Matching with Couples: Stability and Incentives in Large Markets

MATCHING WITH COUPLES: STABILITY AND INCENTIVES IN LARGE MARKETS*

FUHITO KOJIMA
PARAG A. PATHAK
ALVIN E. ROTH

Accommodating couples has been a long-standing issue in the design of centralized labor market clearinghouses for doctors and psychologists, because couples view pairs of jobs as complements. A stable matching may not exist when couples are present. This article's main result is that a stable matching exists when there are relatively few couples and preference lists are sufficiently short relative to market size. We also discuss incentives in markets with couples. We relate these theoretical results to the job market for psychologists, in which stable matchings exist for all years of the data, despite the presence of couples. *JEL* Codes: C78, D47.

I. INTRODUCTION

One of the big twentieth-century transformations of the American labor market involves the increased labor force participation of married women, and the consequent growth in the number of two-career households.[1] When a couple needs two jobs, they face a hard problem of coordination with each other and with their prospective employers. The search and matching process for spouses can involve very different timing of searches and hiring. A couple may be forced to make a decision on a job offer for one member of the couple before knowing what complementary jobs may become available for the other or what better pairs of jobs might become available elsewhere.

*We thank Jin Chen, Rezwan Haque, Fanqi Shi, and especially Dan Barron, Pete Troyan, and Mengxi Wu for superb research assistance. We are also grateful to Delong Meng, Assaf Romm, Ran Shorrer, Joel Sobel, and seminar participants at Calgary, Ecole Polytechnique, Washington University in St. Louis, the ERID Matching Conference "Roth and Sotomayor: Twenty Years After" at Duke University, the Coalition Theory Workshop in Marseilles, the EC conference in San Jose, the 2011 AMES pre-conference at Hanyang, and the NBER Market Design conference for comments. Elliott Peranson and Greg Keilin provided invaluable assistance in obtaining and answering questions about the data used in this article on behalf of National Matching Services and the Association of Psychology Postdoctoral and Internship Centers. The National Science Foundation provided research support.

1. See, for instance, Costa and Kahn (2000) for a description of the trends in the labor market choices for college-educated couples since World War II.

The Quarterly Journal of Economics (2013), 1–48. doi:10.1093/qje/qjt019.

An unusually clear view of this problem can be found in the history of the entry-level labor market for American doctors. Since the early 1900s, new U.S. medical graduates have been first employed as "residents" at hospitals, where they work under the supervision of more senior, licensed doctors. This market experienced serious problems having to do with the timing of offers and acceptances, and this unraveling of the market led to the creation of a centralized clearinghouse in the 1950s that drew high rates of participation (see Roth 1984, 2003; Roth and Xing 1994 for further details). Medical graduates were almost all men throughout this period, but by the 1970s there were enough women graduating from medical school so that it was not unheard of for two new medical graduates to be married to each other.[2] Many couples felt that the existing clearinghouse did not serve them well, and starting in the 1970s, significant numbers of these couples began seeking jobs outside of the clearinghouse.

Roth (1984) argues that this was because the matching algorithm used until then did not allow couples to appropriately express preferences. That paper shows that, in a market without couples, the 1950s clearinghouse algorithm is equivalent to the deferred acceptance algorithm of Gale and Shapley (1962), and that it produces a stable matching for any reported preferences—loosely speaking, this is a matching such that there is no pair of hospital and doctor who want to be matched with each other rather than accepting the prescribed matching.[3] It then observes that the algorithm often fails to find a stable matching when there are couples and argues that a main problem of the mechanism is that (prior to the 1983 match) it did not allow couples to report preferences over pairs of positions, one for each member of the couple. Roth and Peranson (1999) describe the current algorithm, which elicits and uses couples' preferences over pairs of positions and has been used by more than 40 centralized clearinghouses including the American labor market for new doctors, the National Resident Matching Program (NRMP).[4]

2. In the 1967–68 academic year, 8% of the graduates of U.S. medical schools were women. By 1977–78 this fraction had risen to 21%, and by 2008–9 to 49% (Jonas and Etzel 1998; and see http://www.aamc.org/data/facts/charts 1982to2010.pdf).

3. Section III provides a precise definition of our stability concept.

4. See Roth (2007) for a list of these clearinghouses as well as a survey of the literature. See also Sönmez and Ünver (2009).

But the problem is difficult even if couples are allowed to express their preferences over pairs of positions, because there does not necessarily exist a stable matching in markets with couples (Roth 1984). However, some matching clearinghouses regularly entertain high rates of participation and produce matchings that are honored by participants. In fact, it has been reported that there have only been a few occasions in which a stable matching was not found over the past decade in several dozen annual markets (E. Peranson, private communication). Moreover, in the largest of these markets, the NRMP, Roth and Peranson (1999) run a number of matching algorithms using submitted preferences from 1993, 1994, and 1995 and find no instance in which any of these algorithms failed to produce a stable matching. Why do these matching clearinghouses produce stable outcomes from submitted preferences even though existing theory suggests that stable matchings may not exist when couples are present?

This is the puzzle we address, and this article argues that the answer may have to do with the size of the market. We consider a sequence of markets indexed by the number of hospitals. Doctors have preferences, drawn from a distribution over a number of hospitals, and this number is allowed to grow as the total number of hospitals grows, but more slowly. These are preferences after interviews have taken place.[5] When the number of couples grows sufficiently slowly relative to market size, under some regularity conditions, our main result demonstrates that the probability that a stable matching exists converges to 1 as the market size approaches infinity. Moreover, we provide an algorithm that finds a stable matching with a probability that converges to 1 as the market size approaches infinity. We also discuss incentives for doctors and hospitals to report their preferences truthfully to the clearinghouse in markets with couples.

Because our theoretical analysis only provides limit results, we study data on submitted preferences from the centralized

5. The length of doctors preference lists can be expected to grow more slowly than the number of hospitals because other elements of the environment that limit the number of interviews are not also growing (e.g., the number of days between November and February and the time required for interviews remain constant). We do not model frictions involved in interviewing at different hospitals in this article. A model with explicit costs of interviewing would be substantially different than the current framework.

market for clinical psychologists. In the late 1990s, the market evolved from a decentralized one (Roth and Xing 1997) to one employing a centralized clearinghouse, where a key design issue was whether it would be possible to accommodate the presence of couples. Keilin (1998) reports that under the old decentralized system, couples had difficulties coordinating their internship choices. In 1999, clinical psychologists adopted a centralized clearinghouse using an algorithm based on Roth and Peranson (1999), in which couples are allowed to express preferences over hospital pairs. We explore a variation of the Roth-Peranson procedure to investigate the existence of a stable matching for years 1999–2007. Using our algorithm, we are able to find a stable matching with respect to the stated preferences of participants in all nine years. To investigate whether our asymptotic arguments provide a guide for realistic market sizes, we also simulate markets under different assumptions on the number of couples. We draw preferences for agents from a uniform distribution and from a distribution calibrated using data from the clinical psychology market. Our simulations show that a stable matching is more likely to be found in large markets. For example, if the market size, that is, the number of hospital seats, is at least 2,000 and the number of couples is equal to the square root of market size, a stable matching exists at least 96% of the time.

I.A. *Related Literature*

This article is related to several lines of work. First, it is part of research in two-sided matching with couples. Existing studies are mostly negative: Roth (1984) and unpublished work by Sotomayor show that a stable matching does not necessarily exist when there are couples, and Ronn (1990) shows that it may be computationally hard to determine if a stable matching even exists. Klaus and Klijn (2005) provide a maximal domain of couple preferences that guarantees the existence of stable matchings.[6] Although their preference domain has a natural interpretation, our article finds that the submitted preferences of almost

6. Related contributions include Dutta and Massó (1997) and Dean, Goemans, and Immorlica (2006), who investigate the properties of matching models with alternative formulations of couples.

all couples in our psychology market data violate their condition.[7] This empirical fact motivates our appeal to large market arguments.

The second line of studies related to this article is the growing literature on large matching markets. Roth and Peranson (1999) conduct simulations based on data from the NRMP, which include couples and randomly generated data. They find suggestive evidence that in large markets, a stable matching is likely to exist and stable matching mechanisms are difficult to manipulate. We also examine data from the psychologist market and demonstrate that a stable matching exists for all years we have access to data. One of the findings of Roth and Peranson (1999) is that the opportunities for manipulation vanish in large markets if doctor preference lists are bounded as market size grows, but such a result does not hold if each doctor lists all hospitals in her preference list. Moreover, in the market for clinical psychologists and the NRMP, each applicant can physically interview at a small number of potential employers. These two considerations motivate our theoretical analysis of a model in which doctor preference lists are allowed to grow, but in a controlled manner.

Several recent papers have studied incentive issues in matching models with a large number of agents, following the setup and techniques in Immorlica and Mahdian (2005) who consider a one-to-one matching model and establish that hospitals are unlikely to be able to manipulate the doctor-optimal stable mechanism in a large market.[8] Kojima and Pathak (2009) build on and extend this analysis for a many-to-one market. Although the use of large market arguments in our article is similar to these studies, the questions are substantially different. A stable matching always exists in the model without couples, whereas it

7. Preference restrictions are also used to study incentives for manipulation in matching markets. The restriction under which incentive compatibility can be established is often very strong, as shown by Alcalde and Barberà (1994), Kesten (2012), Konishi and Ünver (2006b), and Kojima (2007) for various kinds of manipulations. Similarly, in the context of resource allocation and school choice (Abdulkadiroğlu and Sönmez 2003), necessary and sufficient conditions for desirable properties such as efficiency and incentive compatibility are strict (Ergin 2002; Kesten 2006; Haeringer and Klijn 2009; Ehlers and Erdil 2010). Another approach is based on incomplete information (Roth and Rothblum 1999; Ehlers 2004, 2008; Erdil and Ergin, 2008; Kesten 2010).

8. See also Knuth, Motwani, and Pittel (1990) on the probabilistic analysis of large matching markets.

is not guaranteed to exist when couples are present. This article's use of large market arguments to establish the existence of a stable matching is, as far as we know, new in the matching literature. Our analysis also relaxes some commonly used assumptions in the analysis of large matching markets.

Large market arguments have been used in a number of other recent studies of matching mechanisms (Bulow and Levin 2006; Abdulkadiroğlu, Che, and Yasuda 2008; Manea 2009; Che and Kojima 2010; Kojima and Manea 2010; Budish 2011; Liu and Pycia 2011; Azevedo and Leshno 2012; Lee 2013). We describe subsequent work by Ashlagi, Braverman, and Hassidim (2011) in Section VI. Although the analysis of large markets is relatively new in the matching literature, it has a long tradition in economics. For example, Roberts and Postlewaite (1976) and Jackson and Manelli (1997) show that, under some conditions, the Walrasian mechanism is difficult to manipulate in large exchange economies. Similarly, Gresik and Satterthwaite (1989) and Rustichini, Satterthwaite, and Williams (1994) study incentive properties of a large class of double auction mechanisms.

Finally, a couple preference is a particular form of complementarity, and this article can be put in the context of the larger research program on the role of complementarities in resource allocation. Complementarities have been identified to cause nonexistence of desirable solutions in various contexts of resource allocation. There has been a recent flurry of investigations on complementarities and existence problems in auction markets (Gul and Stacchetti 2000; Milgrom 2004; Sun and Yang 2009), general equilibrium with indivisible goods (Gul and Stacchetti 1999; Bikhchandani and Ostroy 2002; Sun and Yang 2006), and matching markets (Ostrovsky 2008; Hatfield and Kominers 2009; Sönmez and Ünver 2010; Pycia 2012).

The organization of this article is as follows. The next section describes some features of the market for clinical psychologists and lays out a series of stylized facts on matching with couples based on data from this market. Section III defines the model and describes a simple theory of matching with couples in a finite market. Section IV introduces the large market assumptions. Section V states our main results on existence, and Section VI discusses incentives, robustness, and extensions of our results. Section VII concludes.

II. The Market for Internships in Professional Psychology

II.A. Background

The story of how design has been influenced by the presence of couples in the NRMP has parallels in the evolution of the market for internships in professional psychology.[9] Roth and Xing (1997) described this market through the early 1990s. From the 1970s through the late 1990s, this market operated in a decentralized fashion (with frequent rule changes), based on a "uniform notification day" system in which offers were given to internship applicants over the telephone within a specific time frame (e.g., a four-hour period on the second Monday in February). All acceptances and rejections of offers occurred during this period. Keilin (2000, 281) described the system as "problematic, subject to bottlenecks and gridlock, encouraging the violation of guidelines, and resulting in less-than-desirable outcomes for participants."

