MECANISMOS DE ADMISSÃO DE CANDIDATOS ÀS INSTITUIÇÕES. MODELAGEM E ANÁLISE À LUZ DA TEORIA DOS JOGOS*

Marilda Sotomayor**

Resumo

Este trabalho é motivado pela problemática detectada na admissão de candidatos às escolas de pós-graduação em Economia no Brasil. Da análise dos fenômenos observados, à luz da Teoria dos Jogos, concluímos que o insucesso dos procedimentos de admissão de candidatos, até então usados, se deve à descentralização do mercado e à instabilidade dos resultados produzidos. Esta instabilidade é causada quando algum candidato, com média superior a de algum outro candidato designado para uma dada Escola, preferiria ter sido designado para essa Escola. As restrições impostas pelo mercado sugerem um modelo não cooperativo com ausência de comportamento estratégico por parte das Escolas. Para os mercados deste tipo mostramos que dentre todos os mecanismos centralizados de admissão de candidatos à Instituições, que produzem resultados estáveis, existe um e um único que incentiva os participantes a declarar suas verdadeiras preferências. Para qualquer outro caso são analisadas as possibilidades estratégicas dos candidatos e vários resultados teóricos são apresentados.

* Este trabalho foi apoiado parcialmente pela J.S. Guggenheim Memorial Foundation e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico.

** Instituto de Economia Industrial, Universidade Federal do Rio de Janeiro, Av. Pasteur 250, Urca, Rio de Janeiro, RJ.

Agradeço aos professores Renato Fragelli Cardoso e Marco Antonio Cesar Bonomo pelo fornecimento das informações contidas na seção 1.1.

Abstract

This work is motivated by the problems detected in the admission of candidates to graduate Economic Schools in Brazil. From the game theoretic analysis of the observed phenomena, we concluded that the failure of the admission procedures, which has been used up to now, is due to the decentralization of the market and to the instability of the outcomes. This instability is verified when some candidate, c, with a score higher than the score of other candidate assigned to a given school, would prefer to have been assigned to that school. The market restrictions suggest a non-cooperative model in which schools behave straightforwardly. For these kind of markets we show that there is one and only one centralized institution admission mechanism, which yields stable outcomes and incentivates the participants to declare their true preferences, among all such mechanisms. In any other case the strategic possibilities of the candidates are analysed and several theoretic results are presented.

Palavras-Chave: matching estável, estratégia dominante, mecanismos à prova de estratégia

Código JEL: C78

1. Introdução.

O problema da admissão de candidatos à Instituições foi introduzido na literatura em Gale e Shapley (1962) com o nome de "College Admission Problem". São dados um conjunto de Instituições e um conjunto de candidatos, tais como firmas e trabalhadores, universidades e estudantes, médicos e hospitais. Cada candidato lista, em ordem de preferência, aquelas Instituições nas quais desejaria ingressar e cada Instituição lista, em ordem de preferência, aqueles candidatos que ela gostaria de admitir. As Instituições têm cotas, representando o número de vagas disponíveis. Nenhum candidato pode ser admitido em mais de uma Instituição. O problema é encontrar um processo de admissão dos candidatos às Instituições que respeite as preferências de cada participante e as cotas de cada Instituição. Uma tal distribuição é denominada um matching. A noção chave é a de estabilidade. *Um matching exibe uma instabilidade se* *existe um candidato potencial c para alguma Instituição I, tal que c ou não foi designado para nenhuma Instituição ou, se o foi, prefere I à Instituição para a qual foi designado; por sua vez, I ou não preencheu todas as suas vagas ou, se o fez, prefere c a algum dos candidatos designados a ela. Um matching é chamado estável se não exibe instabilidades.*

A existência de matchings estáveis foi demonstrada por Gale e Shapley por meio de um algoritmo, que partindo das preferências dadas chega a um matching estável, num número finito de etapas, que é proporcional ao quadrado do número de candidatos e Instituições. O matching obtido por este modo tem a propriedade de ser o único matching (podem existir muitos) que é o mais preferido por todos os candidatos, dentre todos os matchings estáveis. Este matching é chamado matching estável ótimo para os candidatos. Invertendo-se os papéis, no algoritmo, entre candidatos e Instituições, obtém-se o matching estável ótimo para as Instituições.

Independentemente de Gale e Shapley, a National Resident Matching Programm - NRMP, localizada em Evanston, Illinois, desenvolveu, por um processo de ensaio e erro, ao longo de quase meio século, um procedimento para a distribuição de médicos residentes nos hospitais dos Estados Unidos. A evolução deste processo deu origem, em 1951, a um *mecanismo de matching*, isto é, *uma regra que associa um matching a qualquer conjunto de candidatos e Instituições, com suas cotas e respectivas preferências*. Este mecanismo, que é usado até hoje, utiliza exatamente o algoritmo proposto por Gale e Shapley. A única diferença é que, em vez do matching estável ótimo para os candidatos, ele produz o ótimo para os hospitais.

Algoritmo Gale-Shapley

Cada candidato se candidata à admissão na sua Instituição favorita, isto é, na primeira Instituição de sua lista. As Instituições com excesso de candidatos aceitarão, temporariamente, os melhores candidatos que preencherem as suas vagas e rejeitarão os demais.

Esta é a primeira etapa do algoritmo. Na segunda etapa, os candidatos rejeitados se candidatam à admissão na sua segunda Instituição favorita e novamente as Instituições com excesso de candidatos preencherão as suas vagas, temporariamente, com os melhores candidatos e rejeitarão os demais, etc. O algoritmo terminará quando todo candidato for admitido temporariamente ou tiver sido rejeitado por todas as Instituições de sua lista.

Algoritmo NRMP

Cada Instituição I, com q_I vagas, faz uma oferta de admissão aos q_I candidatos mais preferidos de sua lista. Os candidatos que receberem ofertas de mais do que uma Instituição aceitarão, temporariamente, aquela que eles preferirem e recusarão as outras. Seus nomes são então removidos das listas de todas as Instituições cujas ofertas tenham sido recusadas por eles. Esta é a primeira etapa do algoritmo. As Instituições que não tiverem preenchido suas vagas, novamente, farão uma oferta ao conjunto de candidatos mais preferidos de sua lista de forma a não ultrapassar a sua cota. Os candidatos que receberem novas ofertas, novamente rejeitarão todas, exceto a sua favorita. Esta é a segunda etapa, etc. O algoritmo terminará quando, depois de alguma etapa, nenhuma Instituição puder fazer mais ofertas, ou porque sua cota estiver preenchida ou porque ela já tiver esgotado sua lista.

Os mais simples exemplos (dois candidatos e duas Instituições) mostram que os dois algoritmos acima não necessitam dar o mesmo resultado. Além disso, o matching estável ótimo para os candidatos (resp. Instituições) é o pior matching estável do ponto de vista das Instituições (resp. dos candidatos). (Ver Roth e Sotomayor, 1990).

Um mecanismo de matching, para um mercado de candidatos e Instituições, induz um jogo na forma estratégica, segundo a terminologia usada em Teoria dos Jogos. Os jogadores são os candidatos e as Instituições; o conjunto de estratégias de cada jogador são todas as possíveis listas ordenadas de preferências deste jogador sobre os jogadores de lado oposto. Em cada jogada, cada participante se depara com a situação de decidir que estratégia selecionar. Como se espera que um jogador deva se comportar quando é chamado a tomar suas decisões e a explicação desse comportamento são o assunto de interesse central da Teoria dos Jogos não cooperativos.

Tendo cada participante escolhido sua estratégia, o mecanismo, então, produz um matching. Naturalmente os jogadores avaliam matchings em termos de suas verdadeiras preferências. São estas verdadeiras preferências que determinam seu comportamento.

Em 1981 Dubins e Freedman provaram que o mecanismo de matching estável ótimo para os candidatos é à prova de estratégia para os candidatos, mas não para as Instituições. Isto significa o seguinte: Suponha que todas as Instituições e candidatos tenham informação completa sobre as preferências (estratégias) uns dos outros. A questão é então saber se algum grupo de participantes poderia tirar vantagem disto, usando uma estratégia diferente de suas verdadeiras preferências. Dubins e Freedman mostraram que isto é possível para as Instituições mas não para os candidatos. Mais precisamente, nenhum conjunto de candidatos pode, adulterando suas preferências, forçar um matching que seja preferido por todos os candidatos no conjunto.[1] Em 1985 Roth mostrou que o mecanismo estável ótimo para as Instituições não é à prova de estratégia, nem para as Instituições nem para os candidatos. Que não existe um mecanismo estável à prova de estratégia para os candidatos e Instituições foi provado por Roth em 1982 supondo domínio irrestrito das preferências dos jogadores. De lá para cá, encontrar condições que garantam a existência de um mecanismo estável à prova de estratégia, tem sido objeto de interesse de alguns autores. Nesta direção podemos citar Alcalde e Barberà (1994) que provaram que se o domínio das preferências das Instituições for restrito a uma certa classe de preferências, o mecanismo estável ótimo para os candidatos

[1] A demonstração original deste resultado é bastante longa (cerca de 20 páginas); uma demonstração bem mais curta é apresentada em Gale e Sotomayor (1985a)

é à prova de estratégia.

