

University of Pennsylvania Carey Law School

Penn Law: Legal Scholarship Repository

Faculty Scholarship at Penn Law

10-10-2005

The Residency Match: Competitive Restraints in an Imperfect World

Kristin Madison

University of Pennsylvania Carey Law School

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship



Part of the [Antitrust and Trade Regulation Commons](#), [Health Law and Policy Commons](#), and the [Medical Education Commons](#)

Repository Citation

Madison, Kristin, "The Residency Match: Competitive Restraints in an Imperfect World" (2005). *Faculty Scholarship at Penn Law*. 61.

https://scholarship.law.upenn.edu/faculty_scholarship/61

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.

ARTICLE

THE RESIDENCY MATCH: COMPETITIVE RESTRAINTS IN AN IMPERFECT WORLD

*Kristin Madison**

ABSTRACT

In 2002 physicians filed a lawsuit alleging that “the match,” the more than fifty-year-old system by which medical students and other applicants are assigned to medical residency programs, violates section 1 of the Sherman Act. Last year, without hearings on the issue, Congress found that the match was “highly efficient” and “pro-competitive” and granted a retroactive antitrust exemption for its operation. These seemingly incompatible views invite further analysis of the merits of the residency match from the perspective of public policy. This Article considers the arguments of match advocates and critics, evaluating both theoretical models and empirical evidence of the effects of the match on resident compensation. It rejects the assertion that matching mechanisms are necessarily inefficient, and instead describes factors that should be considered in an assessment of efficiency. The Article concludes that given the role of the residency match in remedying market imperfections, the

* Assistant Professor, University of Pennsylvania Law School and Senior Fellow, Leonard Davis Institute of Health Economics. I thank Christopher Avery, Peter Hammer, Dan Kessler, Fran Miller, Mark Pauly, Ed Rock, Al Roth, Chris Sanchirico, Reed Shuldiner, Jennifer Stack, Raphael Thomadsen, Akila Weerapana, and participants of the University of Pennsylvania Law School and the University of Michigan Law and Economics workshops for their comments and suggestions.

congressional grant of an exemption was justified. It also suggests, however, that further action may be required to ensure that the matching process obtains the maximum possible social benefit.

TABLE OF CONTENTS

I. INTRODUCTION 760

II. MEDICAL EDUCATION AND THE MATCH 767

III. IS THE MATCH ANTI- OR PROCOMPETITIVE? 770

 A. *The Anticompetitive Match* 770

 B. *The Procompetitive Match* 777

 C. *Efficiency in Matching Markets: A Response to Professor Priest* 783

 1. *Does the Match Solve a Prisoner’s Dilemma?* 787

 2. *Do Early Offers Obtain Distributive or Productive Gains?* 790

 3. *Does Inefficiency Arise from Ordinal Ranking in the Match?* 795

 4. *Summary: Factors in an Efficiency Assessment of the Match* 800

 D. *The Role of Compensation in an Efficiency Analysis* 804

IV. THE EMPIRICAL EVIDENCE ON RESIDENT COMPENSATION 816

V. THE ANTITRUST EXEMPTION AND OTHER REGULATORY RESPONSES 826

VI. CONCLUSION 835

I. INTRODUCTION

Medical education has long been a subject of controversy, and with good reason: The stakes associated with the medical education process are unquestionably high. They are high for physicians, who invest considerable money, time, and effort in training, and whose investment return will reflect the nature and quality of training they receive. Stakes are high for teaching hospitals, which depend on medical residents to provide care within their facilities. Stakes are also high for governments, insurance companies, employers, and others who take on the

responsibility of paying for services provided by teaching hospitals. Finally, stakes are high for patients, the immediate beneficiaries of residents' services and the ultimate beneficiaries of residents' training, whose length and quality of life depend on the quality of medical education. With so much at stake for so many, it is not surprising that the interests of many individuals and institutions, including physicians, hospitals, and governments, have shaped the medical education process. It is also not surprising that conflicts over this process have emerged.

One of the most closely watched conflicts in recent years has been the 2002 lawsuit challenging the residency "match" as a violation of section 1 of the Sherman Act.¹ The residency match is the system by which fourth-year medical school students and other applicants are matched to residency programs across the country. Each year, residency candidates submit a list of preferred residency programs, and residency programs submit a list of preferred residency candidates. A computer algorithm designed to satisfy match participants' mutual preferences then assigns individual candidates to residency positions. Pointing to resident compensation that has been relatively static across time and low compared to that of nonresident physicians and other health care professionals, the lawsuit plaintiffs alleged that the match (along with residency program accreditation requirements and the exchange of compensation information) illegally restrained competition.²

1. See Complaint, *Jung v. Ass'n of Am. Med. Colls.*, 300 F. Supp. 2d 119 (D.D.C. 2004) (No. CIV.A.02-0873 PLF) [hereinafter Complaint], <http://www.savethematch.org/pdf/complaint.pdf>. For legal analyses of the suit, see Sanders H. Chae, *Is the Match Illegal?*, 348 NEW ENG. J. MED. 352 (2003); Frances H. Miller & Thomas L. Greaney, *The National Resident Matching Program and Antitrust Law*, 289 JAMA 913 (2003) (explaining the history of the match and the legal principles relevant to the *Jung* suit); Heather S. Crall, Note, *Unreasonable Restraints: Antitrust Law and the National Residency Matching Program*, 82 WASH. U. L.Q. 245 (2004); and Melinda Creasman, Note, *Resuscitating the National Resident Matching Program: Improving Medical Resident Placement Through Binding Dual Matching*, 56 VAND. L. REV. 1439 (2003) (arguing that the *Jung* antitrust claims are meritorious and proposing an alternative matching mechanism).

2. Complaint, *supra* note 1, ¶¶ 1–2. More specifically, the complaint alleged that the defendants violated section 1 of the Sherman Act in three ways. *Id.* ¶¶ 3, 97–103. First, the defendants exchanged information about medical residencies through various mechanisms, including an annual survey that aggregated salary information across residency programs. *Id.* ¶¶ 3, 73–82, 100. Second, the defendants operated or participated in the match. *Id.* ¶¶ 3, 83–86, 100. Third, in implementing accreditation standards, the organization responsible for accrediting residency programs limited the number of residency positions available, hindered resident movement among these positions, encouraged match participation, and reviewed employment terms. *Id.* ¶¶ 3, 87–88, 100. The complaint claimed that the defendants' alleged information exchange, match program, and accreditation activities had "the purpose and effect of artificially fixing, depressing, standardizing and stabilizing resident physician compensation and other

The plaintiffs, physicians who had participated in the match, sought damages and an injunction against antitrust violations.³ Among the defendants were the National Resident Matching Program (NRMP), which operates the match; the Accreditation Council for Graduate Medical Education (ACGME), which accredits residency programs; the Association of American Medical Colleges (AAMC), an NRMP sponsor which conducts an annual survey of resident compensation; and individual, nonpublic teaching hospitals, which participate in the match and employ residents.⁴

A plaintiff victory in the suit would have had profound implications for the graduate medical education system, affecting each of the stakeholders in the medical education process. In the short term, an award of damages would have benefited former residents but depleted the resources of defendant teaching hospitals. In the long term, if the plaintiffs were correct in arguing that the match depresses compensation,⁵ an injunction against the continued operation of the match would have resulted in a further transfer of resources from teaching hospitals to residents. The potential effects of abolishing the residency match, however, would have been confined neither to compensation issues, nor to match participants. Because the residency match affects the output of residency programs, including resident training, patient services, and medical research,⁶ its abolition would have affected society more broadly. Whether these effects would have been on balance beneficial or detrimental remains an open question.