In 1998–99, the Association of Psychology Postdoctoral and Internship Centers (APPIC) switched to a system in which applicants and internship sites were matched by computer. A major debate in this decision was whether a centralized system could handle the presence of couples. In the old, decentralized scheme it was challenging for couples to coordinate their internship choices. Keilin (1998) reports that one partner could be put in the position of having to make an immediate decision about an offer without knowing the status of the other partner. Following the reforms of the NRMP, a new scheme that allowed couples to jointly express their preferences was adopted.

With the permission of the APPIC, the organization that runs the matching process, National Matching Services, provided us with an anonymized data set of the stated rank order lists of single doctors, couples, and hospitals and hospital capacities for the first nine years of the centralized system. Because of privacy concerns, the APPIC data set does not include any demographic information on applicants and only limited identifying information on programs.

9. To be clear, we concentrate on the match run by the APPIC for predoctoral internships in psychology, which involves clinical, counseling, and school psychologists. (This is distinct from the postdoctoral match in neuropsychology.)

II.B. Stylized Facts

This section identifies some stylized facts from the internship market for professional psychologists. Although we do not have as detailed data from the NRMP, we also mention related facts from that market when appropriate using information from the NRMP's annual reports.

The data are the stated preferences of market participants, so their interpretation may require some caution. There are at least two parts to the process by which market participants form their preferences: (1) they determine which applicants or internship programs may be attractive, and participate in interviews; and (2) after interviewing, they determine their rank ordering over the applicants or internship programs they have interviewed. The model and the data do not allow us to say much about the first stage of the application process. In determining where to interview, applicants probably factor in the costs of traveling to interviews, the program's reputation, and a host of other factors. Programs consider, among other things, the applicant's recommendation letters and suitability for their program in deciding whom to interview. Once market participants have learned about each other, they must come up with their rank ordering. For the empirical analysis in this section, we abstract away from the initial phase of mutual decisions of whom to interview, and our interpretation is that the data reflect the preferences formed after interviews. This, and the fact that participants only seem to rank those with whom they have had interviews, likely accounts for the relatively short rank order lists.

Even with this interpretation, the reported post-interview preferences may be manipulations of the true post-interview preferences of market participants because truth-telling is not a dominant strategy for all market participants. However, there are at least two reasons that treating submitted preferences as true preferences may not be an unrealistic approximation. First, as noted in Section VI.A, the organizers of the APPIC match emphasize repeatedly that market participants should declare their preferences truthfully. Second, in our working paper (Kojima, Pathak, and Roth 2010) we provide assumptions under which truth-telling is an approximate equilibrium in large markets.

Table I presents some summary statistics on the market. On average, per year, there are 3,010 single applicants and 19 pairs of applicants who participate as couples. In early years, there

TABLE I
Summary Statistics for Market for Clinical Psychologists

	Total	Mean	Min	25th	Median	75th	Max
Panel A. Length of rank order list (ROL)							
Single applicants	3,010	7.6	1.0	4.0	7.1	10.4	73.1
Couples	19	81.2	7.3	29.4	52.3	115.0	249.9
Distinct programs ranked		10.2	2.0	6.4	9.9	13.0	20.9
Programs	1,094	16.7	1.0	7.6	14.3	23.8	80.9
Panel B: Program capacities							
Capacity	2,721	2.5	1.0	1.0	2.0	3.0	21.4
Panel C: Geographic similarity of preferences							
Single applicants							
# Regions ranked		2.5	1.0	1.0	2.0	3.1	9.3
Couples							
# Regions ranked		3.9	1.0	2.5	3.9	4.9	6.9
Fraction of ROL where both members rank same region		73.4%	29.7%	47.1%	78.2%	99.3%	100.0%

Notes. This table reports descriptive information from the Association of Psychology Postdoctoral and Internship Centers match for 1999–2007. Single applicants' rank order lists consist of a ranking over hospitals, whereas couples indicate rankings over program pairs. Distinct programs ranked are the set of distinct programs ranked by each couple member. Programs include only those that have positive capacity. There are 11 regions, corresponding to the first digit of U.S. ZIP codes and Canada.

were just under 3,000 applicants, but the number of applicants has increased slightly in the most recent years. The number of applicants who participate as couples has remained relatively small, varying between 28 and 44 (i.e., between 14 and 22 couples) which is about 1% of all applicants.[10] On the surface, the small number of couples may appear surprising, but this number represents cases where both couple members look for jobs in professional psychology in the same year. If couple members wish to work in different fields, or even in the same field but in different years, each couple member simply applies as a single applicant. In the NRMP, from 1992 to 2009, on average 4.4% of applicants participated as couples, with slightly more couples participating in the most recent years (NRMP 2009).

Fact 1. Applicants who participate as couples constitute a small fraction of all participating applicants.

10. As the example in the next section shows, even one couple in the market may lead to nonexistence of a stable matching.

Panel A of Table 1 shows the length of the rank order lists for applicants and programs. On average across years, single applicants rank between seven and eight programs. Because there are 1,094 programs on average, this means that the typical applicant ranks less than 1% of all possible programs. Even at the extreme, the length of the longest single applicant's rank order list is about 6.7% of all possible programs. In the NRMP, the length of the applicant preference list is about seven to nine programs, which would be roughly 0.3% of all possible programs.[11] This may not be surprising because an applicant typically ranks a program only after she interviews at the program, and each applicant receives and can travel to only a limited number of interviews.

For couples, each entry in the rank order list is a pair of programs (or being unmatched). The typical rank order list of couples averages 81 program pairs. However, the rank order list of a couple has entries for both members, so there are many duplicate programs. When we consider the number of distinct programs ranked by a couple, it is similar to the number ranked by single applicants: on average, there are about 10 distinct programs listed by each couple member. At the extreme in our data set, the maximum number of distinct programs ranked by a couple member is 1.9% of all programs.

Of course, the fact that a doctor has a short preference list does not mean they prefer to leave the profession if they cannot obtain one of their stated choices. Given our interpretation of preferences as those formed after interviews, the short rank order list means doctors only interview at a fraction of possible hospitals and they may have a complete ranking over these options. In the event that they are unassigned, they either participate in the after-market in which they can learn about additional hospitals or postpone their training for a year, as is commonly done by doing a research year to strengthen one's credentials.

Fact 2. The length of the rank order lists of applicants who are single or couples is small relative to the number of possible programs.

The next issue we examine is the distribution of applicant preferences. In Figure I, we explore the popularity of programs in our data. For each program, we compute the total number of students

11. This information is not available separately for single applicants and those who participate as couples in the NRMP.

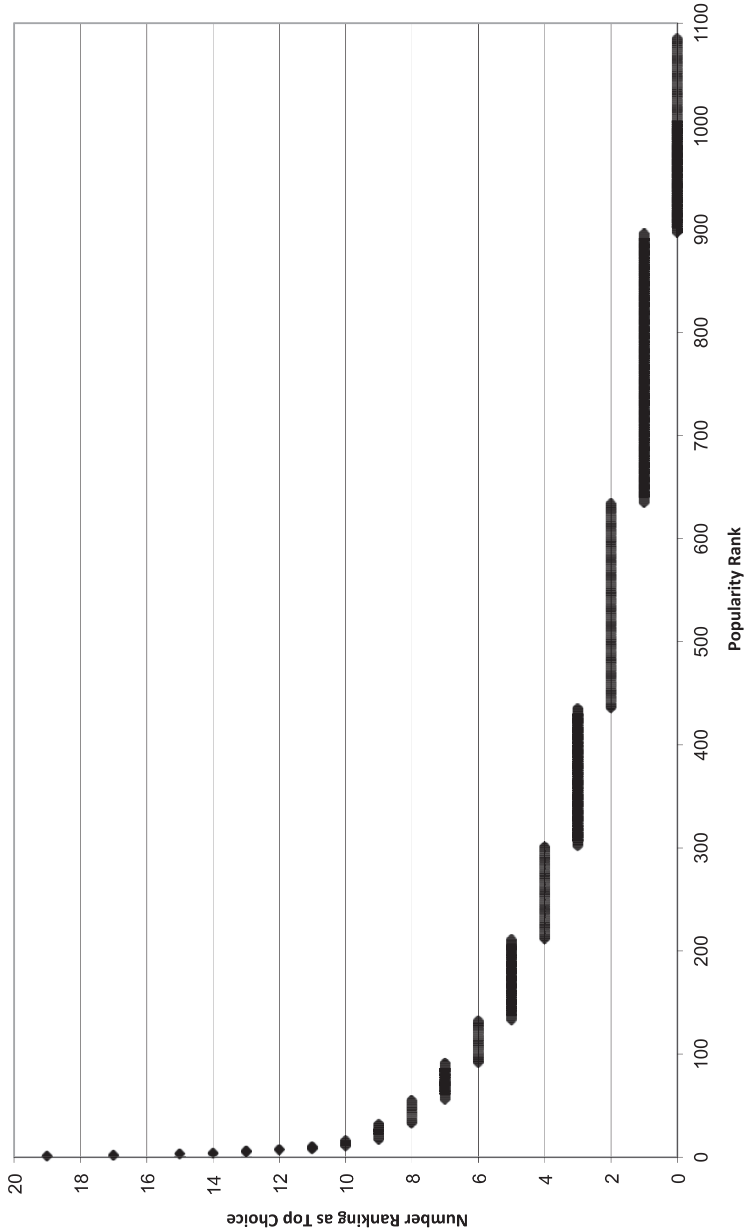

FIGURE I

Popularity of Programs as Top Choice (2003) who rank that program as their top choice. We order programs by this number, with the program with the highest number of top choices on the left and programs that no one ranks as their top choice on the right. Figure I shows the distribution of popularity for 2003. In this year, the most popular program was ranked as the top choice by 19 applicants, and 189 programs are not ranked as a top choice by any applicant. The other years of our data set display a similar pattern. Averaged across all years, the most popular program is ranked as a top choice by 24 applicants, and about 208 programs are not ranked as a top choice by anyone. The fraction of applicants ranking the most popular program as their first choice is only 0.8%. (Recall that these are preferences stated after interviews have been conducted, so it does not preclude the possibility that there are popular programs that receive many applications but only interview a small subset of applicants.)

FACT 3. The most popular programs are ranked as a top choice by a small number of applicants.

The only identifying information we have on programs are geographic regions where they are located. The 11 geographic regions in our data set are 10 regions in the United States, each of which corresponds to the first digit of the zip code of the program's location, and one region for all of Canada. Most programs are concentrated on the West Coast and in the Northeast. In Table I, we report the number of distinct regions ranked by applicants. Half of single applicants rank at most two regions. Couples, on the other hand, tend to rank slightly more regions.

For a given couple rank order list, we also compute the fraction of entries on their submitted list that have both jobs in the same region. On average, 73% of a couple's rank order list is for programs in the same region.

FACT 4. A pair of internship programs ranked by doctors who participate as a couple tend to be in the same region.

In the psychology market, there are about 1,100 internship programs. The average capacity is about 2.5 seats, and more than three-quarters of programs have 3 or fewer spots. The total capacity of internship programs is smaller than the total number of applicants who participate, which implies that each year there will be unmatched applicants. This is also true in the NRMP, where the number of positions per applicant ranges from 0.75 to 0.90 over the period 1995–2009 (NRMP 2009).

Even though there are more applicants than programs, in the APPIC match, there are a sizable number of programs that are unfilled at the end of the regular match. According to the APPIC's statistics, during 1999–2007, on average 17% of programs had unfilled positions. In the NRMP, a similar proportion of programs had unfilled seats. In 2009, for instance, 12% of programs had unfilled positions.[12]

FACT 5. Even though there are more applicants than positions, many programs still have unfilled positions at the end of the centralized match.