No presente trabalho modelamos mercados em que, diferentemente da abordagem até agora usada, as restrições sobre o domínio estratégico das Instituições não são uma hipótese do modelo, mas uma conseqüência de como o mercado é organizado. Essas restrições decorrem do fato de que existem regras segundo as quais as listas de preferências das Instituições sobre os candidatos são construídas. Desta forma o domínio estratégico das Instituições fica restrito às suas "verdadeiras preferências". Em geral os candidatos são listados de acordo com os graus obtidos num conjunto de testes. Se são atribuídos pesos a cada um dos testes, cujos valores variam de uma Instituição para outra, as listas de preferências das Instituições sobre os candidatos podem diferir, também, uma das outras. Caso contrário, a classificação dos candidatos é a mesma para todas as Instituições, produzindo uma única lista ordenada de preferências.

Os mercados de admissão dos candidatos às universidades no Brasil e em outros países sul-americanos, dos economistas às Escolas de pós-graduação em Economia no Brasil, dos professores primários às Escolas Primárias públicas do Estado do Rio de Janeiro, entre outros, se incluem na classe dos mercados mencionados acima. Quando uma grande firma abre um concurso para o preenchimento das vagas em suas filiais, a situação de jogo que se configura é, também, a de um "College Admission" com restrição de domínio estratégico das Instituições.

Em qualquer dessas situações nenhuma Instituição é chamada a decidir sobre qual lista de preferências selecionar, de forma que não tem sentido esperar um comportamento estratégico por parte das Instituições. Portanto o jogo induzido pelo mecanismo de matching usado nesses mercados é estruturalmente diferente do jogo induzido pelo NRMP. O nosso primeiro resultado torna patente esta diferença: O Teorema da Impossibilidade de Roth não se aplica aqui. De fato, para a classe de mercados mencionada acima, o Teorema de Dubins e Freedman implica que o mecanismo que produz o matching estável para os candidatos é à prova de estratégia, para todos os jogadores.

A unicidade de um mecanismo à prova de estratégia, bem como o estudo das possibilidades estratégicas dos candidatos face a um mecanismo manipulável de matchings estáveis, constituem o propósito deste trabalho.[2]

Embora não tenhamos por objetivo relatar observações empíricas sobre os mercados que estamos estudando, estamos voltados para o desenvolvimento de uma teoria capaz de explicar tais observações.

Com este intuito, a próxima seção desta introdução será devotada à descrição de um conjunto de observações que servirão para motivar o material apresentado aqui. Elas se referem ao mercado de admissão dos economistas às Escolas de pós-graduação em Economia no Brasil. Neste mercado os participantes estão presentes ao mesmo tempo, o que o torna factível de ser analisado com o modelo desenvolvido aqui.

Esperamos que esse exemplo ilustre, mais uma vez, o uso da Teoria dos Jogos para explicar e predizer o comportamento de sistemas econômicos do "mundo real".

Este trabalho é organizado como segue: Na seção 2 descrevemos o modelo cooperativo, apresentamos as notações e definições necessárias, bem como enunciamos alguns resultados já conhecidos e demonstramos um novo resultado referente a este modelo. Na seção 3 descrevemos o modelo não-cooperativo e demonstramos a existência e unicidade de um mecanismo estável coletivamente à prova de estratégia. Na seção 4 apresentamos e demonstramos os novos resultados referentes às possibilidades estratégicas dos candidatos. Um exemplo é dado a título de ilustração.

---

[2] Diversos autores têm estudado as possibilidades estratégicas de um dos lados do mercado, face a um mecanismo de matching ótimo para o outro lado, no caso em que os jogadores têm cota unitária. Dentre eles citamos: Gale e Sotomayor (1985a, 1985b), Demange e Gale (1985), Demange, Gale e Sotomayor (1987) e Roth (1982, 1984).

1.1. O Problema da Admissão de Economistas às Escolas de Pós-graduação no Brasil.

Um Pouco de História Institucional

Todos os anos, numa data determinada pela Associação Nacional de Instituições de Pós-graduação - ANPEC, um número entre 400 e 500 graduados em Economia realizam um exame escrito sobre cinco disciplinas: Matemática, Estatística, Microeconomia, Macroeconomia e Economia Brasileira. Esta etapa constitui pré-requisito obrigatório para a admissão numa Escola de Pós-graduação em Economia no Brasil.

As Escolas atribuem pesos a cada uma das cinco disciplinas, que podem diferir de uma Instituição para outra. Por exemplo, a Universidade de Campinas e a Universidade Federal do Rio de Janeiro costumam atribuir à Economia Brasileira um peso maior que o da Matemática; a Escola de Pós-graduação em Economia da Fundação Getúlio Vargas faz justamento o oposto.

Até 1992, ao se inscrever, o candidato submetia uma lista com os nomes de 4 Instituições em ordem de preferência (estrita). Após a realização e correção das provas, cada Instituição recebia uma lista com as médias ponderadas dos candidatos que a tinham escolhido como primeira opção, de acordo com os pesos por ela atribuídos. Portanto, a Escola deveria preencher as suas vagas, cujo número variava de 10 a 20, com os candidatos por ela preferidos, dentre os que a tinham escolhido em primeira opção.

Acontecia que, enquanto as Instituições maiores, mais procuradas, sempre conseguiam preencher as suas vagas, o mesmo não acontecia com as Instituições menores. Estas últimas eram, assim, levadas a fazer suas escolhas entre os candidatos que as tivessem escolhido em segunda, terceira ou quarta opções. Além disso, um outro tipo de problema apareceu, decorrente do fato de algumas Escolas desconsiderarem os candidatos que não as tivessem escolhido em primeira opção e consequentemente ignorarem as médias desses alunos. Basicamente, o problema era que, o candidato que tivesse optado somente por grandes Instituições e tivesse sido preterido por sua Escola preferida, não seria escolhido por nenhuma outra Instituição, mesmo que o seu desempenho nos exames tivesse sido melhor do que o de qualquer candidato escolhido por essas outras Instituições. Esse candidato poderia prejudicar-se mais ainda, caso não fosse possível fazer arranjos subsequentes com outras Instituições.

Para contornar um tal problema, o candidato deveria listar, como primeira opção, a Escola preferida dentre aquelas que fossem alcançáveis para ele, tendo em conta o desempenho que esperava ter nos exames e a competição com os outros candidatos. Este procedimento produzia situações nas quais o candidato, que subestimara seu desempenho, deixava de ser escolhido por uma Escola preferível, por não tê-la listado em primeira opção. Em uma tal situação, Escola e candidato ficavam descontentes.

Problemas da ordem descrita acima levaram a ANPEC a mudar o procedimento em 1992. Os candidatos passaram a indicar os nomes de 3 Instituições, sem contudo revelar suas prioridades; depois de conhecidas as médias ponderadas dos alunos, cada Instituição procurava atrair para si os melhores candidatos que poderiam preencher as suas vagas, dentre aqueles que a tinham indicado em sua lista.

Devido aos casos registrados em 1992 e 1993, de estudantes com nível médio de rendimento nos exames que não foram escolhidos por nenhuma Escola, e de Escolas que não preencheram todas as suas vagas, o número de Instituições indicadas pelos estudantes aumentou para 6, a partir de 1994. Esta medida, que não resolveu o problema, foi considerada de caráter provisório, enquanto se procurava por uma solução mais adequada.

Um novo problema, porém, surgido com a mudança de 1992, vem tomando vulto nestes últimos anos. Tem como causa o tempo de espera entre o momento em que o candidato recebe uma oferta e o momento em que ele deverá decidir se a aceita ou não: Se um candidato recebe uma oferta, digamos, de sua terceira Escola preferida, e é informado que está na lista de espera de sua segunda escolha, então ele estará inclinado a esperar, tanto quanto possível, antes de aceitar a proposta feita, na esperança de, eventualmente, lhe ser oferecida a Escola preferível. No entanto se ele é pressionado a aceitar a oferta antes que o seu "status" na lista de espera de sua segunda escolha seja resolvido, ficará satisfeito se lhe for posteriormente oferecida a Escola preferível. Por outro lado, se ele esperou até o último momento para recusar a oferta de sua terceira Escola preferida, esta ficará insatisfeita se nesse meio tempo o seu candidato preferido da lista de espera tiver se comprometido com outra Escola.

Uma Escola ficaria descontente, ainda, quando um candidato que a tivesse aceito falhasse no seu compromisso, depois de ter recebido uma oferta preferível ou quando, para garantir o preenchimento de suas vagas, tivesse feito ofertas a alunos menos preferidos ...

Todos esses problemas têm sido reconhecidos amplamente pelas Escolas e pela ANPEC mas nenhuma solução foi ainda proposta.

Análise, Conclusões Preliminares e Perguntas a serem Respondidas

O insucesso do procedimento usado pela ANPEC antes de 1992 e a operação desorganizada do mercado nos anos de 92, 93, 94 e 95 são os principais fenômenos observados neste mercado e o entendimento de suas causas é o primeiro passo em direção a adoção de um procedimento duradouro e satisfatório, do ponto de vista dos participantes.

A análise desses fenômenos é muito simples. A existência de um tráfico telefônico entre Escolas e candidatos, algumas vezes voltando atrás em seus acordos verbais, outras, firmando novos acordos, tinha como causa o fato de que os candidatos não eram uniformemente informados sobre quais Escolas já tinham preenchido as suas vagas e as Escolas não eram também uniformemente informadas se um dado candidato ainda estava disponível. Podemos concluir, desta forma, que a desorganização observada nos anos de 93 e 94, no processo de admissão dos candidatos, originara-se da *descentralização do mercado*.

