proportions by Rank			
	Ranked 1st	Ranked 2nd	Ranked 3rd	Ranked 4th
Tidiness	0.1606	0.0723	0.6506	0.1165
Attractiveness	0.2932	0.1647	0.2048	0.3373
Extroversion	0.3133	0.2811	0.0964	0.3092
Mental Outlook	0.2329	0.4819	0.0482	0.2369

In order to investigate the link between the treatments and the responses to the questions, I tested the effect of each treatment using dummy variables in a logit regression for each question. I also included start

time and time spent on the survey as further controls, in case there was a meaningful difference between subjects relating to either of these dimensions.

Differences in Preferences over Characteristics by Treatment Group.

CLAIM 3.1. There is some evidence that the participants in the treatment groups answered differently to the control group about their preferences over characteristics.

Table 4 shows the results of the logit regressions for each question about which option the participant prefers.

TABLE 4. Logit Regressions: Log-odds of choosing each characteristic

Coefficients	M (1) or C (0)	A (1) or V (0)	L (1) or Q (0)	L (1) or E(0)
Intercept	3.9192**** (1.1123)	-0.8642*** (0.3090)	0.5023* (0.2888)	-1.6297**** (0.3675)
Treatment Group 1: Shown 16	-2.6453** (1.3119)	1.1842** (0.4797)	0.0599 (0.4547)	-0.0856 (0.5772)
Treatment Group 2: Shown 6	-0.4334 (1.2519)	1.1872**** (0.3380)	-0.2295 (0.3249)	-0.3286 (0.4442)
Time Spent	0.0081 (0.0066)	0.0024 (0.0026)	-0.0011 (0.0023)	0.0011 (0.0028)
Date Taken	-0.0255 (0.1489)	-0.0233 (0.0575)	-0.0738 (0.0553)	0.0258 (0.0711)

Notes: (1) Standard errors in parenthesis. (2) Significance at: '****' 0.1%; '***' 1%; '**' 5%; '*' 10%. (3) Column 1: Messy or Clean. Column 2: Attractive or Very Attractive. Column 3: Loud or Quiet. Column 4: Logical or Emotional

From the table, participants in treatment group one are less likely to choose that they prefer clean. Also, members of both treatment groups are more likely to choose someone very attractive. This second result is interesting in that it would seem, at first glance that having a partner who is very attractive would be preferable to one that is only attractive, all else equal. Perhaps surprisingly then, almost fifty percent of total participants selected that they would prefer an attractive partner to a very attractive one. Though I can only speculate on why this might be, one possible explanation is that the characteristics of being very attractive is often tied to many other characteristics for many people. If this were the case, choosing an attractive partner might be a substitute for choosing a partner who is less likely to be asked out by others or less likely to be arrogant. However, it may be that when the potential partners are revealed, the participants saw that they could choose someone who was both very attractive, and had other characteristics that they liked as well. If this is true, then revealing information to those participants helped them disentangle the attractiveness characteristic from others. There are other explanations, of course, but the results show a statistically significant change in behavior for the members of the treatment group. This is at least suggestive

evidence that information exposure can change the responses of agents in a preferences over characteristics framework.

Table 5 shows the multinomial logit regression for the characteristic selected as most important, using 'Extroversion' as the baseline. In Columns 2, 3 and 4, we see that the treatment groups have no significant effect on which characteristic was selected as most important. Similar results are found for the second, third and fourth most important characteristic.

TABLE 5. Multinomial Logit: Effect of treatment groups on most important characteristic

Coefficients	Characteristics		
	Tidiness	Mental Outlook	Attractiveness
Intercept	0.7721 (0.4418)	0.7610 (0.4264)	0.7324 (0.4506)
Treatment Group 1: Shown 16	0.1044 (0.6797)	-0.1296 (0.6586)	-0.1554 (0.7030)
Treatment Group 2: Shown 6	0.6609 (0.5295)	0.4399 (0.5241)	0.3189 (0.5507)
Time Spent	-0.0021 (0.0034)	-0.0004 (0.0032)	-0.00132 (0.0035)
Date Taken	-0.1072 (0.0800)	-0.0644 (0.0755)	-0.1545* (0.0901)

Notes: (1) Standard errors in parenthesis. (2) Significance at: ** 10%

Responsive over Characteristics Rankings.