II.C. Stable Matchings in the Market for Psychologists

We next investigate whether a stable matching exists in the market for psychologists.[13] Roughly speaking, this is a matching such that there is no pair of hospital and applicant who prefer each other to the prescribed matching.[14] We use a variant of the procedure by Roth and Peranson (1999) to compute a stable

12. In practice, to place these remaining applicants, both markets have a decentralized aftermarket where positions are filled. In this market, applicants can communicate and informally interview with places they did not initially consider, but under very short time limits. In recent years there have been proposals to eliminate these processes completely (see, e.g., Supplemental Offer and Acceptance Program [SOAP] described at http://www.nrmp.org/soap.pdf, accessed on October 22, 2010.) The model in this article is only about the main match and does not model this decentralized aftermarket, though understanding how it may interact with the main round is an interesting question for future work.

13. The model we analyze in this article allows employers to have preferences over sets of applicants provided that the preferences are responsive. Our data on program rank order lists consist only of preferences over individual applicants. We do not know, for instance, whether a program prefers their first- and fourth-ranked applicants over their second- and third-ranked applicants. To compute a stable matching in the market for psychologists, it is necessary to specify how comparisons between individual applicants relate to comparisons between sets of applicants. For the empirical computation, when comparing sets of applicants D_1 and D_2, we assume that D_1 is more preferred to D_2 if the highest individually ranked applicant in D_1 who is not in D_2 is preferred to the highest individual ranked applicant in D_2 who is not in D_1. This would imply that the first and fourth ranked applicant are preferred over the second and third ranked applicant. (We take advantage of the more flexible formulation of preferences over sets that we employ, compared to that used in practice.)

14. Section III provides a precise definition of our stability concept.

matching.[15] For each of the nine years of data, the first column of Table II shows that a stable matching exists in the market with couples with respect to submitted preferences.[16]

FACT 6. A stable matching with respect to submitted preferences exists in all nine years in the market for psychologists.

We also compare the assignment of single applicants at the stable matching we find in a market with couples to their assignment in the applicant-optimal stable matching in a market without couples in Online Appendix Table A1. Adding couples to the market in principle could affect the assignment received by many single applicants, in practice it has little effect. This can be seen by comparing the overall distribution of choice received for single applicants in a stable matching in markets without couples and with couples. Moreover, Online Appendix Table A2 reports the exact number of single applicants who receive a less preferred assignment in the market with couples. On average, there are 19 couples or 38 applicants who participate as couples in the market and because of their presence, only 63, or 2% of single applicants obtain a lower choice. This corresponds to about three single applicants obtaining a different assignment per couple.

FACT 7. Across stable matchings, most single applicants obtain the same position in the market without couples as in the market with couples.

15. Our variation has a different sequencing of applications from single applicants and couples than that described in Roth and Peranson (1999). That paper gives some evidence that these sequencing decisions have little effect on the success of the procedure.

16. We focus on a particular stable matching in the market with couples, because we are unable to compute the entire set of stable matchings. There may be a reason to suspect that this set is small. In Online Appendix Table A3, we compute stable matchings in the market without couples and find that very few applicants and programs have different assignments across the applicant-optimal and program-optimal stable matchings. Moreover, there is evidence from real-life applications of the college admissions model without couples that the size of the stable set is small: see, for example, Roth and Peranson (1999), table 9, which examines the market for thoracic surgeons, and Pathak and Sönmez (2008), p. 1645, on Boston school choice.

TABLE II
STABLE MATCHINGS IN APPIC MARKET

			Sequential couples algorithm: reasons for failure			
Year	(1) Stable matching found	(2) Stable matching found by SCA	(3) Applicant applies to program where couple assigned (%)	(4) Couple applies to program where couple assigned (%)	(5) Applicant displaces a couple member (%)	(6) Couple displaces a couple member (%)
1999	Yes	Yes	100	0	0	0
2000	Yes	No	100	100	100	74
2001	Yes	Yes	100	0	0	0
2002	Yes	Yes	100	95	0	0
2003	Yes	Yes	0	0	0	0
2004	Yes	Yes	100	100	52	52
2005	Yes	Yes	0	0	0	0
2006	Yes	No	100	52	100	0
2007	Yes	No	100	52	100	0

Notes. This table reports the source of failure of the sequential couples algorithm. The algorithm fails when a doctor applies to a program where a couple member is assigned and is more preferred than the couple member. The table reports the outcomes of 1,000 permutations of the ordering of couples.

III. A Simple Theory of Matching with Couples

III.A. Model

A matching market consists of hospitals, doctors, and their preferences. Let H be the set of hospitals and $\emptyset$ be the outside option for doctors. Define $\tilde{H} = H \cup \{\emptyset\}$. S is the set of single doctors, and C is the set of couples of doctors. Each couple is denoted by $c = (f, m)$, where f and m denote the first and second members of couple c, respectively. When we need to refer to the members of a specific couple c, we sometimes write (f_c, m_c). Let $F = \{f \big| (f, m) \in C \text{ for some } m\}$ and $M = \{m \big| (f, m) \in C \text{ for some } f\}$ be the sets of first and second members that form couples. Let $D = S \cup F \cup M$ be the set of doctors.

Each single doctor $s \in S$ has a preference relation R_s over $\tilde{H}$. We assume that preferences are strict: if hR_sh' and $h'R_sh$, then $h = h'$. We write hP_sh' if hR_sh' and $h \neq h'$. If $hP_s\emptyset$, we say that hospital h is acceptable to single doctor s.

Each couple $c \in C$ has a preference relation R_c over $\tilde{H} \times \tilde{H}$, pairs of hospitals (and being unmatched). We assume that preferences of couples are strict with P_c denoting the asymmetric part of R_c. If $(h, h')P_c(\emptyset, \emptyset)$, then we say that pair (h, h') is acceptable to couple c. We say that hospital h is listed by R_c if there exists $h' \in \tilde{H}$ (so h' may be $\emptyset$) such that either $(h, h')P_c(\emptyset, \emptyset)$ or $(h', h)P_c(\emptyset, \emptyset)$.

Each hospital $h \in H$ has a preference relation over 2^D, all possible subsets of doctors. We assume preferences of hospitals are strict. Let $h \in H$ and κ_h be a positive integer. We say that preference relation $\succeq_h$ is responsive with capacity κ_h if it ranks a doctor independently of her colleagues and disprefers any set of doctors exceeding capacity κ_h to being unmatched (see Online Appendix A.1 for a formal definition). We follow much of the matching literature and assume that hospital preferences are responsive throughout the article. Let R_h be the corresponding preference list of hospital h, which is the preference relation over individual doctors and $\emptyset$. We write dP_hd' if dR_hd' and $d \neq d'$. We say that doctor d is acceptable to hospital h if $dP_h\emptyset$. We write $\succeq_H = (\succeq_h)_{h \in H}$. We refer to a matching market Γ as a tuple $(H, S, C, (\succeq_h)_{h \in H}, (R_i)_{i \in S \cup C})$.

We proceed to define our stability concept in markets with couples. The descriptions are necessarily somewhat more involved than those in the existing literature because we allow for capacities of hospitals larger than one (we elaborate on this issue shortly). First, it is convenient to introduce the concept of hospital choices over permissible sets of doctors. For any set of doctors and couples $D' \subseteq D \cup C$, define

$$\begin{aligned}\mathcal{A}(D') = \{D'' \subseteq D \mid & \forall s \in S, \text{if } s \in D'' \text{ then } s \in D', \\ & \forall c \in C, \text{if } \{f_c, m_c\} \subseteq D'', \text{ then } (f_c, m_c) \in D', \\ & \qquad \text{if } f_c \in D'' \text{ and } m_c \notin D'', \text{then } f_c \in D', \\ & \qquad \text{if } f_c \notin D'' \text{ and } m_c \in D'', \text{then } m_c \in D'\}.\end{aligned}$$

In words, $\mathcal{A}(D')$ is the collection of sets of doctors available for a hospital to employ when doctors (or couples of doctors) D' are applying to it. Underlying this definition is the distinction between applications by individual couple members and those by couples as a whole. For example, if $(f, m) \in D' \cap C$ but $f, m \notin D'$, then the couple is happy to be matched to the hospital if and only if both members are employed together, whereas if $(f, m) \notin D'$ but $\{f, m\} \subseteq D'$, then the couple is happy to have one member matched to the hospital but not together.

For any set $D' \subseteq D \cup C$, define the **choice** of hospital h given D', $Ch_h(D')$, to be the set such that

- $Ch_h(D') \in \mathcal{A}(D')$,
- $Ch_h(D') \succeq_h D''$ for all $D'' \in \mathcal{A}(D')$.

The choice $Ch_h(D')$ is the most preferred subset of doctors among those in D' such that each couple is either chosen or not chosen together if they apply as a couple.[17]

A matching specifies which doctors are matched to which hospitals (if any). Formally, a matching μ is a function defined on the set $\tilde{H} \cup S \cup C$, such that $\mu(h) \subseteq D$ for every hospital h, $\mu(s) \in \tilde{H}$ for every single doctor s, and $\mu(c) \in \tilde{H} \times \tilde{H}$ for every couple c where

- $\mu(s) = h$ if and only if $s \in \mu(h)$ and
- $\mu(c) = (h, h')$ if and only if $f_c \in \mu(h)$ and $m_c \in \mu(h')$.

When there are only single doctors in D', the set $\mathcal{A}(D')$ is simply the set of subsets of D'. Hence the choice $Ch_h(D')$ is the subset of D' that is the most preferred by h. This is the standard definition of $Ch_h(\cdot)$ in markets without couples (see Roth and Sotomayor

17. This formulation of hospital preferences involving couples is more general than currently implemented in practice, where hospitals' preferences are elicited only over individual members of a couple.

1990 for example), and hence the current definition is a generalization of the concept to markets with couples.

A matching is individually rational if no player can be made better off by unilaterally rejecting some of the existing partners (see Online Appendix A.1 for a formal definition). We define different cases of how the relevant small coalitions can block a matching as follows:

(1) A pair of a single doctor s and a hospital $h \in H$ is a block of μ if $hP_s\mu(s)$ and $s \in Ch_h(\mu(h) \cup s)$.[18]

(2) (a) A coalition $(c, h, h') \in C \times \tilde{H} \times \tilde{H}$ of a couple and two hospitals,[19] where $h \neq h'$, is a **block** of μ if
- $(h, h')P_c\mu(c)$,
- $f_c \in Ch_h(\mu(h) \cup f_c)$, and
- $m_c \in Ch_{h'}(\mu(h') \cup m_c)$.[20]

(b) A pair $(c, h) \in C \times H$ of a couple and a hospital is a block of μ if
- $(h, h)P_c\mu(c)$ and
- $\{f_c, m_c\} \subseteq Ch_h(\mu(h) \cup c)$.

A matching μ is stable if it is individually rational and there is no block of μ.

1. Discussion of the Solution Concepts. Models of matching with couples where hospitals have multiple positions are a particular form of many-to-many matching because each couple may seek two positions.[21] Various definitions of stability have been proposed for many-to-many matching, which differ based on the assumptions on what blocking coalitions are allowed (Sotomayor 1999, 2004; Echenique and Oviedo 2006; Konishi and Ünver

18. We denote a singleton set $\{x\}$ simply by x whenever there is no confusion.

19. Note that the definition of a blocking coalition of a couple and two hospitals includes the case in which only one member of the couple changes hospitals, so it might be thought of as being a coalition of a couple and one hospital, with the hospital whose employee doesn't move being a passive participant in the blocking coalition.

20. For the purpose of this definition, we adopt a notational convention that $Ch_\emptyset$ is an identity function, so the condition $d \in Ch_\emptyset(\mu(\emptyset) \cup d)$ is satisfied for any μ and $d \in D$.

21. More precisely, Hatfield and Kojima (2008, 2010) point out that the model is subsumed by a many-to-many generalization of the matching model with contracts as analyzed by Hatfield and Milgrom (2005).

2006a). Consequently, there are multiple possible stability concepts in matching with couples. The present definition of stability allows us to stay as close to the most commonly used pairwise stability as possible, by assuming away deviations involving large groups. Ruling out large coalitions appears to be reasonable because identifying and organizing large groups of agents may be difficult.

It is nevertheless important to understand whether our analysis is sensitive to a particular definition of stability. To address this issue, in Online Appendix A.2 we present an alternative definition of stability that allows for larger coalitions to block a matching. We show that the results of this article hold under that definition as well.