Quanto ao insucesso do procedimento antes de 92, a base de nossa explicação é que o procedimento usado produzia matchings instáveis, devido ao fato dos estudantes submeterem uma lista incompleta de Instituições aceitáveis. Num procedimento que produz resultados instáveis, os pares, candidato-escola, insatisfeitos, farão melhor se ignorarem o procedimento e fizerem seus arranjos independemente. A persistência de tais situações tem como conseqüência o descrédito por parte dos participantes, o que compromete a reputação do sistema e, consequentemente, incentiva a substituição do procedimento usado por outro. (Veja uma ilustração desta análise no final da seção 4)

A descentralização do mercado nos procedimentos de 92 a 95 e a limitação do número de Instituições indicadas pelos candidatos são também responsáveis pela instabilidade dos matchings obtidos nesses anos.

Desta forma, a adoção de um mecanismo centralizado que produza matchings estáveis em relação às preferências submetidas parece ser a solução para a problemática detectada neste mercado. Entretanto, um problema que poderia ocorrer é que um agente bem informado poderia ser capaz de, adulterando suas preferências, influenciar o resultado final em benefício próprio. Ora, se alguns participantes adulterarem suas preferências, então o resultado obtido, embora estável segundo as preferências submetidas, poderá não ser estável com relação às preferências verdadeiras. (Um mecanismo que incentiva os agentes a sempre revelar suas verdadeiras preferências é chamado *não manipulável* ou *coletivamente à prova de estratégia*. Estas considerações sugerem que a solução do problema é *um mecanismo estável à prova de estratégia*.

A existência de um mecanismo estável à prova de estratégia para o mercado que estamos analisando, decorre do fato de que as Escolas submetem o seu vetor de pesos antes dos candidatos realizarem seus exames. Então, sem conhecer o resultado dos testes, as Escolas não têm como deduzir suas preferências sobre os seus futuros candidatos. Assim, se, por exemplo, o verdadeiro vetor de pesos de uma Escola é π, ela não tem nenhuma base racional para acreditar que revelar π' em vez de π será mais vantajoso para ela. Consequentemente podemos contar que cada Escola sempre revelará o seu verdadeiro vetor de pesos e não terá sentido esperar um comportamento estratégico por parte delas. Assim, o Teorema de Dubins e Freedman implica que, para o mercado em questão, um mecanismo que produza o matching estável ótimo para os candidatos é à prova de estratégia. Isto significa dizer que se o algoritmo Gale-Shapley for usado, os candidatos não poderão fazer melhor do que submeter suas verdadeiras preferências e o matching resultante será estável com relação a estas preferências.

Neste ponto as seguintes perguntas se tornam naturais:

1) Será que o mecanismo que produz o matching estável ótimo para os candidatos é o único mecanismo estável à prova de estratégia?

2) A utilização de um mecanismo que produz o matching estável ótimo para os candidatos poderia "beneficiar" os candidatos em detrimento das Escolas?

3) Para onde irão os melhores alunos se um mecanismo estável for usado?

4) Desde que cabe à uma associação de Instituições decidir qual mecanismo adotar, poderia ser que fosse determinado que o procedimento a ser usado fosse o que produz o matching estável ótimo para as Escolas. Neste caso haveria algum conjunto de estratégias que os candidatos pudessem usar para reverter o resultado a seu favor?

5) Se a resposta à pergunta anterior é afirmativa, quando poderemos esperar que o matching produzido pelo mecanismo seja sempre estável em relação às verdadeiras preferências?

6) Em que condições é possível, para um candidato, manipular um mecanismo estável (que não produz o matching estável ótimo para os candidatos), em proveito próprio? Que tipo de informações ele necessita?

7) No caso em que manipular um mecanismo em proveito próprio exige uma quantidade muito grande de informação sobre as preferências dos outros participantes, que conselho se poderia dar a um candidato sobre a escolha de sua estratégia, caso o mecanismo produzisse o matching ótimo para as Instituições?

Uma Prévia da Explicação Proposta

Mostraremos no Teorema 2 que se *o matching produzido por um mecanismo estável não é o ótimo para os candidatos, então este mecanismo não será à prova de estratégia; haverá sempre algum candidato para o qual a honestidade nãoé a melhor política.* Este resultado responde à pergunta número 1.

Quanto à pergunta número 2, foi demonstrado em Gale e Sotomayor (1984), que o conjunto de candidatos não admitidos por nenhuma Escola é o mesmo, quer o matching utilizado seja o ótimo para os candidatos, o ótimo para as Escolas ou qualquer outro matching estável, e as Escolas preenchem a mesma fração de suas cotas em todos os matchings estáveis. Além disso, Roth provou em 1986 que se uma Escola não preencher todas as suas vagas em um dos matchings, ela será designada ao mesmo conjunto de candidatos em qualquer matching estável. Portanto a utilização do matching ótimo para os candidatos (resp. Escolas) não beneficia os alunos com médias mais baixas (resp. as Escolas menos procuradas) e não prejudica as Escolas mais procuradas (resp. os alunos com médias mais altas).

Mas para onde vão os melhores alunos? Veremos no Teorema 1 que *em qualquer matching estável, as Escolas mais procuradas preencherão sempre suas vagas e com os melhores candidatos.*

As respostas às perguntas restantes tem a ver com as possibilidades estratégicas dos candidatos e com a espécie de equilíbrio estratégico que pode resultar. Quanto à pergunta número 4, provaremos no Teorema 3, que *se o matching produzido não é o ótimo para os candidatos, estes sempre poderão forçar qualquer resultado estável que prefiram fracamente ao matching produzido pelo mecanismo, usando estratégias de equilíbrio.*[3] Em particular eles poderão obter o matching estável ótimo para o seu lado. Neste caso, o Teorema 5, deste trabalho, mostra que o matching estável ótimo para os candidatos pode ser alcançado usando-se estratégias de equilíbrio forte.[4]

Quanto à pergunta número 5 temos a dizer que, desde que os candidatos terão sempre um incentivo para adulterar suas preferências quando o matching produzido pelo mecanismo estável não for o ótimo para eles, é natural esperar que o matching resultante não seja sempre estável com respeito às verdadeiras preferências. No entanto, se esperarmos que equilíbrios ocorram, veremos, no Teorema 4, *que qualquer conjunto de estratégias de equilíbrio produz um matching estável com relação às verdadeiras preferências.* Assim, se algum conjunto de estratégias der a algum candidato uma escola que ele goste mais do que a que ele obteria no matching estável ótimo para os candidatos, o matching produzido não será estável e, por conseguinte, o conjunto de estratégias jogadas não será um equilíbrio. Então algum outro candidato poderá mudar o matching a seu favor escolhendo uma estratégia diferente.

Quanto à pergunta número 6, no caso em que o mecanismo não seleciona o matching estável ótimo para os candidatos, aconselhar um candidato a escolher uma estratégia de equilíbrio será inútil se ele não tiver as informações necessárias sobre as preferências dos outros jo-

---

[3] Um conjunto de estratégias, uma para cada jogador, é chamado um equilíbrio (ou equilíbrio de Nash) se nenhum jogador pode se beneficiar mudando a sua estratégia dado que os outros mantém as suas.

[4] Dizemos que um conjunto de estratégias, uma para cada jogador, forma um equilíbrio forte se é um equilíbrio e se nenhuma coalizão de jogadores pode alcançar um payoff melhor para todos os seus membros, tendo estes membros mudado as suas estratégias.

gadores. Entretanto podem existir situações em que um candidato c não necessita conhecer que Instituições obteria nos matchings ótimos para encontrar uma forma de manipular o mecanismo. Mostramos, no Lema 1, que a informação requerida é *escolher, pelo menos, uma Instituição que seja estritamente preferida àquela designada pelo mecanismo e que não seja mais preferida que aquela obtida no matching ótimo para os candidatos.* A Proposição 7 prova que *um mecanismo estável é manipulável por c se e somente se a Instituição à qual c é designado pelo mecanismo é menos preferida, por c, àquela obtida no matching estável ótimo para os candidatos.*

Para as situações em que é extremamente complicado obter as informações necessárias, poderá ser útil saber que estratégias não escolher, isto é, que classes de estratégias para os candidatos podem ser consideradas ruins, no sentido de serem dominadas por outras estratégias disponíveis. Segundo as conclusões do Teorema 6 apresentado aqui, *não listar a Escola favorita em primeiro lugar é uma estratégia estritamente dominada no jogo induzido pelo mecanismo que produz o matching estável ótimo para as instituições.*[5] Reciprocamente, o Teorema 7 mostra que *qualquer permutação das Instituições aceitáveis segundo as preferências verdadeiras, que mantém fixa a Instituição favorita, não é uma estratégia dominada.*

Desta forma, listar a escola favorita em primeiro lugar é o conselho mais prático que podemos dar a um candidato.

Roth e Rothblum (1995) estudaram o problema de como aconselhar um candidato quando ele tem muito pouca informação sobre as preferências dos outros participantes e todas as Instituições tem apenas uma vaga. Eles provaram que se um candidato está submetendo uma lista de preferências com muito poucas Instituições,

[5] Dizemos que uma estratégia E para o candidato c estritamente domina uma outra estratégia E' para o mesmo candidato, se o payoff de c quando ele joga E é maior ou igual ao seu payoff quando ele joga E', independentemente das estratégias jogadas pelos outros jogadores, e é estritamente maior que o seu payoff quando ele joga E' para, pelo menos, um conjunto de estratégias para os outros jogadores.

então ele deve revelar suas verdadeiras preferências sobre todas as Instituições às quais ele se aplicar. Caso contrário ele deve submeter uma estratégia obtida das verdadeiras preferências pela eliminação de todas as Instituições menos preferidas que uma dada Instituição.