The second question to be answered was whether participants selected rankings that were responsive over characteristics and if those rankings corresponded to the answers they submitted for their preferences over characteristics.

DEFINITION 3.1. A ranking is **responsive over characteristics** for agent i if for any set of characteristics C , and characteristics a and b such that $a \succ_i b$, then $C \cup a \succ_i C \cup b$.

Since each potential partner in the survey was defined solely as a combination of characteristics, then with responsive over characteristic rankings for any agent i , person A , defined as $C \cup a$, is preferred if person B , defined as $C \cup b$, if and only if $a \succ_i b$. For example, if person A is {messy, attractive, loud, logical} and person B is {clean, attractive, loud, logical} then $A \succ_i B$ if and only if {messy} \succ_i {clean}.

CLAIM 3.2. No participant among treatment group one chose a ranking that was responsive over characteristics or that matched the ranking generated by their answers about their preferences over characteristics.

Searching among all possible generated rankings, which are by construction responsive over characteristics, no agent of the 83 in treatment group one chose a ranking that could be generated by any combination of answers to the preferences over characteristics questions. As no agent selected any responsive over characteristic ranking, it is also clear that their ranking did not match their own individual generated ranking, which is one instance of such a ranking.

DEFINITION 3.2. A ranking is **partially responsive over characteristics** for agent i if that person ranks a subset of the possible partners, and the relative order of that subset can be found in a responsive over characteristics ranking.

Participants in treatment group two were asked to rank only six of the total sixteen potential partners. To satisfy partial responsiveness over characteristics, the relative order must be found in some responsive over characteristics ranking. This means that if some partner is preferred to another in the ranking of the subset, that partner must also be ranked better than the other in a responsive over characteristics ranking, though there is no restriction on how they must be ranked when compared to a partner not in the subset.

CLAIM 3.3. No participant among treatment group two chose a ranking that was partially responsive over characteristics or one for which the relative order of the partners in the subset was found in his or her generated ranking.

One explanation for Claim 2 is that ranking sixteen agents is too difficult for a participant to keep track of, given that each agent has four characteristics. If this were true, however, it might be expected that having to rank fewer agents would allow the participant to more accurately rank the subset according to his preferences over characteristics. This seems not to be the case. Although it is possible that six agents to be ranked is still simply too many for a participant to organize, it seems more likely that, in both cases, participants did not choose a responsive ranking for some other, unknown reason. This is especially evident in the fact that these participants are not merely failing to match their own generated ranking based on their answers to the questions about preferences over characteristics, but actually failing to come up with any responsive ranking at all.

Claims 4 and 5 analyze the issue of selecting a responsive ranking more deeply.

CLAIM 3.4. Among treatment group one, only three of 83 participants selected a partner for their most preferred partner that matched the most preferred partner of their generated ranking. Only nine selected a least preferred partner that matched the least preferred partner of their generated ranking.

CLAIM 3.5. Among treatment group two, only eleven of 83 participants selected a most preferred partner in their ranking of the subset that remained most preferred among the subset in their generated preferences.

These results are perhaps more surprising than Claims 2 and 3. Claim 4 states that characteristics of the partner selected as most preferred, who is made up only of those four characteristics that make it most preferred, are not the same four characteristics chosen as preferred when the participant is asked about their preferences over characteristics. Suppose an agent chooses a partner who is {clean, attractive, loud, logical}. This is all of the information known about that partner. It is curious that that person would not answer that they prefer a partner to be clean over messy, attractive over very attractive and so on. A similar process is occurring for the least preferred. Despite choosing someone whose combination of characteristics makes them least preferred, the characteristics themselves do not follow the same pattern. Claim 5 shares a similar theme: participants do not seem to pick their most preferred partner based on their preferences over characteristics. While an argument could be made that Claims 2-5 might be a result of mistakes or lack of attention from participants, the prevalence of the lack of agreement in the rankings suggests another cause. The results point to some sort of correlation in the preferences, where a combination of two or more characteristics makes that combination more or less desirable than the characteristics individually. It cannot be discounted that these results may also just be a particular outcome related to the survey design and choice of characteristics, but this is at least evidence that a market designer should be careful when choosing a preferences over characteristics approach; he or she may not be accurately representing the real rankings of the agents in the market.