Most studies in matching with couples have focused on the case in which every hospital has capacity 1.[22] Following the standard definition of stability in such models (see Klaus and Klijn 2005, for instance), we say that a matching μ is unit-capacity stable if

(1) μ is individually rational,
(2) there exists no single doctor-hospital pair s, h such that $hP_s\mu(s)$ and $sP_h\mu(h)$, and
(3) there exists no coalition by a couple $c = (f, m) \in C$ and hospitals (or being unmatched) $h, h' \in \tilde{H}$ with $h \neq h'$ such that $(h, h')P_c\mu(c)$, $fR_h\mu(h)$, and $mR_{h'}\mu(h')$.[23]

Our concept of stability is equivalent to the unit-capacity stability as defined already if every hospital has responsive preferences with capacity 1. To see this, first observe that condition (3) of unit-capacity stability is equivalent to the nonexistence of a block as defined in condition (2a) of our stability concept. Moreover, condition (2b) of our stability concept is irrelevant when each hospital has capacity 1 because a hospital with capacity 1 never prefers to match with two members of a couple. Finally, the remaining conditions for unit-capacity stability have direct counterparts in our

22. Some papers consider multiple positions of hospitals but treat a hospital with capacity larger than 1 as multiple hospitals with capacity 1 each. This approach is customary and usually innocuous when there exists no couple because most stability concepts are known to coincide in that setting (Roth 1985). However, the approach has a consequence if couples exist because it leads to a particular stability concept. A different modeling approach is pursued by McDermid and Manlove (2009).

23. We adopt the notational convention that $dR_{\emptyset}d'$ for any $d, d' \in D \cup \emptyset$.

definition of stability. Thus the stability concept employed in this article is a generalization of the standard concept to the case where hospitals have multiple positions.

Also note that our stability concept is equivalent to the standard definition of (pairwise) stability when there exist no couples. More specifically, condition (2) of our stability concept is irrelevant if couples are not present, and condition (1) is equivalent to the nonexistence of a blocking pair which, together with individual rationality, defines stability in markets without couples.

III.B. The Existence Problem with Couples

We illustrate how the existence of couples poses problems in the theory of two sided matching. To understand the role of couples, however, it is useful to start by considering a matching without couples. In that context, the (doctor-proposing) deferred acceptance algorithm always produces a stable matching (Gale and Shapley 1962).

ALGORITHM 1 (DOCTOR-PROPOSING DEFERRED ACCEPTANCE ALGORITHM) Input: a matching market $(H, S, (\succeq_h)_{h \in H}, (R_s)_{s \in S})$ without couples.

Step 1: Each single doctor applies to her first choice hospital. Each hospital rejects its least preferred doctor in excess of its capacity and all unacceptable doctors among those who applied to it, keeping the rest of the doctors temporarily (so doctors not rejected at this step may be rejected in later steps).

In general, Step t: Each doctor who was rejected in Step $(t-1)$ applies to her next highest choice (if any). Each hospital considers these doctors and doctors who are temporarily held from the previous step together, and rejects the least preferred doctors in excess of its capacity and all unacceptable doctors, keeping the rest of the doctors temporarily (so doctors not rejected at this step may be rejected in later steps).

The algorithm terminates at a step where no doctor is rejected. The algorithm always terminates in a finite number of steps. At that point, all tentative matchings become final. Gale and Shapley (1962) show that for any given market without couples, the matching produced by the deferred acceptance algorithm is stable. Furthermore, they show that it is the doctor-optimal stable matching, the stable matching that is weakly preferred to any other stable matching by all doctors.

By contrast, stable matchings do not necessarily exist even when there is only one couple in the market (shown by Roth 1984 and an unpublished work by Sotomayor). This fact is illustrated in the following example, based on Klaus and Klijn (2005).

EXAMPLE 1. Let there be a single doctor s and a couple $c = (f,m)$ as well as two hospitals h_1 and h_2, each with capacity 1. Suppose the acceptable matches for each agent, in order of preference, are given by:

$$R_c : (h_1, h_2) \qquad R_s : h_1, h_2$$
$$\succeq_{h_1} : f, s \qquad \succeq_{h_2} : s, m.$$

We illustrate that there is no stable matching in this market, by considering each possible matching.

1. Suppose $\mu(c) = (h_1, h_2)$. Then single doctor s is unmatched. Thus single doctor s and hospital h_2 block μ because s prefers h_2 to her match $\mu(s) = \emptyset$ and h_2 prefers s to its match $\mu(h_2) = m$.
2. Suppose $\mu(c) = (\emptyset, \emptyset)$.
   (a) If $\mu(s) = h_1$, then (c, h_1, h_2) blocks μ because couple c prefers (h_1, h_2) to their match $\mu(c) = (\emptyset, \emptyset)$, hospital h_1 prefers f to its match $\mu(h_1) = s$, and hospital h_2 prefers m to its match $\mu(h_2) = \emptyset$.
   (b) If $\mu(s) = h_2$ or $\mu(s) = \emptyset$, then (s, h_1) blocks μ because single doctor s prefers his first-choice hospital h_1 to both hospital h_2 and $\emptyset$ whereas h_1 prefers s to its match $\mu(h_1) = \emptyset$.

Klaus and Klijn (2005) identify a sufficient condition to guarantee the existence of a stable matching called weak responsiveness. A couple's preferences are said to be responsive if an improvement in one member's assignment is an improvement for the couple. Preferences are said to be weakly responsive if the requirement applies to all acceptable positions.[24] The preferences of couples in Example 1 do not satisfy this condition. If, for instance, the couple's preferences are $(h_1, h_2), (h_1, \emptyset), (\emptyset, h_2), (\emptyset, \emptyset)$, in order of preference, then it satisfies responsiveness and a stable matching exists. Klaus and Klijn (2005) write that "responsiveness

24. See Klaus, Klijn, and Nakamura (2009) for formal definition.

essentially excludes complementarities in couples' preferences." They showed that:

(1) if the preferences of every couple are weakly responsive, then there exists a stable matching.
(2) if there is at least one couple whose preferences violate weak responsiveness while satisfying a condition called "restricted strict unemployment aversion," then there exists a preference profile of other agents such that preferences of all other couples are weakly responsive but there exists no stable matching.

Their second result says that the class of weakly responsive preferences is the "maximal domain" of preferences. That is, it is the weakest possible condition that can be imposed on individual couples' preferences that guarantees the existence of stable matchings.[25]

There seem to be many situations in which couple preferences violate weak responsiveness. One reason may be geographic, as stated as Fact 4: both programs ranked as a pair by a couple tend to be in the same geographic region. For example, the first choice of a couple of medical residents may be two residency programs in Boston and the second may be two programs in Los Angeles; one member working in Boston and the other working in Los Angeles could be unacceptable because these two cities are too far away from each other. The coordinator of the APPIC matching program writes in Keilin (1998, 602) that "most couples want to coordinate their internship placements, particularly with regard to geographic location." This suggests that violation of weak responsiveness due to geographic preferences is one of the representative features of couple preferences.[26]

To further study this question empirically, we analyze the data on the stated preferences of couples from the APPIC. Stated preferences of couples may not be their true preferences because the truth-telling is not necessarily their dominant strategy. However, there are reasons that couples may not want to misrepresent their preferences. First, it may be complicated for a couple to determine a profitable deviation. Moreover, truth-telling may

25. Hatfield and Kominers (2009) show that the substitutes condition is a maximal domain in the absence of restricted strict unemployment aversion.

26. For an investigation of decision making among couples in the market for new Ph.D. economists, see Helppie and Murray-Close (2010).

be focal, especially since clearinghouse organizers explicitly encourage participants to report their preferences truthfully. Finally, Kojima, Pathak, and Roth (2010) provide conditions under which truth-telling is an approximate equilibrium when there are large numbers of market participants. During years for which we have data (1999–2007), preferences of only 1 couple out of 167 satisfy weak responsiveness. Thus the data suggest, in light of the results of Klaus and Klijn (2005), that it is virtually impossible to guarantee the existence of a stable matching in such markets with couples based on a domain restriction of preferences.

However, the fact that the preferences of the overwhelming majority (166 out of 167) of couples violate weak responsiveness does not mean that a stable matching does not exist in the psychologist market. Stable matchings have been found in many labor markets despite the presence of couples, and as we described in Section II.C, we find a stable match for each of the nine years of the psychology market for which we have data. This motivates our desire to understand what market features enable the existence of stable matchings most of the time, when the known sufficient conditions on couples' preferences do not guarantee existence.

III.C. Sequential Couples Algorithm

The original deferred acceptance algorithm does not incorporate applications by couples. We consider an extension of the algorithm, which we call the sequential couples algorithm. Although we defer a formal definition to Online Appendix A.3 for expositional simplicity, we offer an informal description as follows.

(1) run a deferred acceptance algorithm for a submarket composed of all hospitals and single doctors, but without couples;
(2) one by one, place couples by allowing each couple to apply to pairs of hospitals in order of their preferences (possibly displacing some doctors from their tentative matches); and
(3) one by one, place singles who were displaced by couples by allowing each of them to apply to a hospital in order of her preferences.

We say that the sequential couples algorithm succeeds if there is no instance in the algorithm in which an application is made to a hospital where an application has previously been made by a member (or both members) of a couple except for the couple who is currently applying. Otherwise, we declare a failure and terminate the algorithm.

Failure of the sequential couples algorithm does not mean that a stable matching does not exist. Therefore, in practice, a matching clearinghouse would be unlikely to declare failure when the sequential couples algorithm fails, but would instead consider some procedure to try to assign the remaining couples and find a stable matching. This is the main idea behind the Roth-Peranson algorithm (Roth and Peranson 1999), which is the basis for the mechanism used in the NRMP, APPIC, and other labor markets. If the sequential couples algorithm succeeds, then the Roth-Peranson algorithm produces the matching reached by the sequential couples algorithm. However, the sequential couples algorithm and the Roth-Peranson algorithm are different in two aspects.[27]

First, where the sequential couples algorithm fails, the Roth-Peranson algorithm proceeds and tries to find a stable matching. The algorithm identifies blocking pairs, eliminating instances of instability one by one, in a manner similar to Roth and Vande Vate (1990). Note that since a stable matching does not necessarily exist in markets with couples, the Roth-Peranson algorithm could cycle without terminating. However, the algorithm forces termination of a cycle and proceeds with processing other applicants. This sometimes ultimately results in a stable matching, and sometimes no stable matching is found. Second, in the Roth-Peranson algorithm, when a couple is added to the market with single doctors, any single doctor who is displaced by the couple is placed before another couple is added. By contrast, the sequential couples algorithm holds any displaced single doctor without letting her apply, until it processes applications by all couples.[28]

The reason we focus on this simplified procedure is that the success of the sequential couples algorithm turns out to be

27. A complete description of the Roth-Peranson algorithm, specifically how the algorithm terminates cycles and proceeds with processing, is not publicly available, but a more detailed description than the one provided here is offered by Roth and Peranson (1999), and a flowchart of the algorithm appears in Roth (2013).

28. As we will point out subsequently, our result also holds if we follow the sequencing of doctors as in the Roth-Peranson algorithm. We chose the current definition of the sequential couples algorithm for expositional simplicity.

sufficient to verify the existence of a stable matching (the proof is in Online Appendix A.3).

LEMMA 1. If the sequential couples algorithm succeeds, then the resulting matching is stable.

To illustrate the main idea of Lemma 1, we consider how the sequential couples algorithm proceeds for the market in Example 1. In step 1 of the algorithm, we run the doctor-proposing deferred acceptance algorithm in the submarket without couples. Single doctor s proposes to hospital h_1 and is assigned there. Then in step 2, we let couple c apply to their top choice (h_1, h_2). Couple member f is preferred to s by h_1, and couple member m is preferred to a vacant position by h_2. Thus f and m are tentatively assigned to h_1 and h_2, respectively while s is rejected. Then in step 3, we let s apply to her next highest choice. In this case, she applies to hospital h_2, where a couple member m has applied and been assigned before. At this point we terminate the algorithm and declare that it has failed.

To see why declaring a failure of the sequential couples algorithm is useful, suppose that we hypothetically continue the algorithm by allowing h_2 to reject m because h_2 prefers s to m. Then the couple prefers being unassigned rather than having only f be matched to h_1, so doctor f would like to withdraw his assignment from hospital h_1. Suppose we terminate the algorithm at this point once f becomes unmatched. Then the resulting matching assigns no doctor to h_1 and s to h_2. This matching is unstable because doctor s can block with hospital h_1. On the other hand, if we continue the algorithm further by allowing s to match with h_1, then the resulting matching is identical to the one obtained at the end of step 1 of the sequential couples algorithm. This suggests that reasonable algorithms would cycle without terminating in this market.