Congress made no attempt to resolve this question publicly when in 2004 it created an antitrust exemption for the residency match.⁷ Versions of the Pension Funding Equity Act of 2004

terms of employment." *Id.* ¶ 101.

3. *Id.* ¶¶ 1, 10–13.

4. *Id.* ¶¶ 15–17, 22–50. In February 2004, the district court allowed the suit to go forward with respect to defendants National Residency Matching Program (NRMP), Accreditation Council for Graduate Medical Education (ACGME), and Association of American Medical Colleges (AAMC), as well as many of the individual teaching hospitals. *Jung*, 300 F. Supp. 2d at 173–74. The court granted the motions to dismiss some defendants, including the American Hospital Association, the American Medical Association, the Council of Medical Specialty Societies, and the American Board of Medical Specialties. *Id.* These four organizations and the AAMC together sponsor the NRMP. Complaint, *supra* note 1, ¶¶ 17–21. The court later dismissed the case. *See infra* note 13 and accompanying text.

5. *See supra* note 2.

6. *See* Accreditation Council for Graduate Medical Education, ACGME Fact Sheet, http://www.acgme.org/acWebsite/newsRoom/newsRm_factSheet.asp (last visited Oct. 1, 2005) (describing the functions of residency programs).

7. Pension Funding Equity Act of 2004 § 207, 15 U.S.C.A. § 37b (West Supp. 2005).

passed both the House and Senate with no mention of any residency-related antitrust exemption.⁸ It was not until the final bill emerged from the conference committee that a section captioned “Confirmation of Antitrust Status of Graduate Medical Resident Matching Programs” first appeared.⁹ This section, which applies to conduct both before and after the law’s passage,¹⁰ states that “[i]t shall not be unlawful under the antitrust laws to sponsor, conduct, or participate in a graduate medical education residency matching program,”¹¹ and prohibits the admission of match-related evidence to support an antitrust claim.¹² The match antitrust suit defendants moved for judgment on the pleadings, and in August 2004, the district court dismissed the case.¹³ Despite its powerful effect, the statute was signed into law without debate concerning the antitrust exemption either in the Senate Judiciary Committee or on the House or Senate floors,¹⁴ a state of affairs that elicited strong objections from several senators. Senator Kohl, for example, said that

[i]n general it is bad policy to provide exemptions to the antitrust laws. . . . We should have had the opportunity to debate this issue and determine whether there was any merit to the exemption, rather than see the exemption mysteriously appear on an unrelated bill. It appears that this provision, enacted in this way, is nothing more than a giveaway to one particular special interest. Without judging

8. See Pension Funding Equity Act of 2003, H.R. 3108, 108th Cong. (as passed by House, Oct. 8, 2003); Pension Stability Act, H.R. 3108, 108th Cong. (as passed by Senate, Jan. 28, 2004).

9. See H.R. REP. NO. 108-457, at 17 (2004) (Conf. Rep.).

10. See 15 U.S.C.A. § 37b(c) (“This section shall take effect on April 10, 2004, shall apply to conduct whether it occurs prior to, on, or after April 10, 2004, and shall apply to all judicial and administrative actions or other proceedings pending on April 10, 2004.”).

11. *Id.* § 37b(b)(2).

12. *Id.* (“Evidence of any of the conduct described in the preceding sentence shall not be admissible in Federal court to support any claim or action alleging a violation of the antitrust laws.”).

13. *Jung v. Ass’n of Am. Med. Colls.*, 339 F. Supp. 2d 26, 46–47 (D.D.C. 2004). After determining that the plaintiffs’ claims of antitrust conspiracy hinged on allegations related to the match, and that the newly enacted statute prohibited the admission of match-related evidence, the court dismissed the case. *Id.* at 36–39. The court also rejected the plaintiffs’ constitutional challenges to the antitrust exemption statutory provisions. *Id.* at 40–46. In February 2005, the court rejected a request to reconsider the dismissal; the plaintiffs subsequently filed notice to appeal. Myrle Croasdale, *Judge Upholds Dismissal of Match Lawsuit*, AM. MED. NEWS, Feb. 14, 2005, at 20, 20.

14. See 149 CONG. REC. H9285 (daily ed. Oct. 8, 2003); 150 CONG. REC. S223 (daily ed. Jan. 26, 2004).

the merits of the issue, we should have had an opportunity to explore it and make that decision for ourselves.¹⁵

This Article assesses the merits of the antitrust exemption by analyzing whether the residency match is desirable from a policy perspective. Only with a thorough understanding of the potential effects of the residency match can it be determined whether the exemption was improvidently granted and should therefore be modified or eliminated.¹⁶ Building on an analysis of arguments on both sides of the residency match debate, this Article explores the implications of the residency match for social welfare.

At the core of the plaintiffs' complaint is the allegation that the residency match has depressed compensation;¹⁷ mean resident compensation levels are mostly in the \$40,000 range.¹⁸ The role of the residency match in determining these compensation levels, however, is unclear. Match rules restrict the timing of offers and govern the matching of residents to programs, but say nothing whatsoever about resident stipends.¹⁹ The complaint alleged that the match "enabled employers to obtain resident physicians without . . . a bidding war,"²⁰ but failed to detail the precise mechanism by which the match depresses compensation.

This Article explores several ways in which the rules of the match, though silent on compensation issues, may nonetheless affect compensation levels. Match rules restricting the formation

15. 150 CONG. REC. S3979 (daily ed. Apr. 8, 2004). While there was no public debate on the exemption, it is likely that there was considerable discussion in private among legislators, legislative staff, and representatives of the affected parties.