Predictors of Responsiveness.

While a natural third question is which factors contribute to a participant being more likely to answer with a responsive or partially responsive over characteristics ranking, the lack of responsive observations make it difficult to perform any substantive analysis. By combining both all the positive observations, there are 22 unique observations where the participant selected the same most preferred, least preferred or most preferred among the subset partner as in his or her generated list. Note here that these are not 22 responsive rankings, merely 22 out of 166 that showed the barest evidence of agreement between the submitted and generated rankings. Using a logit, we find only one significant predictor, time spent on the survey. The coefficient of time spent of .0003 is significant at the ten percent level. While this is significant, and suggests that there is a very small positive effect on the odds of matching one of these three partners when the participant spends more time on the survey, the effect is so small and the observation count slightly dubious,

I hesitate to call this meaningful or to make any inference from this result. Instead, the question remains open in regard to which factors lead to a increase in the likelihood of a responsive submitted ranking.

Revealed Preferences in Treatment Group Exposed to all Partners.

In Table 4, I found significant results that members of treatment group one, who were shown all potential partners, had statistically different answers when asked to report preferences for tidiness and attractiveness. I found that they were more likely to choose that they preferred someone messy and someone very attractive than the control group. While it is impossible to know with certainty the reason for this difference, one possible explanation that might fit these two particular characteristics is that they might, in the real world, often be attached to many other characteristics that are not covered by the survey questions. For instance, it may be that participants perceive messy people as disorganized and unable to accomplish things effectively. Another explanation is that perhaps there is complementarity between characteristics within the survey. For instance, it could be that very attractive people tend to be outgoing and loud. Either of these would affect how participants reported their preferences.

These types of explanations may help explain the discrepancy between expectation and response. One might think from an observers perspective that, all else equal, almost all people should choose someone they find very attractive over someone who is only attractive to them. However, the survey responses reveal that only about 54% of all participants chose very attractive between the two and only 33% of the control group selected very attractive. Given that treatment group one showed significantly higher rates of selecting very attractive, one possible conclusion, among many, is that showing participants the full list of potential partners helps them separate the characteristics in their mind when they were then asked to choose between them. Another explanation could simply be that agents did not take the time to properly evaluate the agents they were ranking based on their characteristics.

As a test of these explanations and conclusions, I constructed revealed preference observations for each participant in treatment group one. For each person, I analyzed their submitted ranking of the partners and compared the rank of the two partners with the same three characteristics plus attractive and very attractive. For instance, an agent “chose” very attractive in an observation if the partner [messy, very attractive, loud, logical] was ranked above the partner [messy, attractive, loud, logical]. Based on this, I generated eight observations per participant, all the possible combinations of three other binary characteristics.

I found that while 68.7% of the treatment group selected very attractive directly, participants only chose very attractive in the revealed preference data 49.2% of the time, much closer to the rate of choosing very attractive in the other groups. The distribution of how many times each agents chose very attractive is presented in Table 6.

From this table, we see that individuals were rarely consistent in choosing very attractive over attractive in the revealed preference data. To test for complementarities between the characteristics within the survey,

TABLE 6. Distribution of Frequency of Choosing Very Attractive

Times Each Individual Chose VA	Frequency
0	1
1	4
2	6
3	18
4	23
5	22
6	8
7	1
8	0

I also tested if the presence of any particular characteristic made an individual more or less likely to choose rank the very attractive partner higher than the attractive one. These results are presented in Table 7.