The idea of declaring failure of the sequential couples algorithm is to avoid a situation like the foregoing example and turns out to be a useful criterion for judging whether the algorithm produces a stable matching. Of course, the algorithm sometimes fails even if there exists a stable matching, so its success is only a sufficient condition for the existence of a stable matching. What is remarkable is that looking at this particular sufficient condition turns out to be enough for establishing that a stable matching exists with a high probability in the environment we study here. Moreover, there is a sense in which it is necessary to use an algorithm that finds a stable matching only in some instances, rather than one that always finds a stable matching whenever it exists. Ronn (1990) shows that the problem of determining whether a market with couples has a stable matching is computationally hard (NP-complete). The result suggests that it may be inevitable to employ an approach that does not always find a stable matching like our sequential couples algorithm.

Example 1 illustrates that the sequential couples algorithm does not necessarily succeed and suggests that markets of any finite size would allow such a failure. We instead consider a large market environment with a random component in the preferences of the market participants. Our contribution is to demonstrate that, with high probability, the sequential couples algorithm succeeds, and hence a stable matching exists in this environment.

IV. Large Markets

IV.A. Random Markets

We have seen that a stable matching does not necessarily exist in a finite matching market with couples. To investigate how often a stable matching exists in large markets, we introduce the following random environment. A random market is a tuple $\tilde{\Gamma} = (H, S, C, \succeq_H, k, \mathcal{P}, \rho)$, where k is a positive integer, $\mathcal{P} = \{p_d(\cdot)\}_{d \in D}$ is a collection of probability distributions for each doctor d on H, and ρ is a function that maps two preferences over $\tilde{H}$ to a preference list for couples (explained shortly). Each random market induces a market by randomly generating preferences of doctors as follows:

Preferences for Single Doctors: For each single doctor $s \in S$,

Step 1: Select a hospital h with probability $p_s(h)$. List this hospital as the top-ranked hospital of single doctor s.

In general,

Step $t \leq k$: Select a hospital h with probability $p_s(h)$. Repeat until a hospital is drawn that has not been previously drawn in steps 1 through $t-1$ or every hospital h such that $p_s(h) > 0$ has already been drawn. List this hospital as the t^{th} most preferred hospital of single doctor s.

In other words, to construct the preference list, we draw hospitals repeatedly without replacement (at most) k times according to distribution $p_s(\cdot)$. Single doctor s finds these k hospitals acceptable, and all other hospitals unacceptable. For example, if $p_s(\cdot)$ is the uniform distribution on H, then the preference list is drawn from the uniform distribution over the set of all preference lists of length k.

Preferences for Doctors who are Couples: Couples' preferences are formed by drawing preferences, R_f and R_m, for each doctor in the couple $c = (f, m)$. R_f and R_m are constructed from distributions $p_f(\cdot)$ and $p_m(\cdot)$ following the process used to generate preferences for a single doctor already described.

To construct the preference list for the couple $c = (f, m)$, define $\rho(R_f, R_m)$ to be a preference of the couple with the following restriction: if (h_1, h_2) is acceptable according to $\rho(R_f, R_m)$, then $h_1 R_f \emptyset$ and $h_2 R_m \emptyset$. This is the only restriction we place on ρ.

Preferences for Hospitals: Each hospital h has a responsive preference relation defined over sets of doctors $\succeq_h$ such that all doctors are acceptable.

1. Discussion of Modeling Choices. The model assumes that doctors' preferences are drawn independently from one another in a statistical sense. However, doctors' preferences are not necessarily drawn from identical distributions and hence can be substantially different from one another. This modeling assumption can capture various possibilities for doctor preferences that may be important empirically. For example, the assumption allows a situation in which some popular hospitals are listed with higher probability than others by many doctors. It also allows preference distributions where doctors from different regions prefer programs from their own region to those outside of their region. Moreover, the assumption also allows for couples to have systematically different views on desirability of hospitals than single doctors. Allowing the doctors' preferences to differ in this way makes the assumption more general than the identical preference distribution assumption in Immorlica and Mahdian (2005), Kojima and Pathak (2009), and Manea (2009). In Section VI.C, we discuss the assumption in more detail and examine ways it can be relaxed.

The function ρ is a mapping that outputs a preference relation for each couple (f, m) given the pair of preferences R_f and R_m over $\tilde{H}$. One could interpret $\rho(R_f, R_m)$ as describing the outcome of household bargaining when preferences of the members are R_f and R_m, respectively. For example, the function ρ can represent a process in which any pair of hospitals that are too far away from each other is declared unacceptable, which seems to be consistent with the observed rank order lists of couples described earlier. We remain agnostic about ρ except that a hospital pair (h, h') is weakly acceptable for the couple under $\rho(R_f, R_m)$ only when h and h' are listed under R_f and R_m, respectively. In other words, no hospital appears in the preference list of a couple unless it is considered by the relevant member of the couple. Note that this, of course, does not impose that the couples preferences are weakly responsive. All our results are unchanged if we allow the function ρ to vary across different couples, but we model a common function ρ for all couples for expositional simplicity. Moreover, our results also hold when couples draw their preferences jointly from some distribution over pairs of hospitals.

Some NRMP participants who participate as couples are advised to form preferences by first forming individual rank order lists after interviewing with programs. Then, these lists serve as an input into the joint ranking of the couple. For instance, medical students who are couples at the University of Kansas Medical School are suggested to make a list of all possible program pair combinations from both individual rank order lists by computing the difference between the ranking number of the program on each individual's rank order list and trying to minimize this difference in their joint rank order list. This would be one example of a ρ function.[29]

29. The details on this advice are available online at http://medicine.kumc.edu/school-of-medicine/osa/residency-information/couples-match.html (accessed July 31, 2013). The clearinghouse for new doctors in Scotland only allows couple members to submit individual rank order lists, in contrast to the model we analyze here. In that context, their mechanism combines these lists into a preference over pairs for the couple using their individual lists and a table of positions that are determined to be geographically compatible by the mechanism. See the discussion of the Scottish Foundation Allocation Scheme at http://www.dcs.gla.ac.uk/~pbiro/applications/uk_sfas.html (accessed July 31, 2013).

The probabilistic structure we place on doctor preferences is unneeded for hospital preferences. Rather, hospital preferences can be arbitrary except for two important restrictions. First, hospital preferences are assumed to be responsive, as in much of the literature on two-sided matching. The labor market clearinghouses that motivate our study impose this restriction by eliciting preferences over individual doctors.

The second important assumption on hospital preferences is that hospitals find all doctors acceptable. We make this assumption so that there are enough hospitals that can actually hire doctors in large markets. At first glance, this assumption seems violated in the data from the market for clinical psychologists as no program submits a rank order list of all doctors (i.e., as seen in Table I, the average number of doctors listed in a hospital's preference list is 16.7 in our APPIC data). However, the programs rank most doctors who have ranked them, and an equivalent assumption is that each hospital finds acceptable only doctors who list that hospital, for example, because hospitals (and doctors) will only rank doctors (and hospitals) they have interviewed. Clearly the existence result follows under this alternative assumption, because any stable matching at the original preference profile is also stable under the modified preference profile. The results also follow, at additional notational complexity, in a model where at least a constant fraction of hospitals find all doctors acceptable.

IV.B. Regular Sequence of Random Markets

To analyze limit behavior of the matching market as the market becomes large, we consider a sequence of markets of different sizes. A sequence of random markets is denoted by $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \ldots)$, where $\tilde{\Gamma}^n = (H^n, S^n, C^n, \succeq_{H^n}, k^n, \mathcal{P}^n, \rho^n)$ is a random market in which $|H^n| = n$ is the number of hospitals. Consider the following regularity conditions.

DEFINITION 1. A sequence of random markets $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \ldots)$ is *regular* if there exist $\lambda > 0,\ a < \frac{1}{2},\ b > 0,\ r \geq 1$, and $\gamma < \frac{1-2a}{r\lambda}$ such that for all sufficiently large n,

1. $|S^n| \leq \lambda n,\ |C^n| \leq bn^a$,
2. $k^n \leq \gamma \log(n)$,
3. $\frac{p_d(h)}{p_d(h')} \leq r$ for all doctors d in D^n and hospitals h, h' in H^n.

Condition 1 requires that the number of single doctors does not grow much faster than the number of hospitals. Moreover, couples do not grow at the same rate as the number of hospitals and instead grow at the slower rate of $O(n^a)$ where $a < \frac{1}{2}$. This condition is motivated by Fact 1 that the number of couples is small compared with the number of hospitals or single doctors. Note that the assumption also implies that the total number of applicants $|S^n| + |C^n|$ is of order at most n and is consistent with either more doctors than hospitals or fewer. Condition 2 assumes that the length of doctors' preference lists does not grow too fast when the number of market participants grows.[30] This assumption is motivated by Fact 2 that the length of rank order lists of doctors is small relative to the number of programs. Allowing the length of doctor's preference lists to vary does not change any of our results as long as each doctor's rank order list is no longer than $\gamma \log(n)$. Condition 3 requires that the popularity of different hospitals (as measured by the probability of being listed by doctors as acceptable) does not vary too much, as suggested by Fact 3.

V. Existence of Stable Matchings

As seen in Example 1, a stable matching does not necessarily exist when some doctors are couples. However, there is a sense in which a stable matching is likely to exist if the market is large as formalized by our main result:

Theorem 1. Suppose that $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \ldots)$ is a regular sequence of random markets. Then the probability that there exists a stable matching in the market induced by $\tilde{\Gamma}^n$ converges to 1 as the number of hospitals n approaches infinity.

We defer the formal proof to Online Appendix A.3 and describe the argument here. Our proof involves analysis of the sequential couples algorithm in a regular sequence of random markets. By Lemma 1, we know that a stable matching exists whenever the algorithm succeeds. Our proof strategy is to show that the probability that the sequential couples algorithm succeeds converges to 1 as the market size approaches infinity.

30. In Kojima, Pathak, and Roth (2010), we considered the case where the doctor rank order list is bounded. We thank the referees for suggesting that we relax this assumption to allow for a growing rank order list length.

Suppose that there are a large number of hospitals in the market. Given our assumptions on the distribution of couples' preferences, different couples are likely to prefer different pairs of hospitals. Hence, in step 2 of the algorithm, members of two distinct couples are unlikely to apply to the same hospital. In such an instance, this step of the algorithm tentatively places couples without failure. Given that, it suffices to show that the single doctors displaced in steps 2 and 3 (if any) are likely to be placed without applying to a hospital where a couple has applied. To show this, first we demonstrate that if the market is large, then it is a high-probability event that there are a large number of hospitals with vacant positions at the end of step 2 (even though there could be more applicants than positions: see Proposition 4.2 in the Online Appendix).[31] Then, any single doctor is much more likely to apply to a hospital with a vacant position than to one of the hospitals that has already received an application by a couple member. Since every doctor is acceptable to any hospital by assumption, a doctor is accepted whenever an application is made to a vacant position. With high probability the algorithm places all the single doctors in step 3, resulting in a success. Together with Lemma 1, we conclude that if the market is large enough, then the probability that there exists no stable matching can be made arbitrarily small. This completes the argument.

As explained in Section III.C, the sequential couples algorithm is similar to but slightly different from the Roth-Peranson algorithm in the order of which doctors apply to hospitals. However, it is clear from the proof that the argument can be modified for the Roth-Peranson algorithm. Therefore, we have the following result as a corollary.

COROLLARY 1. Suppose that $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \ldots)$ is a regular sequence of random markets. Then the probability that the Roth-Peranson algorithm produces a stable matching in the market induced by $\tilde{\Gamma}^n$ converges to 1 as the number of hospitals n approaches infinity.

31. Note that the feature that there are many hospitals with vacant positions is consistent with Fact 5, which states that there are many resident programs with vacant positions in practical matching markets.

VI. Incentives, Robustness, and Extensions

VI.A. Incentives

The previous section establishes our main result on the existence of a stable matching with respect to reported preferences that follow certain distributional assumptions. In practice, however, preferences are private information of market participants, and the matching clearinghouse needs to elicit this information. Thus a natural question is whether there is a mechanism that induces participants to report true preferences and produces a stable matching with respect to the true preferences.