16. Senator Jeff Bingaman, who has criticized the creation of the exemption and argued that the exemption does not apply to claims of price fixing, has said in a floor statement that he "will seek a future opportunity to raise this issue before this body." 150 CONG. REC. S5223 (daily ed. May 11, 2004). In addition, the Antitrust Modernization Commission, a twelve-member commission appointed by the President and congressional leaders, Antitrust Modernization Commission Act of 2002 § 11054(a)(1)–(5), 15 U.S.C. § 1 (Supp. II 2002), has proposed evaluating whether industry-specific exemptions, including the residency match exemption, should be eliminated or time-limited. Memorandum from the Immunities & Exemptions Working Group to All Comm'rs 2–6 (Dec. 21, 2004), <http://www.amc.gov/pdf/meetings/ImmunitiesandExemptions.pdf>.

17. Complaint, *supra* note 1, ¶¶ 73, 83–84.

18. ASS'N OF AM. MED. COLLS., 2004 AAMC SURVEY OF HOUSESTAFF STIPENDS, BENEFITS & FUNDING: NOVEMBER 2004 REPORT 6 tbl.2 (2004), <http://www.aamc.org/data/housestaff/hss2004report.pdf>. See *infra* Part IV for an assessment of the empirical data on resident compensation.

19. See NRMP, Match Participation Agreement for Applicants and Programs § 1.0 (2005), http://www.nrmp.org/res_match/policies/map_main.html (describing the functions of the matching program and noting that "the NRMP is not involved in establishing the terms of any residency or fellowship agreement").

20. Complaint, *supra* note 1, ¶ 84.

of early contracts undermine the credibility of early offers and other expressions of interest, limiting the ability of residency candidates to use them to solicit other offers or to bargain for increased compensation.²¹ Furthermore, several models in the theoretical economics literature indicate that under certain conditions, matching mechanisms can result in lower compensation, suggesting that there may be a basis for the plaintiffs' claim that the match is anticompetitive.²²

At the same time, however, the match facilitates competition by overcoming market failures that reduce the probability of forming efficient matches. Centralized matching mechanisms help ensure efficient pairings of residency programs and residents by delaying the competitive process, thus improving match participants' access to information, and by taking full account of participants' preferences.²³ If Congress was correct in its characterization of the matching process as "highly efficient" and "pro-competitive,"²⁴ then by facilitating efficient pairings the match expands the "output" of resident-residency program relationships. This increase in output could conceivably support an increase in resident compensation relative to a world without the match, and would certainly increase social welfare.²⁵

In a recent article, however, Professor George Priest has argued that mechanisms that impose restrictions on offer timing, like many other types of competitive restraint, actually reduce aggregate welfare.²⁶ To the extent that Priest's arguments apply to the residency match, they reinforce the plaintiffs' claims that the match is anticompetitive. This Article argues, however, that while Priest's work provides important insights, it does not conclusively demonstrate that early-offer restrictions reduce welfare, particularly in the context of medical residencies. While some match participants might gain from early-offer competition, a significant portion of these gains would result from the redistribution of resources among market participants, rather than an increase in aggregate welfare. An overall assessment of

21. See *infra* Part III.A.

22. See *infra* Part III.A.

23. See *infra* Part III.B.

24. One of the congressional findings is that "[a]ntitrust lawsuits challenging the matching process, regardless of their merit or lack thereof, have the potential to undermine this highly efficient, pro-competitive, and long-standing process." 15 U.S.C.A. § 37b(a)(1)(E) (West Supp. 2005).

25. See *infra* Part III.B.

26. George L. Priest, *Reexamining the Market for Judicial Clerks and Other Assortative Matching Markets*, 22 YALE J. ON REG. 123, 204 (2005); see also *infra* Part III.C (responding to Professor Priest's article).

welfare requires a careful weighing of the many factors affecting the benefits and costs associated with timing restrictions.

Much of the work examining the effects of matching mechanisms, both supportive and critical, has focused on the nature of the matching process and the quality of matches achieved, rather than on the potential interaction between matching mechanisms and compensation. This focus is natural, given that the match itself places no direct constraints on compensation. But if the match does affect compensation, it is important to consider the relationship between these effects and efficiency. This Article explains that institutional features of the residency market tend to limit the impact of lower compensation levels, and suggests that any such impact may be outweighed by the benefits reaped through the use of the matching mechanism.

Critics of the match may argue that such a sanguine appraisal of the effects of the match is inappropriate, given the low level of resident compensation relative to that of other health care professionals.²⁷ These differences may result from a variety of factors unrelated to the match, however, including differences in productivity levels, nonmonetary compensation in the form of training, and the value of the opportunity to enter the medical profession. While not conclusive, empirical comparisons of markets with and without matching mechanisms support the argument that the match does not significantly affect compensation.

Ultimately, this Article's analysis suggests that Congress was justified in stepping into the residency match controversy. Given its decision, however, Congress also has an obligation to assure that society gains the maximum possible benefit from the match. Toward this end, this Article proposes that Congress adopt measures that would require programs to offer binding sample contracts prior to the match and facilitate the creation of individually-negotiated contracts.

Part II of this Article explains the medical education process and the mechanics of the residency match. Part III explores the effects of the match on competition and efficiency. Section A discusses ways in which the match might depress compensation, while section B explains the argument that matching mechanisms benefit their participants by generating higher-quality matches than could otherwise be obtained. Section C expands on the analysis in the previous sections by evaluating

27. See discussion *infra* Part IV (comparing resident compensation with that of other health care providers).

the extent to which the match might be an economically efficient institution, using as a starting point Professor George Priest's recent work analyzing assortative matching markets.²⁸ Section D considers the relationship between compensation and efficiency in the context of the residency match. Part IV assesses the empirical evidence on resident compensation. Part V recommends that Congress preserve the antitrust exemption, but proposes additional regulations that would facilitate competition. Part VI concludes.

II. MEDICAL EDUCATION AND THE MATCH

To understand the conflict over the match, it is important to first understand the rigorous and lengthy formal training process that transforms college graduates into physicians. It begins with enrollment in medical school.²⁹ While medical school curriculums vary, they generally include two years of basic science education in areas such as anatomy, biochemistry, and microbiology.³⁰ This coursework is followed by two years of clinical rotations, during which students work with patients in supervised settings in practice areas such as internal medicine, psychiatry, and surgery.³¹

The training required to become a fully licensed physician extends beyond the clinical exposure in medical school. Medical school graduates must first receive the more intensive training offered by residency programs, which focus on developing the clinical skills and detailed knowledge necessary to provide high-quality medical care within a specialized field.³² Providing services to patients is an integral part of the educational process.³³ The length of residency programs varies by specialty; internal medicine residencies last three years, while a basic

28. Priest, *supra* note 26.

29. To prepare for medical school, college students generally take numerous basic sciences courses, including biology, chemistry, and physics. See Am. Med. Ass'n, *Becoming an MD: How Do You Become a Physician?*, <http://www.ama-assn.org/ama/pub/category/14365.html> (last visited Oct. 1, 2005). Most medical schools consider Medical College Admission Test (MCAT) scores as part of the highly competitive admissions process; the MCAT tests proficiency in the biological and physical sciences, among other areas. See Ass'n of Am. Med. Colls., *About the MCAT*, <http://www.aamc.org/students/mcat/about/start.htm> (last visited Oct. 1, 2005).