TABLE 7. Probit: Relationship between characteristics of partner and choosing very attractive

Coefficients	Chose VA Estimates
Intercept	-0.0156 (0.2244)
Partner was Quiet	-0.2747*** (0.0987)
Partner was Emotional	-0.4112**** (0.0987)
Partner was Clean	0.1498 (0.0987)

(1) Standard errors in parenthesis. (2) Significance at: **** 0.1%; *** 1%; ** 5%; * 10%

Table 7 shows that there is evidence of complementarity between characteristics. Participants were more likely to rank the very attractive partner higher than the attractive partner if that partner was also loud and logical. The survey also revealed that after exposing participants to the potential partners, they selected that they preferred very attractive almost 70% of the time. This might provide evidence to match the intuition that people prefer someone very attractive to them, all else equal, and that this is a more accurate estimate of how often participants prefer very attractive partners, and that the revealed preference and data from the control group represent a preferences with complementarities between characteristics.

Of course, it is impossible to know for sure based solely on the data, but at least the evidence is clear that when ranking the agents, participants were not using responsive over characteristics thinking. This further confirms the findings in Claims 1 through 4, and provides some explanation for why agents were unable to submit responsive over characteristics rankings of partners.

3.4. Conclusion

Using a preferences over characteristics method of eliciting preferences, as opposed to having agents submit rankings directly, offers some advantages and disadvantages. One advantage is that the agents do not need to learn all the information about the other agents, the market designer can provide it or use it on their behalf. This additional information, however, may effect how agents report their preferences over characteristics if they have access to it. In this paper, I find that there is some evidence that participants in two randomly assigned treatment groups do alter their reported preferences over characteristics after being shown some or all of the potential partners on the other side of the market. Though only suggestive evidence at best when taken out of the context of the survey, it may be useful for practitioners to understand, such as for an online dating site that might reveal some or all of its member profiles before asking a new user about their characteristics and preferences over characteristics of others.

A major disadvantage of the preferences over characteristics approach is that it limits the rankings that can be used in a matching algorithm to the class of preference lists that is responsive over characteristics. While the extent of the impact of this restriction is unclear and highly dependent on the particular market, agents and mechanism, this paper finds that there is at least evidence that the restriction can be quite binding. I find that no agent of the 166 who might have chosen a ranking that was responsive or partially responsive over characteristics actually chose such a ranking. Indeed, almost all agents did not even choose a most preferred agent that matched the most preferred agent generated by their own responses to the questions about preferences over characteristics. I show some evidence to explain why rankings were not responsive over characteristics using revealed preference data from treatment group one. This suggests that in some contexts, agents have preferences that cannot be represented by a ranking generated from preferences over characteristics, or that participants in the survey had difficulties forming a connection between questions regarding the characteristics and a ranking that may result from their answers. This, too, is important for a practitioner to recognize and be aware of.

While these results are interesting in the context of the survey and when considering the preferences over characteristics approach, it is important to recognize the limitations of the data and the survey experiment. External validity is an important consideration here. I will only say that my findings offer suggestive evidence, at best, for the results I've found in regard to the wider question of the role of information exposure in preference formation and elicitation or the likelihood of agents having responsive over characteristics rankings of agents. Any market designer considering using a preferences over characteristics method of eliciting preferences or sharing information directly or indirectly with agents in the market should carefully consider the features of the market and be cognizant of the potential issues and impact that these actions may have.

This is not to say there is no place for a preferences over characteristics approach. As with any economics decision, the benefits must be weighed against the costs. Preferences over characteristics allows for agents to

communicate less information to the matching administrator and leverages the private information of that administrator. It is certainly possible that when ranking all agents is impossible, such as in extremely large markets like online dating sites, eliciting preferences over characteristics could still be a viable alternative.