2. Modelo Cooperativo Formal e Resultados Preliminares.

Existem dois conjuntos finitos e disjuntos de agentes: C, com m candidatos, e I, com n Instituições. A cada candidato c está associado um vetor $G_c = (G_{c1}, G_{c2}, \cdots, G_{cr})$, onde $0 \leq G_{cj} \leq L$, para todo $j = 1, 2, \cdots, r$ e r e L são inteiros positivos. A cada Instituição i está associado um vetor $p_i = (p_{i1}, p_{i2}, \cdots, p_{ir})$, onde $\sum p_{ij} = 1$, com $j = 1, 2, \cdots, r$ e $p_{ij} \geq 0$, para todo j. Podemos pensar que cada candidato se submete a um exame sobre r disciplinas. (No caso atual da ANPEC, $r = 5$). Cada G_{cj} representa o grau que o candidato c obteve no exame da disciplina j. Em geral é comum representar os graus por um número entre 0 e 10. Neste caso $L = 10$. Cada p_{ij} representa a fração do peso que a Instituição i atribui à disciplina j. A média ponderada de c segundo i é $p_i \cdot G_c = p_{i1}G_{c1} + \cdots + p_{ir}G_{cr}$. Assim sendo, i prefere o candidato c ao candidato c' se e somente se $p_i \cdot G_c > p_i \cdot G_{c'}$. Usaremos a notação $c >_i c'$ para indicar que i prefere (estritamente) c a c'. Para a preferência não estrita usaremos $\geq_i$. Se $p_i \cdot G_c = p_i \cdot G_{c'}$ o desempate é feito por algum critério estabelecido à priori (por exemplo, a ordem alfabética). Cada Instituição i possui uma cota q_i, representando o número de vagas de que dispõe. Diremos que um candidato c é aceitável para i se $p_i \cdot G_c \geq K$. Assim K representa a média mínima requerida para que um candidato seja aceitável para qualquer Instituição i. A interpretação desta condição é obvia: se não for satisfeita para algum candidato c, então i preferirá deixar uma vaga não preenchida a admitir c. Estaremos assumindo que as preferências das Instituições sobre grupos de candidatos satisfazem à seguinte propriedade: se A é um conjunto de candidatos com cardinalidade menor do que q_i e c e c' são dois candidatos não pertencentes ao conjunto A, então $A \cup \{c\} >_i A \cup \{c'\}$ se e somente se $c >_i c'$. Todo candidato c tem uma ordenação total e estrita sobre o conjunto $I \cup \{c\}$. A posição em que ele se coloca na ordenação tem o significado de que as únicas Instituições em que gostaria de ingressar são aquelas que ele prefere a si mesmo. Denotaremos por $P(c)$ a lista de c. Então $P(c) = i_1, i_2, i_3$ significa que as únicas Instituições que o candidato c aceitaria são i_1 e i_2, nesta ordem. Escreveremos $i >_c i'$ para indicar que o candidato c prefere (estritamente) i a i', e $i >_c c$ para indicar que o candidato c prefere (estritamente) i a não ser admitido por nenhuma Instituição. A preferência não estrita será denotada por $\geq_c$. A Instituição i é aceitável para o candidato c se $i >_c c$. Se $A(j)$ é um conjunto para todo agente $j \in B$, denotaremos o conjunto dos $A(j)$ por A_B. Dado o sistema de preferências P, denotamos por $P/P'(c)$ o sistema de preferências obtido de P pela substituição de $P(c)$ por $P'(c)$, matendo as outras preferências inalteradas. O mercado é assim dado por $M = ((I, p, K, q), (C, G), P_C)$, onde p é a matriz de pesos, $n \times r$; q é o vetor de cotas das Instituições; G é a matriz dos graus, $r \times m$, e P_C é o conjunto das preferências dos candidatos. Para os propósitos cooperativos o que se necessita conhecer a respeito dos participantes são somente as suas preferências. Neste caso podemos denotar o mercado, simplesmente, por $M = (C, I, P, q)$, onde $P = \{P_C, P_I\}$ e P_I é o conjunto das listas de preferências das Instituições obtidas a partir de p, G e K. Assim sendo, o nosso modelo é idêntico ao "College Admission Model" de Gale e Shapley.

A função do mercado é encontrar um matching entre os candidatos e as Instituições. Cada candidato é designado a, no máximo, uma Instituição; cada Instituição pode admitir, no máximo, sua cota de candidatos. O candidato que não for admitido por nenhuma Instituição será designado a si mesmo e será chamado "auto-designado"; uma Instituição que tiver algum número de vagas não preenchidas será designada a si mesmo em cada uma dessas vagas. Um matching é bilateral, no sentido que, se um candidato é designado a uma Instituição então o candidato pertence ao grupo ao qual a Instituição foi designada. Formalmente:

Definição 1. Um *matching* μ é uma função do conjunto $C \cup I$ no conjunto das famílias não ordenadas de elementos de $C \cup I$ tal que:

1. $\mu(c)$ tem um elemento para todo candidato c e se $\mu(c) \notin I$ então $\mu(c) = c$;
2. $\mu(i)$ tem q_i elementos para toda Instituição i, e se o número de candidatos em $\mu(i)$ for menor do que q_i, digamos k, então $\mu(i)$ contém $q_i - k$ cópias de i;
3. $\mu(c) = i$ se e somente se c estiver em $\mu(i)$.

Assim $\mu(c) = i$ denota que o candidato c foi designado à Instituição i pelo matching μ, $\mu(i) = [c_1, c_2, i, i]$ denota que a Instituição i, com cota de 4, admitiu os candidatos c_1 e c_2 e tem duas vagas não preenchidas. Representaremos matchings graficamente, por exemplo,

$$\mu: \quad \begin{array}{ccccccc} c_1 & c_2 & i_1 & \quad & c_4 & c_5 \\ \backslash & | & / & & | & | \\ & i_1 & & & i_2 & c_5 \end{array}$$

representa um matching no qual a Instituição i_1, com três vagas, admite dois candidatos c_1 e c_2; a Instituição i_2, com uma vaga, admite o candidato c_4, e o candidato c_5 é auto-designado.

Definição 2. Um matching μ é *individualmente racional* se para todo par (c, i) tal que $\mu(c) = i$, c é aceitável para i e i é aceitável para c.

Definição 3. Um matching μ é *estável* se é individualmente racional e não existem um candidato c e uma Instituição i, não associados pelo matching μ e tal que i) $i >_c \mu(c)$ e ii) $c >_i \tau$, para algum τ em $\mu(i)$.

Portanto, num matching estável, todos os pares associados pelo matching são mutuamente aceitáveis. Além disso, não existem um candidato e uma Instituição, não designados um ao outro pelo matching, tal que o candidato prefere a Instituição àquela para a qual foi designado e, paralelamente, o candidato obteve uma média ponderada que o qualifica para a Instituição em questão e é superior a de algum candidato designado à Instituição pelo matching. Um candidato c e uma Instituição i, como na definição 3, são ditos *bloquear o matching* μ.

Definição 4. O matching estável μ_C é chamado *ótimo para os candidatos* se $\mu_C(c) \geq_c \mu(c)$ para todo candidato c e para todo matching estável μ. Analogamente se define o matching estável ótimo para as Instituições, μ_I.

O Teorema 1 responde à pergunta: Para onde vão os melhores candidatos? Quem vai para as melhores Instituições?

Teorema 1. Suponha que existam um conjunto de Instituições $I' = \{i_1, \cdots, i_k\}$ e um conjunto de candidatos $C' = \{c_1, \cdots, c_t\}$, com $t = \Sigma q_j$ para $j = 1, 2, \cdots, k$, tais que I' é o conjunto das k primeiras Instituições aceitáveis da lista de qualquer candidato de C' (em qualquer ordem) e C' é o conjunto dos t primeiros candidatos aceitáveis da lista de qualquer Instituição em I' (em qualquer ordem). Então, se μ é um matching estável, $\mu(C') = I'$.

Isto é, se existe um consenso sobre o conjunto das k melhores Instituições então é esperado que elas sejam as preferidas dos melhores candidatos. O teorema diz que t vagas dessas k Instituições serão preenchidas pelos t melhores candidatos, em qualquer matching estável.

Demonstração: Seja μ um matching estável. Como o número total de vagas disponíveis no conjunto I' é igual ao número de candidatos em C', basta mostrar que $\mu(C') \subseteq I'$. Então, seja $c \in C'$. Se $\mu(c) \notin I'$ temos que alguma Instituição i_j em I' admitiu algum candidato c' fora de C' ou tem uma vaga não preenchida. Em qualquer caso, i_j prefere c a algum elemento de $\mu(i_j)$ (que pode ser i_j). Por outro lado, c prefere i_j a qualquer Instituição fora de I', em particular c

prefere i_j a $\mu(c)$, (que pode também ser c). Logo (c, i) bloqueia μ, o que é uma contradição. $\square$

Observação 1. Se no Teorema 1 tivermos $t > \Sigma q_j$ para $j = 1, 2, \cdots, k$, então $\mu(i) \subseteq C'$ para todo $i \in I'$. A demonstração deste fato, que é tão simples quanto a do Teorema 1, é deixada ao leitor.

Como foi dito na Introdução, a existência de matchings estáveis para este modelo foi demonstrada por Gale e Shapley, através de um algoritmo que produz o matching estável ótimo para os candidatos. A demonstração desse resultado e dos demais enunciados abaixo podem ser encontradas em Roth e Sotomayor (1990).

Proposição 1. (Gale e Sotomayor). Sejam μ um matching estável, $C(\mu)$ o conjunto de candidatos admitidos a alguma Instituição segundo μ e N_i o número de candidatos admitidos pela Instituição i. Então o conjunto $C(\mu)$ e os números N_i são os mesmos para todo matching estável.