One motivation for studying this question comes from the market for psychologists. The following advice is given to participants by clearinghouse organizers:[32]

> IMPORTANT: There is only one correct "strategy" for developing your Rank Order List: simply list your sites based on your true preferences, without consideration for where you believe you might be ranked by them. List the site that you want most as your #1 choice, followed by your next most-preferred site, and so on.
>
> The previous paragraph is so important that we are going to repeat it: simply list your sites based on your true preferences.

Similar recommendations are made in other labor markets with couples. Below is the advice for participants offered by the NRMP.[33]

> Programs should be ranked in sequence, according to the applicant's true preferences.... It is highly unlikely that either applicants or programs will be able to influence the outcome of the Match in their favor by submitting a list that differs from their true preferences.

32. "FAQ for Internship Applicants" on the APPIC website, http://www.appic.org/match/faqs/applicants/rank-order-lists#q2 (accessed July 30, 2013).

33. "Rank Order Lists" on the NRMP website, http://www.nrmp.org/fellow/rank_order.html (accessed November 11, 2009).

In these quotes, market participants are advised to report their true preferences to the matching authority. This advice may have been based on the incentive properties of stable mechanisms in markets without couples. For instance, without couples, in the doctor proposing deferred acceptance algorithm, truth-telling is a dominant strategy for doctors. However, in markets with couples, this is no longer the case.

In the working paper version (Kojima, Pathak, and Roth 2010), we consider the issue of incentives after participants have conducted their interviews. The question we study is whether given a particular mechanism which finds a stable matching with high probability, do applicants have an incentive to truthfully rank the programs that interviewed them? At a first glance, a positive result seems elusive: there exists no mechanism that is stable and strategy-proof even without couples. However, we provide conditions under which truth-telling is an approximate Bayes-Nash equilibrium in a large regular market.

VI.B. Robustness

Theorem 1 is an asymptotic result, and the probability that a stable matching exists is not guaranteed to reach 1 in any finite market. On the other hand, any particular market has only finite numbers of market participants. In this section, we examine the relevance of the asymptotic result to the applications which motivate our analysis.

1. Speed of Convergence. The first issue is the order (speed) of convergence. Our model is quite general in terms of a number of parameters, making it elusive to establish an appealing result about the order of convergence.

However, the proof of the theorem allows us to evaluate the convergence probability for each given sequence of random markets, and this enables us to obtain a sharp result for special cases. In Online Appendix A.3.1, we show that if the number of couples is bounded and the length of each doctor's preference list is bounded along the sequence of random markets (that is, $a = 0$ and $k^n \leq k$ for some constant k in Definition 1), then the probability that there is no stable matching approaches 0 at least with the rate of convergence $O\left(\frac{1}{n}\right)$.

The key to this observation is inequality (19) in Online Appendix A.3, which provides a lower bound of the probability that a sequential couples algorithm successfully finds a stable matching. As explained there, the convergence order result is obtained through a bounding exercise of this inequality for the special case. Note however that inequality (19) itself holds generally for any choice of parameters as long as our regularity and thickness conditions hold. Therefore, one can evaluate inequality (19) to obtain an order of convergence more generally. In that sense, our analysis may shed light on the speed of convergence given any relevant information about the sequence of random markets.

There is an important sense in which our bounds do not exactly match the bounds required for finite markets like APPIC with 2,000 participants largely because of constants in expression (19). To see whether the qualitative insights from the theory are relevant for actual market sizes, we turn to simulation evidence shortly.

2. Proportion of Market Participants Who Are Couples. Our model made a number of assumptions, some of which could be relaxed. Perhaps the strongest assumption in our analysis relates to the growth rate of couples. Empirically, there are few couples in actual markets, but this fact does not directly imply the appropriate growth rate for a sequence of markets. Subsequent work by Ashlagi, Braverman, and Hassidim (2011) studies a model with similar features. They state that a stable matching exists (with a slightly different notion of stability) when couples are allowed to grow at rate $n^{1-\epsilon}$ for $\epsilon > 0$. Their work differs from ours in that they examine the implications of considering a particular sequence of proposals by couples and find the order that is least likely to generate existence problems.[34]

3. Length of Doctor Rank Order Lists. As stated in Definition 1, our result holds even when the number k^n of hospitals that a doctor finds acceptable grows without a bound as the market size

34. See also Biró and Irving (2010) for simulations and analysis related to a special case of the couples problem that arises in a medical labor market in Scotland, in which hospitals rank all applicants (including the individual members of couples) according to a common exam score. Biró and Irving show that the problem of determining if a stable matching exists remains computationally hard even in this special case, but simulations show that the probability of the set of stable matchings being empty is low when the proportion of couples is low.

n grows, as long as its growth speed is slow enough. More specifically, our proof goes through if $k^n \leq \gamma \log(n)$ where $\gamma < \frac{1-2a}{16\lambda}$.

Our result is a limit one based on probability bounds. As such, our asymptotic prediction is not directly applicable to any finite economy. That being said, we can use the above expression to make some qualitative predictions. To do so, note that the upper bound of γ is decreasing in a and λ: Recall that a is the parameter representing the growth speed of couples, and λ is the upper bound of the ratio of the number of doctors to the number of hospitals. Thus the comparative statics here suggests that asymptotic existence holds for a faster growth speed of doctor preference list length when the number of couples grows more slowly and the number of doctors are smaller. The caveat of this argument is, of course, that our result provides only a lower bound of the existence probability, so our exercise here cannot be conclusive.

4. Performance of Sequential Couples Algorithm in APPIC. We found a stable matching in the APPIC data set for all nine years using the algorithm described in Section B.2 in the Online Appendix. This algorithm is a variant of the sequential couples algorithm (SCA) with particular cycling and termination rules. Here we investigate whether the intuition expressed from the formal analysis of the SCA provides insight into existence for the clinical psychology market. Recall that we define the SCA as succeeding if there is no instance in the algorithm in which an application is made to a hospital where an application has previously been made by a member (or both members) of a couple except for the couple who is currently applying. Under this definition, failure of the SCA does not imply that a stable matching does not exist. Our formal results focused on this definition because it was sufficient to derive formal properties in the market with couples.

To examine how the sequential couples algorithm performs for the APPIC market, in column (2) of Table II we report on whether there is an ordering of couples for which a stable matching is found using the SCA. The definition of failure used in our proof provides only a lower bound, and for Table II we investigate a stronger definition of failure that still works for our purposes. We declare failure when a single doctor or couple applies to a program where a couple member is assigned and this applicant is more preferred than the couple member. This definition is preferred because declaring failure when an applicant applies to a program where a couple is assigned is relatively common in the APPIC data set, even though it suffices for our formal argument and a stable matching continues to be found even when it occurs. If the SCA fails under this slightly stronger definition of failure, it also fails under the definition used in our proof. Even though the SCA by itself is a relatively naive procedure, remarkably, a simple application of the algorithm together with the appropriate sequencing of couples finds a stable matching in six out of nine years.[35]

The fact that SCA finds a stable matching two thirds of the time even in a finite market suggests that analyzing its properties may provide intuition for why stable matchings are found in practice. The key phenomena are doctor's short rank order lists, the small number of couples, and the limited overlap between preferences in doctor's rank order lists. Indeed, in two of the six successful years, 2003 and 2005, it is never the case that an applicant even applies to a program where a couple is assigned as shown in column (3).[36] In another two of those successful years, there is limited overlap in couple's rank order lists. In two years (1999 and 2001) a couple does not apply to a program in the SCA where another couple is assigned, as shown in column (4).

5. Simulations Varying Market Size and Number of Couples. Data from the APPIC market does not provide guidance on how the likelihood of existence changes with market size and the number of couples in the market. We next turn to simulations of markets under different assumptions on size, the number of couples, and the distribution of preferences. Figure II reports the fraction of markets where we find a stable matching for a one-to-one matching market with preferences drawn from a uniform distribution. We set single doctors' rank order lists to have length 10 and couples' joint rank order lists to have length of about 40 on average and focus on varying the fraction of couples

35. Because the ordering of couples is arbitrary and we do not want our conclusions to be influenced by a particular ordering, in these columns we investigate what happens for 1,000 random permutations of the ordering of couples and declare success if we find a stable matching for at least one of those orderings.

36. This is consistent with the definition of failure that is sufficient for our statement of Lemma 1.

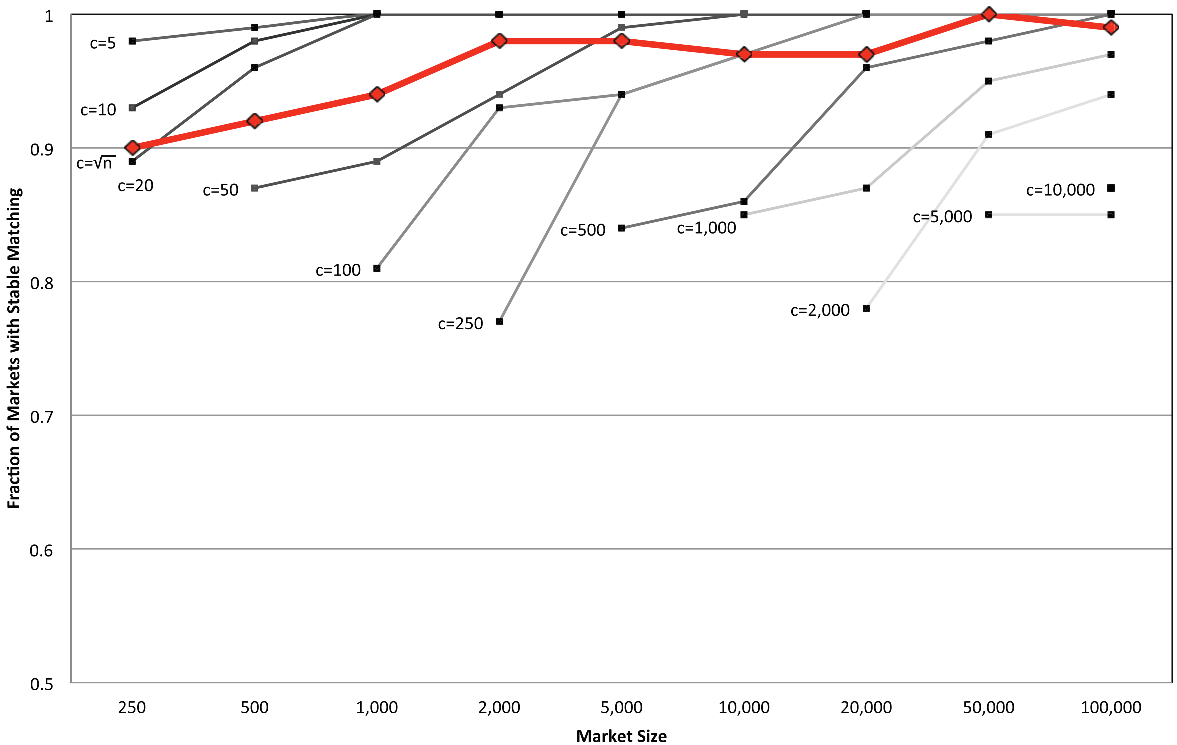

FIGURE II

Market with a Stable Matching with Preferences from Uniform Distribution in the market.[37] (Online Appendix Section B.1 precisely describes how we construct market primitives for each market size.)

Each line in Figure II corresponds a particular number of couples in the market as we vary market size. For a given number of couples, as market size increases, the fraction of simulations for which we find a stable matching increases. For instance, with 20 couples, we always find a stable matching under each simulation once the market size is greater than 1,000. For an economy like the NRMP, where over the past decade about 600 applicants are couples (corresponding to 1,200 couple members) and the market size averages 26,000, the probability a stable matching is found is roughly 95%. Finally, the thicker shaded line in Figure II shows the fraction of markets with a stable matching when the number of couples is equal to $\sqrt{n}$, corresponding to an upper bound of the growth rate in Theorem 1. Once the market size is at least 2,000, a stable matching exists in at least 96% of the markets.