Bibliography

- Abdulkadiroglu, Atila and Tayfun Sönmez**, “School choice: A mechanism design approach,” *The American Economic Review*, 2003, *93* (3), 729–747.
- Abdulkadiroğlu, Atila, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez**, “The Boston public school match,” *American Economic Review*, 2005, pp. 368–371.
- Azevedo, Eduardo M and Eric Budish**, “Strategy-proofness in the Large,” 2013.
- Betts, Julian R**, “Does school quality matter? Evidence from the National Longitudinal Survey of Youth,” *The Review of Economics and Statistics*, 1995, pp. 231–250.
- Buckwalter, J.G., S.R. Carter, G.T. Forgatch, T.D. Parsons, and N.C. Warren**, “Method and system for identifying people who are likely to have a successful relationship,” May 11 2004. US Patent 6,735,568.
- Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn**, “Constrained School Choice: An Experimental Study,” *The American Economic Review*, 2010, *100* (4), 1860–1874.
- Card, David and Alan B Krueger**, “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States,” *The Journal of Political Economy*, 1992, *100* (1), 1–40.
- Chen, Yajing**, “When is the Boston mechanism strategy-proof?,” *Mathematical Social Sciences*, 2014, *71*, 43–45.
- Chen, Yan and Tayfun Sönmez**, “School choice: an experimental study,” *Journal of Economic theory*, 2006, *127* (1), 202–231.
- Ehlers, Lars**, “Truncation strategies in matching markets,” *Mathematics of Operations Research*, 2008, *33* (2), 327–335.
- Gale, David and Lloyd S Shapley**, “College admissions and the stability of marriage,” *American Mathematical Monthly*, 1962, *69* (1), 9–15.
- and **Marilda Sotomayor**, “Some remarks on the stable matching problem,” *Discrete Applied Mathematics*, 1985, *11* (3), 223–232.
- Gonczarowski, Yannai A, Noam Nisan, Rafail Ostrovsky, and Will Rosenbaum**, “A stable marriage requires communication,” in “Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms” SIAM 2015, pp. 1003–1017.
- Gross, Betheny, Michael DeArmond, and Patrick Denice**, “Common Enrollment, Parents, and School Choice: Early Evidence from Denver and New Orleans. Making School Choice Work Series,” *Center on Reinventing Public Education*, 2015.
- Haeringer, Guillaume and Flip Klijn**, “Constrained school choice,” *Journal of Economic Theory*, 2009, *144* (5), 1921–1947.
- Hanushek, Eric A**, “The economics of schooling: Production and efficiency in public schools,” *Journal of economic literature*, 1986, *24* (3), 1141–1177.
- Hastings, Justine S, Jeffrey Weinstein, and Richard Van Weelden**, “Preferences, Information, and Parental Choice Behavior in Public School Choice,” *NBER Working Paper Series*, 2007, p. 12995.
- , **Thomas J Kane, and Douglas O Staiger**, “Parental preferences and school competition: Evidence from a public school choice program,” Technical Report, National Bureau of Economic Research 2005.
- Immorlica, Nicole and Mohammad Mahdian**, “Marriage, honesty, and stability,” in “Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms” Society for Industrial and Applied Mathematics 2005, pp. 53–62.
- Kojima, Fuhito and M Utku Unver**, “The Boston school choice mechanism,” Technical Report, working