30. See Ass'n of Am. Med. Colls., *Curriculum Directory*, <http://services.aamc.org/currrdir/about.cfm> (last visited Oct. 1, 2005) (describing the medical education process).

31. *Id.*

32. AM. MED. ASS'N, *GRADUATE MEDICAL EDUCATION DIRECTORY* 9 (2003).

33. See *id.* at 10 (explaining that "[t]he education of resident physicians relies on an integration of didactic activity in a structured curriculum with diagnosis and management of patients").

surgery program may last five years.³⁴ Training within a subspecialty may extend physicians' formal education process by several more years.³⁵ A physician seeking specialized training in cardiology, for example, might first participate in an internal medicine residency program and then spend three years in a cardiovascular disease program.³⁶

Before physicians may practice independently, they must obtain a license from the state in which they intend to practice.³⁷ Standards for licensure vary by state, but generally include graduation from medical school, successful completion of an exam, and completion of at least one year of postgraduate training.³⁸ Residents are therefore not generally permitted to practice independently at the beginning of their residencies, but they may become fully licensed before the end of their residencies.³⁹

Most residency positions are assigned through the residency match. In 2005, the match involved 3,813 programs offering 24,012 positions, and 31,862 applicants, including 15,308 2005 graduates of American medical schools.⁴⁰ In their fourth year of

34. *Id.* at 423 (charting the length of time required for various medical specialties and subspecialties).

35. *Id.*

36. *Id.* at 92 (general requirements for internal medicine subspecialties); *id.* at 97–98 (specific requirements for residency education in cardiovascular disease); *id.* at 423 (program lengths).

37. *See, e.g.*, N.Y. State Educ. Dep't, License Requirements, <http://www.op.nysed.gov/medlic.htm> (last visited Oct. 1, 2005) ("Any use of the title 'Physician' or practice of medicine within New York State requires licensure.").

38. *See, e.g., id.* (describing New York's requirements); N.D. State Bd. of Med. Exam'rs, Physician Licensure Requirements, http://www.ndbomex.com/MD_Req.htm (last visited Oct. 1, 2005) (listing North Dakota's requirements).

39. While obtaining a license does not generally require successful completion of an entire residency program, obtaining board certification in a discipline does. Residencies in approved hospitals have been required by specialty boards since the 1930s. KENNETH M. LUDMERER, TIME TO HEAL: AMERICAN MEDICAL EDUCATION FROM THE TURN OF THE CENTURY TO THE ERA OF MANAGED CARE 87 (1999). These requirements continue today. For example, the American Board of Internal Medicine requires that physicians complete thirty-six months of training in an ACGME-accredited residency (or a residency accredited by specified Canadian organizations) to obtain certification. Am. Bd. of Internal Med., Certification Policies, Internal Medicine Policies, <http://www.abim.org/cert/policiesim.shtm> (last visited Oct. 1, 2005). While board certification is not a requirement for licensure or for practice in a specific area, it may be relevant to a patient's choice of provider, to a hospital's decision about the scope of a physician's practice within its facility, to an insurer's decision about the makeup of its provider panel, or to an insurer's decision about the provision of malpractice insurance.

40. NRMP, About the NRMP, http://www.nrmp.org/about_nrmp/index.html (last visited Oct. 1, 2005) [hereinafter NRMP, About the NRMP]. The match involves more than just fourth-year medical students and first-year residency positions. Match participants also include previous American medical school graduates seeking residency positions and physicians educated outside of American medical schools. NRMP,

medical school, students submit applications to residency programs. Most programs permit electronic submissions of applications and other supporting materials, including personal statements, recommendations, and transcripts, beginning in September.⁴¹ Programs invite selected applicants to on-site interviews, during which programs and students can collect more information about each other.⁴² Then, sometime in January or February, programs submit to the NRMP confidential ordered lists of their preferred applicants, and students submit confidential ordered lists of their preferred programs.⁴³ There is no limit to the number of applicants or programs that may be listed.⁴⁴ Match rules prohibit participants from making commitments prior to the match.⁴⁵

In March, a computer algorithm assigns matches of programs and students based on the rank-order lists that have been submitted.⁴⁶ The algorithm begins by tentatively assigning an applicant to the applicant's first choice program, if that program has listed the applicant and still has positions available.⁴⁷ If the first choice program has not listed the applicant, the algorithm will attempt to assign the applicant to his or her second choice program.⁴⁸ Similarly, if the first choice

Independent Applicants, http://www.nrmp.org/res_match/special_part/ind_app (last updated Aug. 2004). Also, some residency positions offer advanced training and are not open to first-year residents. See AM. MED. ASS'N, *supra* note 32, at 39 (discussing advanced training programs not available to first year residents in the field of anesthesiology).

41. AAMC, Electronic Residency Application Service, Timeline, Deadlines and Timing Issues, <http://www.aamc.org/students/eras/timeline/start.htm> (last visited Oct. 1, 2005).

42. See NRMP, The Application Process, http://www.nrmp.org/res_match/about_res/application_process.html (last visited Oct. 1, 2005) (referring to programs' interview requirements).

43. See NRMP, Dates of NRMP Matches, http://www.nrmp.org/about_nrmp/schedule.html (last visited Oct. 1, 2005) (listing the dates that rank-order lists open and close). The match also accommodates couples seeking to coordinate their position searches. See NRMP, U.S. Seniors: Couples, http://www.nrmp.org/res_match/special_part/us_seniors/couples.html (last visited Oct. 1, 2005) (describing the couples search algorithm).

44. See Miller & Greaney, *supra* note 1, at 914 (explaining that up to fifteen preferences are included in the registration fee for applicants, and that applicants may rank additional preferences for thirty dollars per listing).

45. NRMP, *supra* note 19, § 6.0.

46. See NRMP, 2006 Main Match Schedule, http://www.nrmp.org/res_match/yearly.html (last visited Oct. 1, 2005) (indicating that match results will be posted on March 16, 2006); NRMP, How the Matching Algorithm Works, http://www.nrmp.org/res_match/about_res/algorithms.html (last visited Oct. 1, 2005) [hereinafter NRMP, How the Matching Algorithm Works] (explaining the matching algorithm).