Proposição 2. (Roth). Seja μ um matching estável. Se a Instituição i não preencheu todas as suas vagas em μ, então ela admitirá o mesmo conjunto de candidatos em qualquer matching estável.

Proposição 3. (Dubins-Freedman). Seja P uma estrutura de preferências para os candidatos e Instituições. Seja P' obtida de P pela substituição de $P(c)$ por $P'(c)$ para todo $c \in C' \subseteq C$. Então não existe nenhum matching estável para P', que seja preferido a μ_C por todos os membros de C'.

Isto é, nenhum conjunto C' de candidatos pode, adulterando suas preferências, obter um resultado melhor para todos os seus membros, que o que obteria se todos revelassem suas verdadeiras preferências, quando o matching estável ótimo para os candidatos é usado.

Proposição 4. (Gale e Sotomayor). Suponha que $C \subseteq C'$, que μ_C é o matching estável ótimo para os candidatos em (C, I, P, q) e que μ'_C é o matching estável ótimo para os candidatos em (C', I, P', q), onde $P' = P$ em C. Então,

$$\mu'_C(i) \geq_i \mu_C(i), \quad \text{para toda Instituição } i.$$

Observação 2. Decorre da demonstração da Proposição 4 que não somente $\mu'_C(i) \geq_i \mu_C(i)$, mas também que i prefere fracamente todo candidato em $\mu'_C(i)$ a algum candidato em $\mu_C(i)$.

Proposição 5. (Roth e Sotomayor). Sejam μ e μ' matchings estáveis. Se $\mu(i) >_i \mu'(i)$, para alguma Instituição i, então $c >_i c'$ para todo c em $\mu(i)$ e c' em $\mu'(i) - \mu(i)$.

Ou seja, se i prefere o grupo de candidatos designados a ela por μ ao grupo de candidatos designados a ela por μ', então *ela prefere qualquer candidato do primeiro grupo a qualquer candidato do segundo grupo que não estiver no primeiro.*

Proposição 6. (Gale e Shapley). Seja μ um matching estável. Então $\mu(c) \geq_c \mu_I(c)$, para todo c em C, e $\mu(i) \geq_i \mu_C(i)$, para todo i em I.

3. O Modelo Não-Cooperativo Formal e um Mecanismo Coletivamente à Prova de Estratégia.

Nesta seção consideramos uma Central que, para dois dados conjuntos C e I, recebe as listas de preferências $Q(c)$, para todo c em C e que, por algum procedimento, determina as listas de preferências $P(i)$ para todo i em I. Em seguida usa as preferências ordenadas dos agentes e suas cotas como "input" para algum algoritmo que produz como "output" final um matching estável para $(C, I, \{Q_C, P_I\}, q)$. A função H que a cada possível mercado $(C, I, \{Q_C, P_I\}, q)$ associa um matching $\mu = H(C, I, \{Q_C, P_I\}, q)$, entre C e I é denominada um *mecanismo de matching*. Se o matching produzido é sempre estável com respeito a Q_C e P_I o mecanismo de matching é dito estável; se é sempre o matching estável ótimo para um dos lados, o mecanismo é chamado estável ótimo para este lado.

Fixemos agora os conjuntos C e I. Então a Central fixa P_I e q. Neste caso as únicas variáveis são as preferências Q_C. Denotaremos por $D(c)$ o conjunto de todas as listas ordenadas de preferências de c, para todo c em C. $P(c)$ denotará a lista ordenada das preferências verdadeiras de c em C. Quando cada candidato c em C é chamado a decidir que lista de preferências $Q(c) \in D(c)$ informar à Central, configura-se uma situação de jogo na forma estratégica induzido pelo mecanismo H. Os jogadores são o conjunto de candidatos; o conjunto de estratégias disponíveis para cada candidato c é dado por $D(c)$. O algoritmo usado pela Central toma P_I e q como dados e para cada conjunto Q_C de preferências informadas determina um matching estável entre C e I. Isto é, se h é a função resultado que descreve como o algoritmo usado pela Central transforma o conjunto das preferências Q_C no matching μ, então $\mu = h(Q_C) = H(C, I, \{Q_C, P_I\}, q)$. Denotaremos por h_c a coordenada de h referente a c.

Definição 5. Dizemos que um mecanismo é *à prova de estratégia* se, em todo jogo induzido pelo mecanismo, nenhum jogador puder se beneficiar por usar como estratégia um conjunto de preferências diferente das suas verdadeiras preferências. Um mecanismo é *coletivamente à prova de estratégia* se em todo jogo induzido pelo mecanismo nenhum conjunto de jogadores puder, adulterando suas verdadeiras preferências, obter um resultado mais vantajoso para cada um de seus membros.

Portanto se um mecanismo estável é coletivamente à prova de estratégia, em qualquer jogo induzido por este mecanismo os participantes têm o incentivo de revelar suas verdadeiras preferências, o que garante a estabilidade dos resultados em relação a estas preferências.

Podemos agora enunciar o resultado principal desta seção:

Teorema 2. Um mecanismo de matching estável H é coletivamente à prova de estratégia se e somente se seleciona o matching estável ótimo para os candidatos, para cada mercado.

Para a demonstração deste resultado necessitamos do Lema 1 abaixo. Segundo este resultado, truncar as verdadeiras preferências em qualquer ponto entre $\mu_C(c)$ e $\mu_I(c)$ produz um resultado melhor do que $\mu_I(c)$. De fato,

Lema 1. Seja $c \in C$ e $i \in I$ tal que $\mu_C(c) \geq_c i \geq_c \mu(c) >_c c$, onde μ é algum matching estável para (C, I, P, q). Seja $P'(c)$ obtida da verdadeira preferência, $P(c)$, pela remoção de todas as Instituições menos preferidas que i. Então, $\mu'(c) \geq_c i$, para todo matching μ', estável para $(C, I, P/P'(c), q)$. Consequentemente, se $\mu_C(c) \geq_c i >_c \mu(c)$, temos que $\mu'(c) >_c \mu(c)$, para todo matching μ' estável para $(C, I, P/P'(c), q)$.

Demonstração: Basta observar que μ_C é estável em relação à $P' = P/P'(c)$. Assim, como c não é auto-designado em μ_C, não será auto designado em μ', para todo matching μ' estável para $(C, I, P/P'(c), q)$. Portanto, para todo tal matching μ', $\mu'(c) \geq_c i$, pela construção de $P'(c)$.

A outra afirmação é, agora, imediata. □

Demonstração do Teorema 2. Suponha que exista um mercado para o qual H seleciona um matching estável μ, diferente de μ_C. Vamos mostrar que H é manipulável por qualquer candidato que obtenha resultados diferentes em μ e μ_C. Então seja c um tal candidato, isto é, $\mu_C(c) >_c \mu(c) \geq_c c$. O Lema 1 aplicado a $i = \mu_C(c)$ implica que existe uma estratégia para c, $P'(c)$, tal que se

$\mu' = H(C, I, P/P'(c), q)$ então c prefere $\mu'(c)$ a $\mu(c)$. Por conseguinte H não é à prova de estratégia.

Reciprocamente, se H produz o matching estável ótimo para os candidatos, a Proposição 3 garante que nenhum candidato ou grupo de candidatos pode manipular o mecanismo. □

O nosso próximo resultado dá uma condição necessária e suficiente para que um candidato possa manipular um dado mecanismo estável.

Proposição 7. Seja H um mecanismo estável tal que $H(C, I, P, q) = \mu$. Seja $c \in C$. Então H é manipulável pelo candidato c se e somente se $\mu_C(c) \neq \mu(c)$.

Demonstração: Se H é manipulável por c então $\mu_C(c) \neq \mu(c)$ pela Proposição 3. Para a recíproca use o Lema 1 para $i = \mu_C(c)$ e conclua que truncando as verdadeiras preferências em $\mu_C(c)$, o candidato c obtém um resultado mais preferível do que $\mu(c)$.

4. Possibilidades Estratégicas dos Candidatos.

4.1. Comportamento de Equilíbrio.

Nesta seção estaremos considerando um *mecanismo de matching estável* H (não necessariamente o ótimo para os candidatos ou para as Instituições), adotado por uma Central, para ser usado por um conjunto de candidatos C, um conjunto de Instituições I, com seu vetor de cotas q e suas verdadeiras preferências P_I determinadas pela Central. Seja $J = (C, D_C, h, P_C)$ o jogo estratégico induzido por H.

No caso em que as Instituições têm cota 1, revelar as verdadeiras preferências é uma estratégia dominante para as Instituições quando o mecanismo sempre seleciona o matching estável ótimo para as Instituições. Nesta situação, os Teoremas 3, 4, 5, 6 e 7 desta seção são análogos aos obtidos em Gale e Sotomayor (1985 b).

As demonstrações dos Teoremas 3, 5 e 7, apresentadas aqui, seguem as linhas das demonstrações de Gale e Sotomayor, com as devidas adaptações. Entretanto, as demonstrações dos Teoremas 4 e 6, embora utilizem os argumentos centrais usados em Gale e Sotomayor, foram obtidas graças à Proposição 5, que não tem paralelo no modelo de matching um-a-um.

Cumpre observar que as conclusões dos Teoremas 6 e 7 se aplicam também às situações em que as Instituições têm comportamento estratégico.