- paper 2010.
- _____ and **Parag A. Pathak**, “Incentives and Stability in Large Two-Sided Matching Markets,” *The American Economic Review*, 2009, *99* (3), 608–627.
- Kumano, Taro**, “Strategy-proofness and stability of the Boston mechanism: An almost impossibility result,” *Journal of Public Economics*, 2013, *105*, 23–29.
- Levitt, Steven D and John A List**, “What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?,” *The Journal of Economic Perspectives*, 2007, *21* (2), 153–174.
- Oyer, Paul**, *Everything I Ever Needed to Know about Economics I Learned from Online Dating*, Harvard Business Review Press, 2014.
- Pathak, P.A. and T. Sonmez**, “School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation,” *American Economic Review*, 2013, *103* (1), 80–106.
- Pathak, Parag**, “School Assignment and School Effectiveness,” *NBER Reporter*, 2014, *2014* (1), 10–13.
- Pathak, Parag A and Tayfun Sönmez**, “Leveling the playing field: Sincere and sophisticated players in the Boston mechanism,” *The American Economic Review*, 2008, *98* (4), 1636–1652.
- _____ and _____, “School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation,” *American Economic Review*, 2013, *103* (1), 80–106.
- Peterson, Robert A**, “On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-analysis,” *Journal of Consumer Research*, 2001, *28* (3), 450–461.
- Pittel, Boris**, “The average number of stable matchings,” *SIAM Journal on Discrete Mathematics*, 1989, *2* (4), 530–549.
- _____, “On likely solutions of a stable marriage problem,” *The Annals of Applied Probability*, 1992, *2* (2), 358–401.
- Roth, Alvin E**, “The economics of matching: Stability and incentives,” *Mathematics of operations research*, 1982, *7* (4), 617–628.
- _____, “What Have We Learned from Market Design?,” *The Economic Journal*, 2008, *118* (527), 285–310.
- Roth, Alvin E. and Elliott Peranson**, “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design,” *The American Economic Review*, 1999, *89* (4), 748–780.
- Roth, Alvin E and Marilda A Oliveira Sotomayor**, *Two-sided matching: A study in game-theoretic modeling and analysis* number 18, Cambridge University Press, 1992.
- _____ and **Uriel G Rothblum**, “Truncation strategies in matching markets-in search of advice for participants,” *Econometrica*, 1999, *67* (1), 21–43.
- Sahajdack, Thomas**, “Incentives in a Matching with Characteristics: Can Designers Create Better Dates?,” 2015.
- _____, “Matching over Characteristics with Incomplete Information,” 2016.
- Segal, Ilya**, “The communication requirements of social choice rules and supporting budget sets,” *Journal of Economic Theory*, 2007, *136* (1), 341–378.
- Teo, Chung-Piaw, Jay Sethuraman, and Wee-Peng Tan**, “Gale-shapley stable marriage problem revisited: Strategic issues and applications,” *Management Science*, 2001, *47* (9), 1252–1267.

Appendix

Example 2.

Let four men have combinations of two characteristics: wealth and humor.

- Man 1: {Rich, Funny}
- Man 2: {Poor, Funny}
- Man 3: {Poor, Boring}
- Man 4: {Rich, Boring}

A woman is asked three questions:

- “Do you want a partner who is rich or poor?”
- “Do you want a partner who is funny or boring?”
- “Which is more important, wealth or humor?”

If poor and funny are chosen as the desired traits, an agent who is rich and boring can never be ranked above an agent who has one of the desired traits

Preference list	Acceptable	Answers needed
$2\{P,F\} > 1\{R,F\} > 3\{P,B\} > 4\{R,B\}$	yes	Poor, Funny, Wealth is more important than humor
$2\{P,F\} > 3\{P,B\} > 1\{R,F\} > 4\{R,B\}$	yes	Poor, Funny, Humor is more important than wealth
$2\{P,F\} > 4\{R,B\} > 3\{P,B\} > 1\{R,F\}$	no	Impossible

This example shows how the responsiveness in preferences created by the questions limits the preferences that can be represented once a particular set of characteristics and questions is chosen. A different three questions over different characteristics could allow for the last line of the table to be a possible generated preference list.

Example 3.

The mechanism in this example will be the one that simply maximizes total compatibility scores in the market. A couple score is calculated by taking the minimum of the two individual compatibility scores. The mechanism has many undesirable properties; there is no regard for stability and it is easy to show that the mechanism is not strategy-proof. Additionally, I will simplify the message space to twenty “yes or no” questions. An agent will answer these questions about himself and what he desires a partner to answer as in the OkCupid matching questions. Each time the partner’s answers match, she gains 5% in compatibility score, out of 100% total. There will be three men and three women. Suppose preferences are

$$m_1 : w_1 - 100\%; w_2 - 95\%; w_3 - 60\% \quad w_1 : m_1 - 95\%; m_2 - 90\%; m_3 - 85\%$$

$$m_2 : w_1 - 95\%; w_3 - 80\%; w_2 - 75\% \quad w_2 : m_2 - 100\%; m_3 - 85\%; m_1 - 75\%$$

$$m_3 : w_2 - 100\%; w_3 - 95\%; w_1 - 95\% \quad w_3 : m_2 - 90\%; m_3 - 85\%; m_1 - 80\%$$