47. NRMP, How the Matching Algorithm Works, *supra* note 46.

48. *Id.*

program has listed the applicant but the program is full and has ranked all of its tentatively matched candidates more highly than the applicant, the algorithm will attempt to assign the applicant to his or her second choice program.⁴⁹ On the other hand, if the first choice program is full, but the program prefers the applicant to at least one of its tentatively matched candidates, the algorithm will tentatively assign the applicant to the program.⁵⁰ To accommodate this tentative assignment, the algorithm will remove the program's least-preferred tentatively assigned candidate.⁵¹ The algorithm will subsequently attempt to assign the dropped candidate to the next program on the dropped candidate's list.⁵² Through this iterative temporary-assignment process, the algorithm ultimately ensures that each student is admitted to the program he or she prefers the most if it is not filled by students that the program prefers more. Students who are left unmatched at the end of this process may reapply the following year, or they may join the "Scramble" in which they apply for positions that have remained unfilled.⁵³

In 2005, about 94% of U.S. medical school seniors participating in the match were assigned successfully,⁵⁴ the majority to their first-choice program.⁵⁵ Applicants who are matched are required to join their assigned programs.⁵⁶ It is this matching mechanism that is at the center of the antitrust complaint.

III. IS THE MATCH ANTI- OR PROCOMPETITIVE?

A. *The Anticompetitive Match*

Although only recently challenged as a violation of antitrust law, the residency match has been a feature of the medical

49. *Id.*

50. *Id.*

51. *Id.*

52. *Id.*

53. NRMP, Independent Applicants, http://www.nrmp.org/res_match/special_part/ind_app/match_results.html (last visited October 1, 2005) (describing the "Scramble").

54. Press Release, Nat'l Resident Matching Program, U.S. Medical School Seniors Apply to Residency Programs in Record Numbers (Mar. 17, 2005), available at <http://www.aamc.org/newsroom/pressrel/2005/050317.htm>.

55. Among U.S. seniors who were matched, 62.5% received their first choice, 15.0% their second choice, and 8.7% their third choice. The remaining 13.8% were matched with their fourth-ranked program or lower. NRMP, Percent Matches by Rank Number, <http://www.aamc.org/newsroom/pressrel/2005/matchcharts.pdf> (last visited Oct. 1, 2005).

56. See NRMP, *supra* note 19, § 5.1 ("Failure to honor this commitment by either party participating in a match will be a material breach of this Agreement . . .").

education system since 1951.⁵⁷ The motivation for the adoption of the match is controversial. In their complaint, the plaintiffs allege that the match was intended to depress compensation in a market in which the demand for residents exceeded the supply.⁵⁸ Specifically, they allege that

[i]n 1952, the hospitals and other entities employing resident physicians determined that the continuation of free competition in recruiting, hiring, employing and compensating resident physicians was undesirable because the number of available residency positions outpaced the number of available candidates. Employers determined that continued free competition would “bid up” compensation and other terms of employment by which employers commonly compete to attract employees. Creating the matching program enabled employers to obtain resident physicians without such a bidding war⁵⁹

In fact, the number of positions available to first-year medical school graduates exceeded the number of applicants from around the turn of the century, when postgraduate training was first introduced, until the mid-1970s.⁶⁰ The number of positions continues to exceed the number of graduates from U.S. medical schools (although not the total number of applicants).⁶¹ In a competitive market, if demand exceeds supply of a good or service, the price would ordinarily rise, and supply would increase and demand would decrease, until the market equilibrated.⁶² The surplus in residency positions therefore would have placed upward pressure on the compensation offered to applicants. In theory, then, the match may have been an anticompetitive measure adopted in response to this pressure.⁶³

By imposing constraints on the competitive process, the residency match affects the nature of competition but does not

57. Alvin E. Roth, *The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory*, 92 J. POL. ECON. 991, 992, 995–96 (1984).

58. Complaint, *supra* note 1, ¶ 84.

59. *Id.*

60. Roth, *supra* note 57, at 992–93 & 993 n.1.

61. See NRMP, About the NRMP, *supra* note 40 (reporting that in the 2005 match, 15,308 2005 U.S. graduates and 16,554 independent applicants applied for 24,012 positions).

62. See, e.g., PAUL A. SAMUELSON & WILLIAM D. NORDHAUS, ECONOMICS 55–61 (17th ed. 2001) (discussing the general principles of supply and demand).

63. The ratios of applicants to available positions likely differ depending on the nature of the position. To the extent that residency positions are imperfect substitutes for one another, an excess supply of one type of position would not necessarily place upward pressure on wages for another type of position. Programs with unfilled positions, however, would have reason to advocate the adoption of a mechanism that might allow them to obtain residents without having to increase compensation.

eliminate it. Under a match system, programs compete indirectly for residents by competing for high rank-orders. A program might try to increase its place in applicants' rank-orderings by touting all of its positive attributes, including high compensation. Anecdotal evidence in fact suggests that programs have competed based on compensation in the post-match period.⁶⁴ While match rules prohibiting the creation of pre-match contracts preclude a binding commitment to pay any particular level of compensation, programs may disclose intended future program characteristics to applicants.⁶⁵ Students for whom compensation is important are therefore likely to obtain compensation information before the ranking process begins. It seems, then, that the main effect of the match would be to shift the competitive process from the employment offer stage to the rank-order stage. If this were the only effect, the match would impose no downward pressure on wages.

On the other hand, the nature of competition within the match will not necessarily mirror the competition that would occur outside of a match system. Imagine a labor market in which skilled workers are in high demand. Imagine, furthermore, that there is some uncertainty about the quality of applicants. Some applicants will arrive with knowledge, skills, talents, and experience that allow them to generate significant surpluses (profits) for an employer through their efforts in the workplace; those of lesser ability will generate lower surpluses. Even candidates with similar educational backgrounds and formal training may differ widely in their productivity. Assessing this productivity is difficult and costly.

64. See LUDMERER, *supra* note 39, at 193.

Salary became a weapon in the efforts of teaching hospitals to recruit the best house officers, particularly when they were competing against hospitals in the same city or geographical region. . . . Jefferson Medical College, for instance, repeatedly raised its house staff pay scale in response to competition from the other Philadelphia teaching programs. Students and house officers reveled in this competition. . . . [T]hey made their desire for higher incomes known, and leaders of even the most prestigious programs had to listen.

Id. (footnotes omitted).

65. In the past, NRMP rules precluded pre-match contracts but encouraged information disclosure. For example, the August 26, 2002 "Policies of the NRMP" states that "any verbal or written contracts prior to the submission of Rank Order Lists is [sic] a violation of the Match," (policy 8.0) but that "programs are expected to provide complete and accurate information to interviewees, including a sample contract and institutional policies regarding eligibility for appointment to a residency position" (policy 6.2). Policies of the NRMP (2002) (on file with author). In 2003, the NRMP decided to require programs participating in the match to disclose before the rank-order list deadline the contract that successfully matched applicants would be expected to sign. Press Release, Nat'l Resident Matching Program, NRMP Requires Medical Residency Programs to Show Contracts to Applicants (Oct. 31, 2003), *available at* <http://www.nrmp.org/contractpr.pdf>.