O Teorema 2 implica que se o matching produzido pelo mecanismo H não for o matching estável ótimo para os candidatos, então a honestidade não será a melhor política para eles. Neste caso será natural perguntar se haverá algum conjunto de estratégias com a propriedade de que uma vez que sejam adotadas pelos candidatos, não seja vantajoso para nenhum deles mudar a sua estratégia. Isto é, existe algum conselho que poderemos dar aos candidatos que, uma vez que seja adotado, seja ainda um bom conselho segui-lo? Um conjunto de estratégias, uma para cada jogador, que tem esta propriedade é chamada um equilíbrio (ou equilíbrio de Nash). Formalmente,

Definição 6. Um conjunto de estratégias $P^* = \{P^*(c_1), ..., P^*(c_m),\}$ é um equilíbrio para o jogo $J = (C \cup I, \{D_C\}, h, P_C)$ se para cada jogador c em C, $h_c(P^*_C) \geq_c h_c(P^*_C/P'(c))$, para toda estratégia $P'(c) \in D_C$.

Isto é, um conjunto de estratégias, uma para cada jogador, forma um equilíbrio se nenhum jogador pode conseguir um payoff melhor mudando a sua estratégia, assumindo que os outros jogadores não mudam as suas.

A existência de estratégias de equilíbrio é demonstrada no seguinte teorema:

Teorema 3. Seja μ um matching estável para (C, I, P, q) tal que $\mu(c) \geq_c h_c(P_C)$ para todo c em C. Suponha que cada candidato c, tal que $\mu(c) \in I$, escolha a estratégia $P'(c)$ que consiste um listar somente $\mu(c)$ como a única Instituição aceitável; no caso em que $\mu(c) = c$, $P'(c) = c$. O conjunto de estratégias $P'_C = \{P'(c_1), \cdots, P'(c_m)\}$ é um equilíbrio no jogo induzido pelo mecanismo H (e μ é o matching que resulta).

Demonstração: É claro que μ é estável para o mercado $M' = (C, I, \{P'_C, P_I\}, q)$. Além disso, μ é o único matching estável para M', pois qualquer outro matching estável deixaria algum candidato c, anteriormente designado a alguma Instituição por μ, auto designado por esse outro matching. Mas isto, pela Proposição 1, não é possível. Portanto μ é o matching estável produzido pelo mecanismo h quando os jogadores jogam P'_C. Resta agora ver que $P'_C = \{P'(c_1), \cdots, P'(c_m)\}$ é um equilíbrio. Suponha que algum candidato c mude sua lista de preferências de $P'(c)$ para $P''(c)$ produzindo um novo conjunto de estratégias $P''_C = P'_C/P''(c)$, com um novo matching estável $h(P''_C) = \mu'$, que dá a ele uma Instituição $i' = \mu'(c)$, que ele prefere a $\mu(c)$, segundo suas preferências verdadeiras. Entretanto a Instituição i' deve ter preenchido sua cota no matching μ e prefere qualquer candidato designado a ela por μ, ao candidato c, pois caso contrário o par (c, i') causaria uma instabilidade no matching μ em (C, I, P, q). Mas então algum candidato c' que fora designado a i' por μ não está designado a i' por μ'. Como a única Instituição aceitável da lista de c' em P''_C é i', isto significa que c' é auto-designado por μ'. Por conseguinte o par (c', i') bloqueia o matching μ' em $(C, I, \{P''_C, P_I\}, q)$, contradição. □

O Teorema 3 diz que os candidatos podem forçar, por estratégias de equilíbrio, qualquer matching μ que seja estável em relação às verdadeiras preferências e seja fracamente preferido ao resultado do jogo. Em particular podem obter o matching estável ótimo para eles. O seguinte resultado é uma espécie de recíproca:

Teorema 4. Suponha que os candidatos escolham qualquer conjunto de estratégias (listas de preferências), P'_C, que formem um equilíbrio para o jogo induzido por um mecanismo estável H. Então, $h(P'_C) = H(C, I, \{P'_C, P_I\}, q)$ é um matching estável para (C, I, P, q).

Demonstração: Seja $\mu' = h(P'_C)$. Suponha, por absurdo, que (c, i) causa uma instabilidade em μ' em relação às verdadeiras preferências P. Então a Instituição i prefere c a algum candidato c', designado a ela por μ'. Mostraremos que P' não é um equilíbrio. De fato, suponha que em vez de escolher $P'(c)$, c lista somente i. Então ele obterá i, pois seja $\mu'' = h(P''_C)$, onde P''_C difere de P'_C apenas na nova lista de c. Se c não obtiver i, ficará auto-designado em μ''. Pela Proposição 1, c também ficará auto-designado em μ''_C, o matching estável ótimo para os candidatos segundo $\{P''_C, P_I\}$. Seja μ'_C o matching estável ótimo para os candidatos segundo P'. Sabemos que $\mu'_C(i) \leq_i \mu'(i)$, pela Proposição 6. Temos, também, que em ambos os matchings i preenche a mesma fração de sua cota, pela Proposição 1. Então, se $\mu'_C(i) <_i \mu'(i)$, decorre que $\mu'_C(i) - \mu'(i) \not\in \emptyset$. Neste caso, a Proposição 5 implica que i prefere cada candidato em $\mu'(i)$ a todo candidato em $\mu'_C(i) - \mu'(i)$, em particular, i prefere c' a todo candidato em $\mu'_C(i) - \mu'(i)$. Pela transitividade das preferências, i preferirá c a qualquer candidato em $\mu'_C(i) - \mu'(i)$. Em qualquer caso, portanto, existirá $c^* \in \mu'_C(i)$ tal que $c >_i c^*$. Por outro lado, a estabilidade de μ''_C e o fato de que $\mu''_C(c) = c$ implicam que a Instituição i preencherá todas as suas vagas em μ''_C e preferirá qualquer candidato a ela designado por μ''_C ao candidato c, e por transitividade, a Instituição i preferirá qualquer candidato designado a ela por μ''_C a c^*. Mas claramente o matching μ''_C seria estável para $(C - \{c\}, I, P', q)$, onde P' está restrita a $C - \{c\}$. Decorre da Proposição 4 que $\mu'_C(i) \geq_i \mu''_C(i)$ e da Observação 1 que i prefere fracamente qualquer candidato em $\mu'_C(i)$ a algum candidato em $\mu''_C(i)$. Em particular, i prefere fracamente c^* a algum candidato em $\mu''_C(i)$, contradição. Logo $\mu''(c) = i$, o que contraria a hipótese de que P' é um equilíbrio. $\square$

Um outro uso da forma estratégica de um jogo é olhar o comportamento das coalizões. A pergunta agora é: Será que existem sempre equilíbrios fortes para os candidatos no jogo J? O seguinte teorema mostra que sim.

Teorema 5. Seja J o jogo induzido pelo mecanismo estável H. Suponha que cada candidato c escolha a estratégia $P'(c)$ obtida das verdadeiras preferências de c pela remoção de todos as Instituições menos preferidas que $\mu_C(c)$. O conjunto de estratégias $P' = \{P'(c_1), \cdots, P'(c_m)\}$ é um equilíbrio forte para os candidatos no jogo J (e μ_C é o matching que resulta).

Demonstração: Afirmamos que μ_C é o único matching estvel para $(C, I, \{P'_C, P_I\}, q)$, pois claramente μ_C é estável e qualquer outro matching estável μ' deve ter $\mu'(c) \neq \mu_C(c)$ para algum c; portanto, c não é auto-designado por μ'. Então $\mu'(c)$ é preferido por c a $\mu_C(c)$ por construção de P'_C. Desde que μ_C é o matching ótimo para os candidatos segundo P, isto significa que μ' é instável com relação às verdadeiras preferências e portanto é bloqueado por algum par (c', i). Mas por construção de P'_C isto implica que (c', i) bloqueia μ' com relação às preferências P'_C, contradizendo a P'_C-estabilidade de μ'. Agora, desde que μ_C é o único matching estável para $(C, I, \{P'_C, P_I\}, q)$ ele é o matching estável ótimo para os candidatos neste mercado. Se algum subconjunto $C^* \subseteq C$ pudesse obter um payoff melhor para todos os seus membros declarando preferências diferentes, então todos os candidatos em C^* obteriam um payoff melhor do que o dado por μ_C; mas pelo Teorema de Dubins e Freedman isto é impossível. $\square$

Observação 3. Parece razoável considerar a estratégia P' do Teorema 5 como o melhor método de jogar para os candidatos. É (a) um equilíbrio forte, e portanto nenhum candidato ou conjunto de candidatos seria tentado a se desviar de P', e (b) dentre todas as estratgias de equilíbrio esta dá aos candidatos o mais alto payoff possível.

Entretanto, fora os problemas de coordenação entre os jogadores, que podem se originar na escolha de um conjunto particular de estratégias de equilíbrio, uma grande quantidade de informação pode ser requerida para os jogadores conseguirem computar uma tal estratégia. Para computar as estratégias descritas no Teorema 5, por exemplo, cada candidato c necessita conhecer $\mu_C(c)$, que pode ser computado somente com uma grande quantidade de informação sobre as preferências dos outros candidatos e Instituições. Estas informações seriam difíceis de serem obtidas em situações como a que estamos modelando aqui, com um grande número de participantes. Mesmo com toda essa dificuldade, o problema seria enormemente simplificado se pudessemos afirmar que o matching μ_C é o único matching estável obtido de um equilíbrio forte para os candidatos. Entretanto este não é o caso, como mostra o exemplo em Gale e Sotomayor (1985 b), que ilustra um caso particular em que cada Instituição pode aceitar, no máximo, um candidato.