Then the couple compatibility scores are

$$m_1w_1 : 95 \quad m_2w_1 : 90 \quad m_3w_1 : 85$$

$$m_1w_2 : 75 \quad m_2w_2 : 75 \quad m_3w_2 : 85$$

$$m_1w_3 : 60 \quad m_2w_3 : 80 \quad m_3w_3 : 85$$

If all agents report honestly, the mechanism will select $\mu = [(m_1w_1), (m_2w_3), (m_3w_2)]$, for a maximized total couple compatibility of $95+80+85=260$. Let us consider another matching, μ' , where $\mu' = [(m_1w_1), (m_2w_2), (m_3w_3)]$. Here the total is lower, $95+75+85=255$, but it is also clear that w_2 prefers μ' , receiving her 100% individually compatible partner, m_2 , as opposed to m_1 at 85% in μ . At this point, we can observe that w_2 would like to manipulate her preferences, if possible, to improve her own personal payoff by changing the outcome to μ' .

However, consider one last possible matching, μ'' , where $\mu'' = [(m_1w_2), (m_2w_1), (m_3w_3)]$. This match has a total score of $75+90+85=250$ and will certainly never be selected by the mechanism if agents report honestly. We can also note that w_2 prefers μ'' the least, receiving her lowest compatible partner, m_1 at 75%. It is the existence of this matching that blocks w_2 's ability to manipulate the market.

Consider w_2 's strategies if she wishes to manipulate the outcome of the mechanism by lying about her desired traits. She knows she must lower the total compatibility of μ by at least 10 points for the mechanism to select μ' instead. Therefore, she must lie twice in her desired traits to drop her reported compatibility with m_3 to 75% instead of 85%. However, because agents cannot directly manipulate their preference lists, lying about two questions must lower her reported compatibility with m_2 from 100% to 90%. Furthermore, these two lies might increase or decrease her reported score with m_1 . For this example, supposed the two lies caused the reported score to increase from 75% to 85%.

As the couple compatibility scores are calculated with the minimum of the two, this changes the couple compatibilities of the pair (m_1, w_2) to 85% and (m_3, w_2) to 75%. All others remain the same, including (m_2, w_2) as that score is determined by m_2 's low personal compatibility with w_2 . With these new couple compatibility scores, based on the manipulated preferences of w_2 , we can recalculate the totals and find the

new outcome of the mechanism. For match μ , the new total is 250. For match μ' , the new total is 255. Thus, w_2 has succeeded in blocking the match μ , as she had hoped.

However, due to the nature of the alternative message space, she cannot simply adjust her reported compatibility with one other man. Because her two lies have changed the reports about all three men, we must also calculate the score of μ'' , which is now 260 under her false report. The mechanism will select μ'' , resulting in w_2 receiving her least desirable partner under her true preferences. In this market, the restriction on the alternative message space prohibits this agent from improving her outcome through manipulation. A similar argument can be made for m_2 , who also is prevented from improving by deviating in the same manner. As all other agents are receiving their top choices, they, too, have no incentive to deviate. As no agent can improve by deviating, honest reporting is a Nash equilibrium in this market.

It is worth noting that this Nash equilibrium cannot be sustained under direct reporting. In that case, w_2 would simply change her reported compatibility with m_3 , leaving the others unchanged. This would allow her to block the first matching and have the second selected. Here, she clearly has incentive to deviate, receiving her top choice by doing so, instead of her second choice.

Proof of Theorem 1.5.

PROOF. I first show that the set of generated preference lists is strictly contained in the set of direct reporting preferences, $L \subset P$.

This follows from the definition of the two sets.

First, $L \subseteq P$. L contains at most all permutations of a ranking containing all agents to be ranked. P contains all permutations of any subset of agents to be ranked, including the case when all agents are ranked.

Second, $L \not\supseteq P$. By construction, any item in L must rank all agents. Take any element P_i of P such that the length of that preference list, or the number of agents it ranks, is less than n . Such an element $P_i \notin L$.

A manipulation in the case of the alternative message space is a report leading to an $L'_i(r'_i) \neq L_i(r_i)$. Since any $L'_i \in P$, choose $P'_i = L'_i$.

Thus any manipulation in the alternative message space can be replicated in a direct reporting environment.