In this setting, a job candidate may visit several potential employers. A potential employer evaluates the candidate's credentials and decides to make an offer, say for \$50,000, because it believes that the candidate can produce more than \$50,000 in benefit for the employer. The existence of an offer that would be contractually binding if accepted communicates to other similarly situated employers that according to one assessment, the candidate would produce more than \$50,000 in surplus. An employment offer with compensation attached is a signal of the candidate's minimum quality. To the extent that the employers' needs are similar, a competing employer could use this information as a baseline for making its own offer. It could use the existence of outside offers as a low-cost way of sorting more productive from less productive candidates, and then conduct a further assessment to determine whether the candidate would likely produce even more than \$50,000 in surplus for its own organization. If so, it could make the candidate an offer of more than \$50,000. The candidate could then return to the initial offeror, or move on to other potential employers, to solicit a still more lucrative offer. In an ordinary competitive labor market, workers can use offers from one potential employer to extract offers and higher compensation from competing potential employers.⁶⁶

In the residency match setting, residents cannot take full advantage of this type of bargaining in order to increase either the likelihood of competing offers or the amount of compensation, benefits, or other perks. Residents may certainly suggest to programs that their benefits packages are not as attractive as those of other programs, but they cannot negotiate with an employer from a position of strength based on a competing offer.⁶⁷ Programs participating in the match are not permitted to form a contract outside of the match for a position that is to be allocated

66. In this scenario, later potential employers are able to free ride off of the initial potential employer's investment in assessing the quality of the employee. Not only does the initial employer bear the full cost of that investment, but by making an offer it invites competition for the services of the employee. The initial employer could try to limit the employee's ability to seek competing offers by holding the offer open for only a short time, but to prevent the free-riding problem the employer would have to find a way to prevent disclosure of the existence of the offer. The discussion of this scenario assumes that employers would not successfully do so.

67. Dr. Sanders Chae makes this point when he argues that "[w]ithout an offer in hand, students have no power to negotiate with the programs, and programs do not need to compete for residents through salary or other monetary benefits." Chae, *supra* note 1, at 353. This passage overstates the case, however. To the extent that residents care about the immediate monetary benefits of a residency position, programs have an incentive to compete to attract high rankings from students they prefer.

through the match process.⁶⁸ Nor are match participants permitted to solicit information about rank-orderings from one another.⁶⁹ While it is permissible to volunteer ranking information,⁷⁰ and it is conceivable that a program would make an offer prior to the match for a position included in the match, neither action would constitute a credible commitment. Positions are allocated based solely on rank-orderings actually submitted, not on rank-orderings compiled previously or other information.⁷¹ Commitments become binding only after the match takes place.⁷² As a result, any information provided before the match is suspect and cannot be used in the same way that formal offers are used in a more conventional workplace setting—to solicit an offer or improve on a compensation package.

Although there is no formal sanction for deviating from a suggested rank-ordering or benefits package, reputational sanctions may discourage misleading representations.⁷³ A program that widely advertises that it will offer a high level of compensation or adopt a particularly attractive feature, for example, might face a decline in applications from suspicious future applicants if it later reverses its decision. If reputational sanctions are effective, nonbinding commitments may be helpful to a candidate in the same way formal offers would be.

Employment negotiations typically occur at the individual level, however, where reputational sanctions may be less effective. In fact, studies suggest that reputational sanctions (if any) have been unsuccessful in deterring misleading behavior

68. NRMP, Match Participation Agreement for Institutions § 4.2, http://www.nrmp.org/res_match/policies/map_institution.html (last visited Oct. 1, 2005).

69. NRMP, *supra* note 19, § 6.0.

70. *See id.*

71. Section 8.0 of the NRMP's Match Participation Agreement for Institutions reads in part as follows:

It is a material breach of this Agreement and of the Match Participation Agreement . . . for a program that is participating in the Matching Program to make any verbal or written contract for appointment to a concurrent year residency position prior to the Matching Program. In addition, although applicants or programs may volunteer how they plan to rank each other, it is a material breach of this Agreement and of the Match Participation Agreement . . . to request such information. Only the final preferences of programs and applicants, as reflected in their final certified rank order lists, will determine the offering of positions and the placement of applicants through the Matching Program.

NRMP, *supra* note 68, § 8.0.

72. *See* NRMP, *supra* note 19, § 5.1 (noting that a binding commitment is established if a successful match results).

73. *See id.* § 4.3 (indicating that although “programs are expected to provide complete and accurate information to interviewees,” the “NRMP is not responsible for ensuring the accuracy of information exchanged between applicants and programs”).

that occurs within the context of the match. For example, over seventy percent of one survey's respondents agreed with a statement that "[a]pplicants often make dishonest or misleading assurances or statements to programs about their level of interest."⁷⁴ Another survey-based study reports that more than ninety percent of family practice program directors believed that applicants had occasionally lied to them.⁷⁵ Program directors may also engage in misleading behavior; the same family practice program survey found that while almost no directors would lie to candidates asking about the programs' rankings of candidates, the majority would give only a vague response, and a significant number would give a "vague but positive response" to low-ranked applicants.⁷⁶ In this setting, applicants would be unlikely to be able to use information about competing programs' prospective rankings as an effective negotiation tool.⁷⁷

The implication of the match's restrictions and the participants' behavior is that there is little foundation for negotiation with individual candidates.⁷⁸ An applicant cannot wait until a program makes an offer and then try to gain a larger percentage of the surplus created by the match through

74. John Bernard Miller et al., *Communication Between Programs and Applicants During Residency Selection: Effects of the Match on Medical Students' Professional Development*, 78 ACAD. MED. 403, 408 (2003). The majority of students also thought that misleading statements helped improve their position in the match, implying that they believed that program directors rely on their statements in some way despite their lack of credibility. *Id.* at 408 tbl.5. On the other hand, another study reported that over eighty percent of program directors were skeptical of or did not believe a candidate's claim that the program was a highly ranked choice. Peter J. Carek et al., *Recruitment Behavior and Program Directors: How Ethical Are Their Perspectives About the Match Process?*, 32 FAM. MED. 258, 259 (2000).

75. Carek et al., *supra* note 74, at 259.

76. *Id.* at 259 tbl.1. When communicating with low-ranked applicants, 6.0% of program directors said they would respond honestly, 30.1% would give a vague but positive response, 23.2% would give a vague but negative response, 0.3% would lie, and 40.5% would indicate that they were not permitted to share this information with the applicant. *Id.*

77. See Peter J. Carek & Kimberly D. Anderson, *Residency Selection Process and the Match: Does Anyone Believe Anybody?*, 285 JAMA 2784, 2784-85 (2001), for a discussion of ethics within the residency match.