Even though Figure I suggests that doctor preferences are widely distributed across programs, it is possible that the uniform distribution does not adequately capture the distribution of doctor preferences in real markets. In Figure III, we use the APPIC data set to calibrate the distribution for applicant and program preferences. We fit models of doctor and program demand for one another and simulate an economy based on these estimates. (The precise details on how we generate preferences are in the Online Appendix Section B.1.) The pattern of how existence changes with market size and the number of couples is similar between markets with the APPIC-calibrated distribution and the uniform distribution. For a given number of couples, the likelihood of finding a stable matching weakly increases with market size for the 100 markets we simulate.

Overall, Figures II and III suggests that the existence probability increases with market size for a given number of couples. Market sizes and couples fractions like those observed in APPIC or the NRMP display a high likelihood of stable matching existence in our simulation.

37. Roth and Peranson (1999) report simulations varying the length of doctor rank order lists in markets without couples.

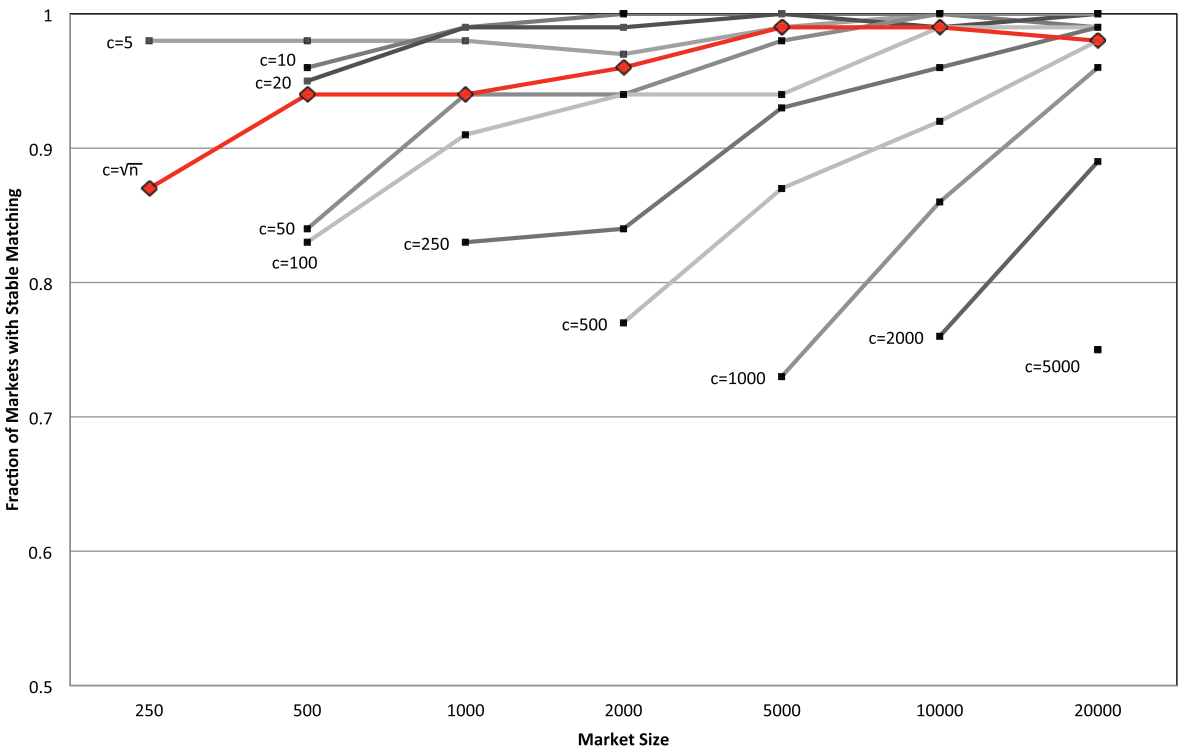

FIGURE III

Market with a Stable Matching with Calibrated Preferences

VI.C. Extensions

This article focuses on regular sequences of random markets and makes use of each condition in our arguments. A notable implication of the model is that, if the market is large, then it is a high-probability event that there are a large number of hospitals with vacant positions, even if there are more applicants than positions (for formal statements, see Proposition 2 in the Online Appendix). Note that the feature that there are many hospitals with vacant positions is consistent with Fact 5.

1. Distribution of Preferences. Our model assumed certain regularities in the way that random markets grow. In particular, there are some nontrivial restrictions on doctors' preference distributions. Of course, some distributional assumptions are needed in large market analysis: For instance, Immorlica and Mahdian (2005) offer an example where preference distributions violate a regularity assumption and their result fails even without couples.

One of the modeling assumptions in the preceding text is that doctors' preferences are drawn independently from one another in a statistical sense. However, doctors' preferences are not necessarily drawn from identical distributions and hence can be substantially different from one another. As discussed in Section IV.A, this modeling assumption can capture various possibilities for doctor preferences that may be important empirically. These include doctors from a particular region ranking own-region programs higher than programs outside of their region. Also, the model allows single doctors and couple members to draw their rankings from different distributions.

In fact, statistical independence was introduced for expositional simplicity, but we do not even need this assumption. On the contrary, the result can be generalized to environments with preference correlation and even aggregate shocks in other variables, such as supply of doctors and hospitals' job openings. To see this point, consider a model in which there is a state variable σ that is drawn randomly from a certain distribution. For each realization of the state variable σ, there is a sequence of random markets $(\tilde{\Gamma}^1(\sigma), \tilde{\Gamma}^2(\sigma), \ldots)$. We assume that $(\tilde{\Gamma}^1(\sigma), \tilde{\Gamma}^2(\sigma), \ldots)$ satisfies the regularity condition for each σ, but allow $\tilde{\Gamma}^n(\sigma)$ to be different for different realizations of σ. Doctor preferences are conditionally independent given σ, but this framework allows for doctor preferences to be correlated through their dependence on the common shock represented by the state variable. Clearly, our result generalizes to this model: the probability that a stable matching exists conditional on state variable σ converge to 1 as the market size approaches infinity for each σ by our previous analysis, and the unconditional probability of existence is merely a weighted average of these conditional probabilities.

This model allows for correlations in doctor preferences as well as other variables that may be useful in applications. For instance, σ could represent the state of the economy in different regions of the country, which then determines how many positions are open in each hospital and which hospitals are popular. In that case, the realization of σ determines the preference distributions from which doctors draw their preference list in such a way that hospitals in regions with positive wage shocks are popular and those in regions with negative wage shocks are unpopular. Similarly, σ could represent changes in funding of medical residency training by Medicare. Medicare is a major source of funding for residency training, which is subject to heated debate and many reform proposals (Rich et al. 2002). Thus the Medicare funding level may be seen as a relevant state variable, which affects parameters such as the number of advertised positions in hospitals and the content of (and hence the popularity of) hospital residency programs.

2. Other Complementarities. As mentioned in Section I, a couple preference is a particular form of complementarity, and this article can be put in the context of the research program on the role of complementarities in resource allocation. Examples could include groups of workers who care about externalities and firms that need a team of workers with complementary skills, among others. Given our analysis, a natural question is whether our asymptotic existence result of stable matchings continues to hold in the presence of these and other forms of complementarities.

To study this issue, consider a model of firm-worker matching, where each firm can hire up to κ workers and each worker can work for at most one firm, where κ is a constant. Suppose that a small fraction of firms draw nonsubstitutable preferences over workers, whereas others draw substitutable preferences that satisfy the law of aggregate demand. Moreover, assume regularity of the sequence of random markets similar to Definition 1. The couple matching is the case with $\kappa = 2$, where we relabel firms and workers as doctors and hospitals, respectively (think of each couple as a firm needing to hire two positions).

Our asymptotic existence result extends to the model already described. To see why, consider a variant of the sequential couples algorithm, which places firms with substitutable preferences first and then places firms with nonsubstitutable preferences one by one. This algorithm succeeds if no firms apply to workers where other firms with nonsubstitutable preferences are already tentatively placed. As in the case of couples, as long as the fraction of firms with nonsubstitutable preferences shrinks sufficiently fast, each firm has a capacity bounded by a constant κ, and each firm finds up to a constant number of workers to be acceptable, then the probability of a success can be shown to converge to 1 by the same argument as the one for Theorem 1. In fact, Ashlagi, Braverman, and Hassidim (2011) consider a similar model with complementarities and present a convergence result (see their Theorem 5).

VII. Conclusion

This article contributes toward understanding the consequences of the complementarities caused by couples in matching markets, a phenomenon that has grown in importance as dual-career households have become a significant part of the labor force. We investigate this issue by studying couples in centralized labor market clearinghouses. Even though a stable matching does not necessarily exist when couples are present, as long as the complementarities caused by couples are small in an appropriate way, our main result is that the market has a stable matching with a high probability. More broadly, our study suggests that not only does large market analysis help understand economies with indivisibilities, it also generates new kinds of results in markets with complementarities, which have been challenging for existing approaches.

We have complemented our theoretical results with analysis of data from the market for psychologists. The stylized facts from the data motivate some of the modeling assumptions. In every year of the data we are able to find a stable matching with respect to the stated preferences. Our simulations show that a stable matching is more likely to be found in large markets. Because the mechanism we analyze is similar to the actual procedures used in markets such as the NRMP for U.S. medical residents, our results help explain why some mechanisms in practice provide a stable matching with high probability despite the presence of couples.

There are a number of additional questions motivated by this article. One question is whether stability itself is the reason for the enduring success of the NRMP and postdoctoral psychology market. Field and laboratory evidence suggests that stability is responsible for the persistence of certain centralized clearinghouses (Roth 1991; Kagel and Roth 2000). Within the context of couples, however, an alternative may simply be to consider a weaker approximate notion of stability, such as a requirement that the number of blocking coalitions be small. Under our assumptions it is obvious that there always exists a matching that is approximately stable in that sense.[38] What is more remarkable and interesting is that the markets we study have exactly stable matchings, and our analysis provides conditions for this fact.

Another question involves the interpretation of preferences. As we have emphasized, the analysis we undertake is after applicants interview for positions. This is perhaps the major reason why applicants' rank order lists are short in a large market. A richer, but substantially different, model could consider a two-stage game where participants have imperfect information about their preferences and first decide where to interview. This type of analysis could provide a way to endogenize the short rank order lists of applicants. While interesting, we expect this sort of analysis to first focus on the decision problem of where to apply, before adding the complications of how to participate in a matching market with couples.[39] Loosely speaking, we would expect that in such a model, most applicants would apply to many positions, and many hospitals would have more applicants than they can interview, while some applicants might receive more

38. To see this point, consider a matching that is stable in the submarket composed of hospitals and single doctors only while keeping all couples unmatched. Clearly the number of coalitions that may block this matching is at most the sum of the lengths of the rank order lists over all couples, which is small in large markets under our maintained assumptions.

39. Related to this, another issue is to examine whether couples may have an incentive to manipulate by pretending to be singles or vice versa (as in Klaus, Klijn, and Masso 2007), or even whether a dual-career joint location problem encourages or discourages doctors from marrying other doctors (see Hurder 2013).

interview invitations than they can or feel a need to accept, so that a good deal of sorting among doctors and hospitals would take place even before interviews were conducted.

A further topic for research is how *decentralized* markets might be organized to handle couples better. For instance, Niederle and Roth (2009a, 2009b) study how the rules regarding exploding offers influence market outcomes. The issues here would involve the formal and informal rules by which couples search for two positions and by which offers and responses are made, so as to increase the efficiency of the market in finding matches when some applicants are looking for pairs of positions.

In summary, labor markets in which the pool of applicants includes two-career households have proved challenging to study even as they have become more common and have demanded adaptation in labor market rules and institutions. Although many open questions remain, the results of the present article suggest that some of the potential problems that couples and market designers face may become more tractable in large markets.

Department of Economics, Stanford University
Department of Economics, Massachusetts Institute of Technology, and NBER
Department of Economics, Stanford University

Supplementary Material

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez, “School Choice: A Mechanism Design Approach,” *American Economic Review*, 93 (2003), 729–747.

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda, “Expanding Choice in School Choice,” (Economic Research Initiatives at Duke Research Paper No. 20, 2008).

Alcalde, José, and Salvador Barberà, “Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems,” *Economic Theory*, 4 (1994), 417–435.

Ashlagi, Itai, Mark Braverman, and Avinatan Hassidim, “Stability in Large Matching Markets with Complementarities,” (unpublished mimeo, MIT Sloan School, 2011).