4.2. Estratégias Dominadas.

Quando um mecanismo seleciona o matching ótimo para os candidatos, um candidato pode confiantemente declarar suas verdadeiras preferências, sem considerar quais são as preferências dos outros participantes, posto que ele está usando uma estratégia dominante. Entretanto, como foi visto aqui, declarar as verdadeiras preferências não é uma boa estratégia para os candidatos quando o mecanismo seleciona o matching estável ótimo para as Instituições. Nesta situação, aconselhar um candidato c a escolher uma estratégia como no Teorema 5, será inútil se ele desconhecer $\mu_C(c)$. Entretanto existem situações em que um candidato c não necessita conhecer $\mu_C(c)$ nem $\mu_I(c)$ para encontrar uma forma de manipular um tal mecanismo. O Lema 1 implica que o que ele precisa saber é se existe uma Instituição i tal que $\mu_C(c) \geq_c i >_c \mu_I(c)$. A obtenção de uma tal informação pode ser simples, como ilustramos a seguir.

Suponha que o candidato c é o $q_{i'}+1$-ésimo elemento da lista de candidatos aceitáveis de i' e está entre os q_i primeiros da lista de aceitáveis de i. Suponha, também, que existe um outro candidato c' que está listado entre os $q_{i'}$ primeiros elementos aceitáveis de i' e é o q_i+1-ésimo elemento da lista de candidatos aceitáveis de i. Esta informação é de conhecimento de c, visto que as listas de preferências das Instituições sobre os candidatos é de conhecimento geral.

Suponha, agora, que c prefere i' a i e tem informação (que pode ter sido fornecida pelo próprio c') de que c' prefere i a i'. Qual deve ser o efeito dessa informação no comportamento estratégico de c?

Primeiramente observe que c obterá, no mínimo, i em μ_I e, no mínimo, i' em μ_C. Como c não tem informações a respeito dos outros candidatos, não tem certeza a respeito dos matchings μ_I e μ_C. Mesmo nestas condições, o Lema 1 garante ao candidato c, no mínimo, i', se este truncar suas verdadeiras preferências em i'.

No exemplo acima, a informação que c obteve a respeito das preferências de apenas um dos candidatos permitiu-lhe concluir que $\mu_C(c) \geq_c i' \geq_c \mu_I(c)$.

Do que aprendemos até agora podemos deduzir que o nível de informação requerida para que um candidato c trunque as suas preferências em alguma Instituição i, sem risco de ficar auto-designado é aquele que lhe permite deduzir que $\mu_C(c) \geq_c i \geq_c \mu_I(c)$.

Para alguns candidatos, saber se $\mu_C(c)$ é diferente de $\mu_I(c)$, é bastante simples, como é o caso do candidato c no exemplo dado; para outros, porém, é extremamente complicado. Nesses casos qualquer truncamento que exclua $\mu_I(c)$ poderá ser desastroso, deixando o candidato auto-designado.

Na falta da informação necessária para que um candidato possa deduzir se $\mu_C(c) \neq \mu_I(c)$, poderia ser útil saber que estratégias não jogar, isto é, que classes de estratégias para os candidatos podem ser consideradas ruins, no sentido de serem dominadas por outras estratégias disponíveis.

Definição 7. Dizemos que uma estratégia E para o candidato c *domina* uma outra estratégia E' para o mesmo jogador se o payoff de c quando ele joga E é maior ou igual ao seu payoff quando ele joga E', independentemente das estratégias jogadas pelos outros jogadores. Dizemos que uma estratégia E para o candidato c *estritamente domina* uma outra estratégia E' para o mesmo candidato, se E domina E' e o payoff de c quando ele joga E é estritamente maior que o seu payoff quando ele joga E' para, pelo menos, um conjunto de estratégias para os outros jogadores.

Se existir somente um candidato c, então é claro que revelar suas verdadeiras preferências estritamente domina qualquer outra estratégia que possua, pois se ele submete suas verdadeiras preferências, obterá a Instituição mais preferida dentre aquelas que o aceitarem. Se ele adultera suas preferências para, por exemplo, $i' > i$ em vez de $i > i'$, ele estará pior no caso onde i e i' são as únicas Instituições que o aceitarem.

Daqui para frente assumiremos que existem, no mínimo, dois candidatos em C. Mostraremos que a estratégia de listar somente uma Instituição, como no Teorema 3, é dominada a menos que esta Instituição seja a primeira da lista verdadeira do candidato.

Teorema 6. Seja J o jogo induzido pelo mecanismo H que produz o matching estável ótimo para as Instituições. Seja c um candidato e i sua verdadeira primeira escolha. Suponhã que c seja aceitável para i. Então, qualquer estratégia $P'(c)$ em que c não lista i em primeiro lugar é estritamente dominada no jogo J.

Demonstração: Seja $P'(c)$ uma estratégia. Mostraremos que $P'(c)$ é estritamente dominada por $P''(c)$ que lista i em primeiro lugar e deixa inalterado o resto de $P'(c)$.

Sejam $\mu'_I = h(P'_C)$ e $\mu''_I = h(P''_C)$, onde P'_C e P''_C diferem apenas nas estratégias de c. Suponha primeiramente que $\mu''_I(c) \neq i$. Então, como c é aceitável para i, a Instituição i deve ter preenchido todas as suas vagas em μ''_I e prefere qualquer candidato designado a ela por μ''_I a c, pois senão (c, i) bloquearia μ''_I. Portanto μ''_I é aceitável em relação às preferências $P'(c)$. Assim, $\mu'_I(i) \geq_i \mu''_I(i)$, da otimalidade de μ'_I. Pela transitividade das preferências e pela Proposição 5 temos que i prefere qualquer candidato em $\mu'_I(c)$ a c. Então, desde que todos os outros elementos diferentes de i são listados na mesma ordem em $P'(c)$ e $P''(c)$, segue que μ'_I é estável segundo P''_C e portanto $\mu'_I(c) = \mu''_I(c)$. Assim c não está pior usando $P''(c)$ em vez de $P'(c)$. (Se $\mu''_I(c) = i$ a conclusão é imediata). Para mostrar que $P''(c)$ estritamente domina $P'(c)$, seja i' o primeiro elemento em $P'(c)$. Então, podemos ver que $\mu''_I(c) = i >_c \mu'_I(c) = i'$, no caso onde i (resp. i') tem c entre os seus q_i (resp. $q_{i'}$) mais preferidos candidatos, o que conclui a demonstração. $\square$

O Teorema 7 afirma que o Teorema 6 descreve todas as estratégias dominadas para um dado candidato.

Teorema 7. Seja J o jogo induzido pelo mecanismo. Seja $P'(c)$ qualquer estratégia em que (a) c lista a sua primeira verdadeira escolha em primeiro lugar, e (b) as Instituições aceitáveis em $P'(c)$ são também aceitáveis na lista das verdadeiras preferências de c, $P(c)$. Então $P'(c)$ não é uma estratégia dominada no jogo J.

Demonstração: Mostraremos que para qualquer outra estratégia $P''(c)$ existem estratégias P_{-c} para os outros jogadores tais que $\mu'_I(c) >_c \mu''_I(c)$, onde μ'_I e μ''_I são os matchings estáveis ótimos para as Instituições para $(C, I, (P'(c), P_{-c}), q)$ e $(C, I, (P''(c), P_{-c}), q)$, respectivamente. Seja i' a verdadeira primeira escolha de c. Existem três casos. Primeiro suporemos que $P''(c)$ também satisfaz a condição (a) acima.

1º caso: i é uma Instituição aceitável em $P'(c)$ mas não em $P''(c)$. Então, para P_{-c} supomos que i lista c entre os seus q_i primeiros candidatos e nenhuma outra Instituição lista c. Então, $\mu'_I(c) = i$ enquanto c é auto-designado em μ''_I. Portanto c prefere μ'_I a μ''_I (aqui é usado o fato que i é aceitável segundo as verdadeiras preferências $P(c)$).

2º caso: i é uma Instituição aceitável em $P''(c)$ mas não em $P'(c)$. Então, para P_{-c} supomos que i lista c entre os seus q_i primeiros candidatos e lista c' como o $q_i + 1$-ésimo, para algum c'; i' lista c' entre os seus $q_{i'}$ primeiros candidatos e lista c como o $q_{i'} + 1$-ésimo; c' prefere i a i' e nenhuma outra Instituição lista c ou c'. Então, $\mu'_I(c) = i' >_c \mu''_I(c) = i$.

3º caso: $P'(c)$ e $P''(c)$ contém o mesmo conjunto de Instituições aceitáveis, mas c prefere i a i^* em $P'(c)$ e i^* a i em $P''(c)$. Então suponha que i' lista c' entre os seus $q_{i'}$ primeiros candidatos e lista c como o $q_{i'} + 1$-ésimo; i lista c entre os seus q_i primeiros candidatos e lista c' como o $q_i + 1$-ésimo candidato; i^* lista c entre os seus q_{i^*} primeiros candidatos e lista c' como o $q_{i^*} + 1$-ésimo; e c' prefere i^* a i' e i' a i e nenhuma outra Instituição lista c ou c'. Agora verifique que $\mu'_I(c) = i'$, enquanto $\mu''_I(c) = i'$.

Vimos que se $P'(c)$ satisfaz (a) e (b) e $P''(c)$ satisfaz (a) então, para algum P_{-c} temos que $\mu'_I(c) >_c \mu''_I(c)$. Se $P''(c)$ não satisfaz (a), então o Teorema 7 garante que existe $P'''(c)$ que domina $P''(c)$ e $P'''(c)$ satisfaz (a). Assim construímos P_{-c} tal que $\mu'_I(c) >_c \mu'''_I(c)$, mas $\mu'''_I(c) \geq_c \mu''_I(c)$, por dominância. Desta forma a demonstração está completa. □

Para finalizar este trabalho apresentamos, a título de ilustração, o exemplo abaixo.