78. On the other hand, it has been asserted that negotiations outside of the match nonetheless occur. One observer offers the following comments about the match:

Here, we find a selection process riddled with backstage deals, negotiated perks, and signing bonuses (of \$10,000 and up). With residency slots far exceeding the number of U.S. medical school graduates, programs battle to entice the "best and the brightest." . . . Meanwhile, applicants are well aware that program directors fear "not filling" and therefore play programs off against each other to wrangle the "best deal."

Fred Hafferty, *Finding Soul in a "Medical Profession of One,"* 28 J. HEALTH POL. POL'Y & L. 133, 147 (2003) (book review) (citations omitted).

negotiations and a threat to go elsewhere. While competition will tend to keep compensation within a reasonable range—applicants may avoid programs with histories or promises of poor compensation packages—applicants do not have the extra leverage that they would have in a conventional labor market characterized by unfettered, individualized competition among both future employers and future employees. Once the match is officially made, the resident is bound to accept the position at the offered salary.⁷⁹

Moreover, academic work focusing specifically on the question of matching and wages suggests that matching mechanisms similar in general form to the NRMP may be associated with lower wages. Ulrich Kamecke demonstrates this point with a theoretical model and concludes that “the competition for interns is not sufficient to force wages all the way up to their competitive level. In equilibrium the interns are in general worse off than in an ideal decentralized market.”⁸⁰ The general intuition underlying this result is that when programs make their wage offers through the match, the stated wage only has to be sufficiently high that the candidate will rank that program more highly than his or her next best choice. In a competitive market, a candidate’s ability to seek employment with a competitor pressures the initial hospital to raise its wages, allowing the candidate to extract more of the surplus created through the employment relationship. Without this opportunity to negotiate, the resident’s pay is not as high as it otherwise would be.⁸¹ A recent working paper by Jeremy Bulow and Jonathan Levin using a different matching model finds that when programs do not make individualized salary offers, salaries will be lower than in a competitive equilibrium, particularly for the most sought-after residents.⁸² The features of these papers’ models do not conform precisely to those of the residency match, but the results suggest the effects the residency match may have on resident compensation. The models are discussed in more detail in subpart III.C.4 below.

79. In addition, programs may solidify their bargaining position by not negotiating individually over salaries. Historically, the characteristics of the position (including duties and compensation) were specified in advance of the match, and were not subject to negotiation. Roth, *supra* note 57, at 995–96.

80. Ulrich Kamecke, *Wage Formation in a Centralized Matching Market*, 39 INT’L ECON. REV. 33, 34–35 (1998).

81. *Id.* at 42, 48.

82. Jeremy Bulow & Jonathan Levin, *Matching and Price Competition* 1–2 (Stanford Univ., Working Paper No. 1818, 2003), available at <http://www.stanford.edu/group/SITE/Levin.03.pdf>.

B. The Procompetitive Match

While some have argued that the match is anticompetitive, others have argued that the match is in fact procompetitive.⁸³ Advocates of the match, including the NRMP itself, cite as its chief advantage its beneficial effect on the timing of residency offers and acceptances.⁸⁴ Professor Alvin Roth, an expert in the economics of matching mechanisms who assisted in a relatively recent redesign of the residency match, has written extensively about market timing questions.⁸⁵ His academic work focuses not on compensation issues, but on the benefits that arise from channeling competition through a matching mechanism.⁸⁶

While the residency antitrust suit complaint argues that the motivation for adopting the match was relief from competition-induced upward pressure on wages,⁸⁷ Roth's historical account of the catalyzing pressures focuses on the timing of offers and acceptances of residency positions.⁸⁸ He explains that as a result

83. For general descriptions of benefits of a matching system, see Miller & Greaney, *supra* note 1, at 915–16, and Creasman, *supra* note 1, at 1462–63. For a discussion of potential benefits (and drawbacks) of a formalized matching process in the context of the judicial clerkship market, see generally Christopher Avery, Christine Jolls, Richard A. Posner & Alvin E. Roth, *The Market for Federal Judicial Law Clerks*, 68 U. CHI. L. REV. 793 (2001).

84. See NRMP, How the NRMP Process Works, http://www.nrmp.org/about_nrmp/how.html (last visited Oct. 1, 2005) (“Because it provides a uniform date for decisions about residency selection for both applicants and programs, the NRMP matches eliminate the pressure that might otherwise fall upon applicants and programs to make decisions before all of their options are known.”).

85. Roth was hired in 1995 to redesign the residency matching system and to evaluate the redesign's potential effects. See Al Roth's Game Theory and Experimental Economics Page, <http://kuznets.fas.harvard.edu/~aroth/alroth.html> (last updated June 16, 2005) (an excellent resource for information about matching). The main feature of this redesign was to move from a “program-proposing” algorithm to an “applicant-proposing” algorithm to address concerns that the algorithm was too favorable to programs and too susceptible to strategic manipulation. Roth's study found that the effects of the subsequently adopted redesign would be small. See Alvin E. Roth & Elliott Peranson, *The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design*, 89 AM. ECON. REV. 748, 748, 754–55, 773 (1999).

86. See Al Roth's Game Theory and Experimental Economics Page, *supra* note 85 (listing Roth's publications).

87. Complaint, *supra* note 1, ¶ 84.

88. Roth, *supra* note 57, at 992–97. Some medical school graduates of the era in which the match was adopted echo Roth's account. In response to an article describing the antitrust issues raised by the match, W. Hardy Hendren writes,

I graduated from medical school in 1952, the year the National Resident Matching Program was initiated. As a medical student, I spent five months working on the Match

The Match was designed to assist medical students in choosing the best possible internship.

W. Hardy Hendren, Letter to the Editor, 348 NEW ENG. J. MED. 2260, 2260 (2003). Another physician writes,

of competition (presumably the same competition over residents discussed by the plaintiffs in their complaint), the date of agreements between applicants and programs began to creep backward.⁸⁹ Each program wanted to make sure it could obtain commitments just a bit before competitor programs.⁹⁰ The effect of this creep, or “unraveling,” as Roth terms the phenomenon in other work on market transaction timing,⁹¹ was that offers were made at increasingly early points in medical school students’ education. By 1944, the standard appointment date had become the beginning of the student’s junior year, two years before the residency would begin.⁹²

This unraveling was problematic because it meant that both students and programs made decisions on the basis of information that they possessed at an early stage in the students’ education. Students who had completed two years of medical school may not have had much exposure to clinical practice, and may not have had much inclination about their own career preferences, which would influence their residency position choices. Residency programs, meanwhile, would have had limited student performance data, impeding their ability to target offers to the most qualified or appropriate candidates.⁹³ Many efforts were made to coordinate appointment dates voluntarily, but they failed.⁹⁴ The unraveling phenomenon and the difficulty of resolving it are likely familiar to lawyers, judges, and legal academics, who have experienced both in the market for federal judicial clerks.⁹⁵ In the market for residents, the problem was

Physicians of my era (I attended medical school from 1948 to 1952) remember seeing senior students sleeping in phone booths waiting for a call from their first-choice hospital but having to respond to the rest of their choices by a deadline that was only days or hours away—typically, 72 hours.