Azevedo, Eduardo M., and Jacob D. Leshno, “A Supply and Demand Framework for Two-Sided Matching Markets,” (unpublished manuscript, Wharton, 2012).

Bikhchandani, S., and J. M. Ostroy, "The Package Assignment Model," *Journal of Economic Theory*, 107, no. 2 (2002), 377–406.
Biró, P., and R. W. Irving, "Stable Matching with Couples—An Empirical Study," (Technical Report TR-2010-319, University of Glasgow, 2010).
Budish, Eric, "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," *Journal of Political Economy*, 119, no. 6 (2011), 1061–1103.
Bulow, Jeremy, and Jonathan Levin, "Matching and Price Competition," *American Economic Review*, 96 (2006), 652–668.
Che, Yeon-Koo, and Fuhito Kojima, "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms," *Econometrica* (2010), 1625–1672.
Costa, Dora, and Matthew Kahn, "Power Couples: Changes in the Locational Choice of the College Educated, 1940–1990," *Quarterly Journal of Economics*, 115 (2000), 1287–1315.
Dean, Brian C., Michael X. Goemans, and N. Immorlica, "The Unsplittable Stable Marriage Problem," *Proceedings of the 4th IFIP International Conference on Theoretical Computer Science* 2006, pp. 65–76.
Dutta, B., and J. Massó, "Stability of Matchings When Individuals Have Preferences over Colleagues," *Journal of Economic Theory*, 75 (1997), 464–475.
Echenique, Federico, and Jorge Oviedo, "A Theory of Stability in Many-to-Many Matching," *Theoretical Economics*, 1 (2006), 233–273.
Ehlers, Lars, "In Search of Advice for Participants in Matching Markets Which Use the Deferred Acceptance Algorithm," *Games and Economic Behavior*, 48 (2004), 249–270.
———, "Truncation Strategies in Matching Markets," *Mathematics of Operations Research*, 33 (2008), 327–335.
Ehlers, Lars, and Aytek Erdil, "Efficient Assignment Respecting Priorities," *Journal of Economic Theory*, 145 (2010), 1269–1282.
Erdil, Aytek, and Haluk Ergin, "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice," *American Economic Review*, 98 (2008), 669–689.
Ergin, Haluk, "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 70 (2002), 2489–2498.
Gale, David, and Lloyd S. Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69 (1962), 9–15.
Gresik, Thomas, and Mark Satterthwaite, "The Rate at Which a Simple Market Converges to Efficiency as the Number of Traders Increases," *Journal of Economic Theory*, 48 (1989), 304–332.
Gul, Faruk, and Ennio Stacchetti, "Walrasian Equilibrium with Gross Substitutes," *Journal of Economic Theory*, 87 (1999), 95–124.
———, "The English Auction with Differentiated Commodities," *Journal of Economic Theory*, 92 (2000), 66–95.
Haeringer, Guillaume, and Flip Klijn, "Constrained School Choice," *Journal of Economic Theory*, 144 (2009), 1921–1947.
Hatfield, John William, and Fuhito Kojima, "Matching with Contracts: Comment," *American Economic Review*, 98 (2008), 1189–1194.
———, "Substitutes and Stability for Matching with Contracts," *Journal of Economic Theory*, 145 (2010), 1703–1723.
Hatfield, John William, and Scott Duke Kominers, "Many-to-Many Matching with Contracts," (Mimeo: Stanford Graduate School of Business and Becker Friedman Institute, 2009).
Hatfield, John William, and Paul Milgrom, "Matching with Contracts," *American Economic Review*, 95 (2005), 913–935.
Helppie, Brooke, and Marta Murray-Close, "Moving Out or Moving Up? New Economists Sacrifice Job Opportunities for Proximity to Significant Others – and Vice Versa," (Working paper, University of Michigan, 2010).
Hurder, Stephanie, "An Integrated Model of Occupation Choice, Spouse Choice, and Family Labor Supply," Harvard University, Department of Economics, 2013, available at http://www.people.fas.harvard.edu/~hurder/papers/SHurderJMP.pdf.
Immorlica, Nicole, and Mohammad Mahdian, "Marriage, Honesty, and Stability," *SODA Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete*

Algorithms (Philadelphia: Society for Industrial and Applied Mathematics, 2005), pp. 53–62.

Jackson, Matthew O., and Alejandro M. Manelli, "Approximately Competitive Equilibria in Large Finite Economies," *Journal of Economic Theory*, 77 (1997), 354–376.

Jonas, Harry S., and Sylvia I. Etzel, "Undergraduate Medical Education," *Journal of the American Medical Association*, 260 (1998), 1063–1071.

Kagel, John, and Alvin E. Roth, "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Field Experiment," *Quarterly Journal of Economics*, 115 (2000), 201–235.

Keilin, W. Gregory, "Internship Selection 30 Years Later: An Overview of the APPIC Matching Program," *Professional Psychology: Research and Practice*, 29 (1998), 599–603.

———, "Internship Selection in 1999: Was the Association of Psychology Postdoctoral and Internship Centers' Match a Success?," *Professional Psychology: Research and Pratice*, 31 (2000), 281–287.

Kesten, Onur, "On Two Competing Mechanisms for Priority-Based Allocation Problems," *Journal of Economic Theory*, 127 (2006), 155–171.

———, "School Choice with Consent," *Quarterly Journal of Economics*, 125 (2010), 1297–1348.

———, "On Two Kinds of Manipulation for School Choice Problems," *Economic Theory*, 51 (2012), 677–693.

Klaus, Bettina, and Flip Klijn, "Stable Matchings and Preferences of Couples," *Journal of Economic Theory*, 121 (2005), 75–106.

Klaus, Bettina, Flip Klijn, and Jordi Masso, "Some Things Couples Always Wanted to Know about Stable Matchings (But Were Afraid to Ask)," *Review of Economic Design*, 11 (2007), 175–184.

Klaus, Bettina, Flip Klijn, and Toshifumi Nakamura, "Corrigendum: Stable Matchings and Preferences of Couples," *Journal of Economic Theory*, 144 (2009), 2227–2233.

Knuth, Donald E., Rajeev Motwani, and Boris Pittel, "Stable Husbands," *Random Structures and Algorithms*, 1 (1990), 1–14.

Kojima, Fuhito, "When Can Manipulations Be Avoided in Two-Sided Matching Markets? Maximal Domain Results," *Contributions to Theoretical Economics*, 7 (2007), article 32.

Kojima, Fuhito, and Mihai Manea, "Incentives in the Probabilistic Serial Mechanism," *Journal of Economic Theory*, 145 (2010), 106–123.

Kojima, Fuhito, and Parag A. Pathak, "Incentives and Stability in Large Two-Sided Matching Markets," *American Economic Review*, 99 (2009), 608–627.

Kojima, Fuhito, Parag A. Pathak, and Alvin E. Roth, "Matching with Couples: Stability and Incentives in Large Markets," (NBER Working Paper 16028, 2010).

Konishi, Hideo, and M. Utku Ünver, "Credible Group Stability in Many-to-Many Matching Problems," *Journal of Economic Theory*, 129 (2006a), 57–80.

———, "Games of Capacity Manipulation in the Hospital-Intern Market," *Social Choice and Welfare*, 27 (2006b), 3–24.

Lee, SangMok, "Incentive Compatibility of Large Centralized Matching Markets," (unpublished manuscript, University of Pennsylvania, 2013).

Liu, Qingmin, and Marek Pycia, "Ordinal Efficiency, Fairness, and Incentives in Large Markets," (UCLA working paper, 2011).

Manea, M., "Asymptotic Ordinal Inefficiency of Random Serial Dictatorship," *Theoretical Economics*, 4, no. 2 (2009), 165–197.

McDermid, Eric J., and David Manlove, "Keeping Partners Together: Algorithmic Results for the Hospitals/Resident Problem with Couples," *Journal of Combinatorial Optimization*, 19 (2009), 279–303.

Milgrom, Paul R. *Putting Auction Theory to Work* (Cambridge: Cambridge University Press, 2004).

Niederle, Muriel, and Alvin E. Roth, "The Effects of a Central Clearinghouse on Job Placement, Wages, and Hiring Practices," in *Labor Market Intermediation*, David Autor, ed. (Chicago: The University of Chicago Press, 2009a), 235–271.

———, “Market Culture: How Rules Governing Exploding Offers Affect Market Performance,” *American Economic Journal: Microeconomics*, 1 (2009b), 199–219.

NRMP (National Resident Matching Program), “Results and Data: 2009 Main Residency Match,” 2009, available at http://www.nrmp.org/data/resultsand data2009.pdf (accessed March 15, 2010).

Ostrovsky, Michael, “Stability in Supply Chain Networks,” *American Economic Review*, 98 (2008), 897–923.

Pathak, Parag A., and Tayfun Sönmez, “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism,” *American Economic Review*, 98 (2008), 1636–1652.

Pycia, Marek, “Stability and Preference Alignment in Matching and Coalition Formation,” *Econometrica*, 80 (2012), 323–362.

Rich, E. C., M. Liebow, M. Srinivasan, D. Parish, J. O. Wolliscroft, O. Fein, and R. Blaser, “Medicare Financing of Graduate Medical Education,” *Journal of General Internal Medicine*, 17, no. 4 (2002), 283–292.

Roberts, Donald John, and Andrew Postlewaite, “The Incentives for Price-Taking Behavior in Large Exchange Economies,” *Econometrica*, 44 (1976), 115–127.

Ronn, E., “NP-Complete Stable Matching Problems,” *Journal of Algorithms*, 11, no. 2 (1990), 285–304.

Roth, Alvin E., “The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory,” *Journal of Political Economy*, 92 (1984), 991–1016.

———, “The College Admission Problem Is Not Equivalent to the Marriage Problem,” *Journal of Economic Theory*, 36 (1985), 277–288.

———, “A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.,” *American Economic Review*, 81 (1991), 415–440.

———, “The Origins, History, and Design of the Resident Match,” *Journal of the American Medical Association*, 289 (2003), 909–912.

———, “Deferred Acceptance Algorithms: History, Theory, Practice and Open Questions,” *International Journal of Game Theory*, 36 (2007), 537–569.

———, “What Have We Learned from Market Design?,” in *The Handbook of Market Design*, Nir Vulkan, Alvin E. Roth, and Zvika Neeman, eds. (Oxford: Oxford University Press, 2013).

Roth, Alvin E., and Elliott Peranson, “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design,” *American Economic Review*, 89 (1999), 748–780.

Roth, Alvin E., and Uriel Rothblum, “Truncation Strategies in Matching Markets: In Search of Advice for Participants,” *Econometrica*, 67 (1999), 21–43.

Roth, Alvin E., and Marilda A. O. Sotomayor *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* (Cambridge: Econometric Society Monographs, 1990).

Roth, Alvin E., and John H. Vande Vate, “Random Paths to Stability in Two-Sided Matching,” *Econometrica*, 58 (1990), 1475–1480.

Roth, Alvin E., and Xiaolin Xing, “Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions,” *American Economic Review*, 84 (1994), 992–1044.

———, “Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists,” *Journal of Political Economy*, 105 (1997), 284–329.

Rustichini, Aldo, Mark Satterthwaite, and Steven Williams, “Convergence to Efficiency in a Simple Market with Incomplete Information,” *Econometrica*, 62 (1994), 1041–1064.

Sönmez, Tayfun, and M. Utku Ünver, “Matching, Allocation, and Exchange of Discrete Resources,” in *Handbook of Social Economics*, Jess Benhabib, Alberto Bisin, and Matthew Jackson, eds. (Amsterdam: Elsevier, 2009).

———, “Course Bidding at Business Schools,” *International Economic Review*, 51, no. 1 (2010), 425–445.

Sotomayor, Marilda A. O., “Three Remarks on the Many-to-Many Stable Matching Problem,” *Mathematical Social Sciences*, 38 (1999), 55–70.

———, “Implementation in the Many-to-Many Matching Market,” *Games and Economic Behavior*, 46 (2004), 199–212.
Sun, Ning, and Zaifu Yang, “Equilibria and Indivisibilities: Gross Substitutes and Complements,” *Econometrica*, 74 (2006), 1385–1402.
———, “Double-Track Adjustment Process for Discrete Markets with Substitutes and Complements,” *Econometrica*, 77 (2009), 933–952.