Exemplo. Considere o mercado de admissão de economistas às Escolas de pós-graduação em Economia no Brasil. Para simplificar suporemos que cada Instituição possui apenas uma vaga. Os candidatos se submetem a exames em matemática, estatística, microeconomia, macroeconomia e economia brasileira. Suponhamos que os candidatos c_1, $c_2, \cdots, c_7$ têm as seguintes listas de preferências, onde $\{i_1, \cdots, i_4\}$ é o conjunto de Instituições aceitáveis para eles.

$$\begin{array}{llllll}
P(c_1): & i_2 & i_1 & i_3 & i_4 & c_1 \\
P(c_2): & i_2 & i_1 & i_3 & i_4 & c_2 \\
P(c_3): & i_2 & i_3 & i_1 & i_4 & c_3 \\
P(c_4): & i_1 & i_3 & i_2 & i_4 & c_4 \\
P(c_5): & i_1 & i_3 & i_2 & i_4 & c_5 \\
P(c_6): & i_3 & i_2 & i_1 & i_4 & c_6 \\
P(c_7): & i_2 & i_3 & i_1 & i_4 & c_7
\end{array}$$

Os graus obtidos pelos candidatos nas diversas disciplinas podem ser vistos na tabela 1; os pesos atribuídos a cada disciplina pelas Instituições $i_1, \cdots, i_4$, na tabela 2; as médias ponderadas, na tabela 3.

Tabela 1

	Matemática	Estatística	Microecon.	Macroecon.	Econ. Bras.
c_1	9	8	9	7	7
c_2	10	10	9	8	8
c_3	7	7	9	8	10
c_4	6	7	6	6	6
c_5	8	7	6	7	8
c_6	8	9	8	6	6
c_7	5	7	6	8	9

Tabela 2

	Matemática	Estatística	Microecon.	Macroecon.	Econ. Bras.
i_1	2	2	2	2	2
i_2	3	2	3	1	1
i_3	3	2	2	2	1
i_4	1	2	2	2	3

Tabela 3

	c_1	c_2	c_3	c_4	c_5	c_6	c_7
i_1	8.0	9.0	8.2	6.2	7.2	7.4	7.0
i_2	8.4	9.3	8.0	6.2	7.1	7.8	6.4
i_3	8.2	9.2	7.9	6.2	7.2	7.6	6.6
i_4	7.8	8.8	8.5	6.2	7.2	7.2	7.4

Suponhamos que qualquer outro candidato tenha média ponderada inferior a 6.2, segundo os pesos de qualquer das Instituições no conjunto $\{i_1, \cdots, i_4\}$. Então, as posições de c_1, $c_2, \cdots, c_7$ nas listas de $\{i_1, \cdots, i_4\}$ são dadas nas seguintes listas de preferências:

$$
\begin{array}{llllllll}
P(i_1): & c_2 & c_3 & c_1 & c_6 & c_5 & c_7 & c_4 \\
P(i_2): & c_2 & c_1 & c_3 & c_6 & c_5 & c_7 & c_4 \\
P(i_3): & c_2 & c_1 & c_3 & c_6 & c_5 & c_7 & c_4 \\
P(i_4): & c_2 & c_3 & c_1 & c_7 & c_5 & c_6 & c_4
\end{array}
$$

Até 1992, o procedimento só considerava as primeiras opções dos candidatos:

$$
\begin{array}{lllll}
P'(i_1): & c_5 & c_4 & & \\
P'(i_2): & c_2 & c_1 & c_3 & c_7 \\
P'(i_3): & c_6 & & & \\
P'(i_4): & \ldots & & &
\end{array}
$$

A distribuição segundo essas preferências seria então:

$$
\mu : \begin{array}{ccccccc}
i_1 & i_2 & i_3 & i_4 & c_1 & c_4 & c_7 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
c_5 & c_2 & c_6 & (c_3) & c_1 & c_4 & c_7
\end{array}
$$

onde c_3 conseguiu i_4 num arranjo subsequente.

Nesta distribuição, c_1, o segundo da lista de i_2 e i_3 e o terceiro da lista de i_1 e i_4, não obteve nenhuma escola. Se c_1 tivesse sido designado para i_1 ou i_3 ambos, escola e candidato, estariam mais satisfeitos. O mesmo pode ser dito para c_3. Pode ser visto também que se c_1 tivesse listado i_1 como primeira opção, em vez de i_2, ele teria obtido i_1.

No procedimento atual podemos esperar que todas as Instituições fazem uma oferta a c_2, quase que simultaneamente, e este, ao receber a oferta de i_2 a aceita e recusa as demais. É igualmente esperado, que nestas circunstâncias, tendo sua primeira oferta recusada, i_3 faça uma oferta à c_1 e i_1 e i_4 façam uma oferta a c_3. Imaginemos agora o seguinte cenário: c_1 sabe que está na lista de espera de i_1 e não foi informado de que i_2 já preencheu sua vaga. Pressionado por i_3 a tomar uma decisão ele recusa a oferta de i_3 no último momento, confiante num futuro melhor. Paralelamente, c_3, mais cauteloso, aceita a oferta de i_1. Tendo sua oferta recusada por c_1, i_3 faz uma oferta a c_3, mas c_3 já está comprometido ... Daí para frente, qualquer que seja o resultado, c_3 e i_3 ficarão insatisfeitos.

No procedimento usado pelo NRMP o matching obtido seria o ótimo para as Instituições:

$$\mu_I : \begin{array}{cccccccc} i_1 & i_2 & i_3 & i_4 & c_4 & c_5 & c_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ c_3 & c_2 & c_1 & c_7 & c_4 & c_5 & c_6 & \cdots \end{array}$$

caso os candidatos tivessem declarado suas verdadeiras preferências.

Se c_3, por exemplo, tivesse alterado apenas a posição de i_4 e i_1 em sua lista e tivesse submetido a lista de preferências resultante: $P'(c_3) : i_2 \quad i_3 \quad i_4 \quad i_1 \quad c_3$, o matching do NRMP teria sido:

$$\mu' : \begin{array}{cccccccc} i_1 & i_2 & i_3 & i_4 & c_4 & c_5 & c_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ c_1 & c_2 & c_3 & c_7 & c_4 & c_5 & c_6 & \cdots \end{array}$$

Neste caso c_3 teria se beneficiado pois teria obtido sua segunda escolha em vez de sua terceira. O matching estável ótimo para os candidatos, segundo as preferências verdadeiras, P, é μ' que poderia ser obtido no NRMP, também, se os candidatos tivessem usado a estratégia de equilíbrio P^* em vez de P, onde

$$
\begin{array}{llllll}
P^*(c_1): & i_2 & i_1 & c_1 & i_3 & i_4 \\
P^*(c_2): & i_2 & c_2 & i_1 & i_3 & i_4 \\
P^*(c_3): & i_2 & i_3 & c_3 & i_1 & i_4 \\
P^*(c_4) = & P(c_4) \\
P^*(c_5) = & P(c_5) \\
P^*(c_6) = & P(c_6) \\
P^*(c_7) = & P(c_7)
\end{array}
$$

Observe neste exemplo as conclusões do Teorema 1: i_1, i_2 e i_3, as três Instituições mais preferidas pelos candidatos, recebem os três melhores candidatos: c_1, c_2 e c_3 em ambos os matchings ótimos.

As conclusões da Proposição 1 podem ser observadas aqui: os candidatos c_4, c_5 e c_6 não foram admitidos por nenhuma escola nos dois matchings ótimos.

Submetido em Setembro de 1995. Revisado em Fevereiro 1996.

Referências

Alcade, J. e Salvador Barberà. 1994. "Top dominance and the possibility of strategy-proof stable solutions to matching markets." *Journal of Economic Theory* (a parecer).

Demange, G. e David Gale. 1985. "The strategy structure of two-sided matching markets." *Econometrica* **53**: 873–888.

Demange, G.; D. Gale e Marilda Sotomayor. 1987. "A further note on the stable matching problem." *Discrete Applied Mathematics* **16**: 217–222.

Dubins, L.E. e D.A. Freedman. 1981. "Machiavelli and the Gale-Shapley algorithm." *American Math. Monthly* **88**: 485–494.

Gale, D. e Lloyd Shapley. 1962. "College admissions and the stability of marriage". *American Math. Monthly* **69**: 9–15.

Gale, D. e Marilda Sotomayor. 1985 a. "Some remarks on the stable matching problem". *Discrete Applied Mathematics* **11**: 223–232.

Gale, D. e Marilda Sotomayor. 1985 b. "Ms. Machiavelli and the stable matching problem". *American Math. Monthly* **92**: 261–268.

Roth, A. 1982. "The economics of matching: stability and incentives". *Math. of Operations Research* **7**: 617–628.

Roth, A. 1984. "Misrepresentation and stability in the marriage problem." *Journal of Economic Theory* **34**: 383–387.

Roth, A. 1985. "The college admissions problem is not equivalent to the marriage problem." *Journal of Economic Theory* **36**: 277–288.

Roth, A. 1986. "On the allocation of residents to rural hospitals: a general property of two-sided matching markets." *Econometrica* **54**: 425–427.

Roth, A. e Marilda Sotomayor. 1989. "The college admissions problem revisited". *Econometrica* **57** 85–101.

Roth, A. e Marilda Sotomayor. 1990. "Two-sided matching. A study in game-theoretic modeling and analysis." *Econometric Society Monographs*, Nº 18, publicado pela Cambridge University Press.

Roth, A. e Uriel Rothblum. 1995. Strategic behavior in low-information matching markets: in search of practical advice for participants (mimeog).