A manipulation in the case of direct reporting is any report $P'_i \neq P_i$. Choose a P'_i such that it is a truncation of P_i , and as a result $P'_i \notin L$.

Thus the reverse is not true.

□

FIGURE 3. Survey Questionnaire

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

How would you rank these potential dates?
The following boxes represent 6 potential dates. Each box lists four characteristics that define that person. The dates are identical otherwise. Please rank them in order from the one you like the most (1) to the one you like the least (6). You'll have to rank all 5 people to advance to the next question. You can drag and drop or double click the options to rank them.

loud, messy, logical, attractive	quiet, messy, emotional, attractive	Most desirable 2 3 4 5 Least Desirable
quiet, clean, emotional, attractive	quiet, clean, logical, very attractive	
loud, clean, logical, very attractive	loud, messy, emotional, very attractive	

Which trait is most important to you?
Please rank the traits based on how important it is to you that your partner matches your preference. The traits fall into these categories:
Extroversion – Loud or Quiet
Tidiness – Messy or Clean
Mental Outlook – Logical or Emotional
Attractiveness – Attractive or Very Attractive
For instance, if you prefer someone loud and it is very important that a partner meets this criterion, you would rank Extroversion towards the top of your list. If you prefer someone messy, but this really doesn't matter to you very much either way, you would rank "Tidiness" towards the bottom of your list.

Extroversion	Tidiness	More Important 2 3 Less Important
Mental Outlook	Attractiveness	

Next

Thomas Sahajdack, University of Illinois – 2016 60% completed

FIGURE 4. Ranking All Sixteen Agents



How would you rank these potential dates?

The following boxes represent 16 potential dates. Each box lists four characteristics that define that person. The dates are identical otherwise. Please rank them in order from the one you like the most (1) to the one you like the least (16). You'll have to rank all 16 people to advance to the next question. You can drag and drop or double click the options to rank them.

quiet, clean, logical, very attractive	quiet, messy, logical, attractive	Most desirable
quiet, messy, emotional, very attractive	loud, clean, logical, very attractive	2
quiet, clean, emotional, very attractive	quiet, clean, logical, attractive	3
loud, clean, emotional, attractive	quiet, clean, emotional, attractive	4
loud, clean, emotional, very attractive	quiet, messy, emotional, attractive	5
loud, clean, logical, attractive	loud, messy, emotional, attractive	6
loud, messy, logical, attractive	quiet, messy, logical, very attractive	7
loud, messy, emotional, very attractive	loud, messy, logical, very attractive	8
		9
		10
		11
		12
		13
		14
		15
		Least desirable

FIGURE 5. Human Subjects Approval

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research

Office for the Protection of Research Subjects
528 East Green Street
Suite 203
Champaign, IL 61820



April 21, 2016

Thomas Sahajdack
Department of Economics
214 David Kinley Hall
1408 Gregory
Urbana, IL 61801

RE: *Information Exposure and Preference Structure in Matching over Characteristics*
IRB Protocol Number: 16783

Dear Dr. Sahajdack:

This letter authorizes the use of human subjects in your project entitled *Information Exposure and Preference Structure in Matching over Characteristics*. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) approved, by expedited review, the protocol as described in your IRB application. The expiration date for this protocol, IRB number 16783, is 04/20/2017. The risk designation applied to your project is *no more than minimal risk*.

Copies of the attached date-stamped consent form(s) must be used in obtaining informed consent. If there is a need to revise or alter the consent form(s), please submit the revised form(s) for IRB review, approval, and date-stamping prior to use.

Under applicable regulations, no changes to procedures involving human subjects may be made without prior IRB review and approval. The regulations also require that you promptly notify the IRB of any problems involving human subjects, including unanticipated side effects, adverse reactions, and any injuries or complications that arise during the project.

If you have any questions about the IRB process, or if you need assistance at any time, please feel free to contact me at the OPRS office, or visit our Web site at <http://oprs.research.illinois.edu>.

Sincerely,

Dustin Yocum, MA, CIP
Human Subjects Research Specialist, Office for the Protection of Research Subjects

Attachment(s): On-line informed consent document and Waiver of Documentation of Informed Consent