The Match was initiated in 1952 by my class, led by Hardy Hendren, now a professor emeritus at Harvard Medical School. The medical students—not the hospitals—arranged the Match.

James A. Pittman, Jr., Letter to the Editor, 348 *NEW ENG. J. MED.* 2260, 2260 (2003) (footnote omitted).

89. Roth, *supra* note 57, at 993–94.

90. *Id.* at 993.

91. See Alvin E. Roth & Xiaolin Xing, *Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions*, 84 *AM. ECON. REV.* 992, 994 (1994) (using the term “unraveling”).

92. Roth, *supra* note 57, at 994.

93. See, e.g., *id.* at 993 (noting that hospitals “had to appoint interns without knowing their final grades or class standings”).

94. *Id.* at 993–97.

95. For a description of the problems associated with early hiring in the clerkship market, including informational deficiencies, see Avery et al., *supra* note 83, at 801–04. Recent efforts by federal appellate judges to mitigate the unraveling problem through a hiring moratorium and an agreement to hire students no earlier than the fall of their

2005]

THE RESIDENCY MATCH

779

ultimately resolved when the AAMC adopted a proposal to refuse to release transcripts or reference letters before the end of a student's junior year.⁹⁶

Another problem soon arose, however.⁹⁷ A program ("Program B") would make an offer to a candidate, but the candidate would not immediately accept it because the program was the candidate's second choice. The candidate would instead wait until he heard from his first choice ("Program A"), while Program B continued to hold open his position. (Meanwhile, Program A may have been keeping the candidate waiting because it was holding open an offer for another candidate.) By holding open its offer in the hope that the candidate would eventually accept it, Program B may have sacrificed an opportunity to make an offer to a third candidate, who may have preferred to join Program B, but who accepted an offer from a third program out of the fear she would not get another offer. To avoid losing the opportunity to recruit other candidates, programs would want to hold open offers for as little time as possible. In 1949, for example, the AAMC proposed that offers should be made just after midnight on November 15, and should be held open at least until noon the same day.⁹⁸ There was pressure for offers to be open for even less time.⁹⁹ This exploding offer phenomenon, also historically present in the market for judicial clerks, is problematic because candidates who do not want to risk losing an offer may be forced to accept it even though a more desirable position might be offered subsequently.¹⁰⁰ The phenomenon also may lead to early acceptances that are later withdrawn, causing considerable disruption for programs that believe they have a commitment satisfactory to all parties.¹⁰¹

The match resolved both of these problems while preserving competition among both programs and students. If all participants commit to the match, the backwards-creep problem

third year of law school are documented at Summary of the Law Clerk Hiring Plan, <http://www.cadc.uscourts.gov/internet/lawclerk.nsf/Content/SUMMARY?OpenDocument> (last visited Oct. 1, 2005). For an article comparing the judicial clerkship and residency markets, see Annette E. Clark, *On Comparing Apples and Oranges: The Judicial Clerk Selection Process and the Medical Matching Model*, 83 GEO. L.J. 1749 (1995).

96. Roth, *supra* note 57, at 994.

97. See *id.* at 994–95 (describing the exploding offer problem).

98. *Id.* at 995.

99. *Id.*

100. For a discussion of exploding offers within the law clerk market, see Avery et al., *supra* note 83, at 816–25.

101. Roth, *supra* note 57, at 994 (noting hospitals' frustration when candidates revoked their acceptances).

is resolved because participants do not form binding contracts outside of the match. This means that binding commitments cannot be made before the spring of medical students' senior year. The match resolves the exploding offer problem by using a mechanism that precludes the existence of such devices. Based on the parties' expressed preferences, the matching algorithm generates a single offer that the future resident is bound to accept.¹⁰² The match thus places bounds on the competitive process, bounds that almost certainly alter the matches that are ultimately formed. The important question for participants is whether the results of this process are preferable to the results of unfettered competition.

Historical data and information from similar markets, such as the market for law clerks, suggest that if competition were completely unconstrained, contracts would almost certainly be established earlier in students' educations.¹⁰³ Earlier in the competitive process, programs have less information about the true quality of students. They have limited information about the students' abilities and no information about the improvement in those abilities over the course of the students' training. Furthermore, while students may not lack information about programs, they may have limited information about their own preferences with respect to these programs, particularly if most of their experience has been in a classroom rather than clinical setting. In addition, other changes in students' lives during their medical school years may alter their preferences with respect to program attributes (geography, for example).

The participants' lack of information means that the preferences they express at an early stage in the offer process may not match the underlying "true" preferences they would have in an environment with more information. If the residency match were held in the spring of a student's second year of medical school, rather than in the spring of the fourth year, the results would be quite different. As more time passes, information accumulates, and the correspondence between expressed preferences and true preferences improves. The later a match is formed, the more likely its results will maximize participant satisfaction. All else equal (an important caveat, as discussed in subpart III.B.4 below), participants in the competitive process will prefer later decisions to earlier decisions.

102. See *supra* Part II (describing the match process).

103. See Avery et al., *supra* note 83, at 805–06 (describing the "backward progression in the time of hiring federal judicial law clerks").

The second way in which the match alters the results relative to the unconstrained competitive process also involves information failures. Imagine that competition over residencies, while otherwise unconstrained, is delayed until the students' fourth year. At this point, programs are fully aware of student characteristics, and their own preferences with respect to those characteristics. Candidates are fully aware of program characteristics, and their own preferences with respect to those characteristics. Each participant in the process could thus generate a complete rank-ordered list of fully-informed preferences; the higher the ranking, the higher the participant's satisfaction from the resulting match. Nonetheless, information in this market is still less than perfect. In particular, students are not ordinarily aware of programs' preferences with respect to other students, or other students' preferences with respect to programs. Programs are not ordinarily aware of students' preferences with respect to other programs, or other programs' preferences with respect to students. Participants are therefore not fully informed about the nature of the market, and more specifically, the likelihood that any particular offer will be made or accepted.

Assume, for example, that a student prefers program A to program B, and program B to program C. If program B makes an exploding offer, the student may accept only because the student is not sufficiently familiar with A's and other students' preferences to determine whether A will eventually make him or her an offer. This situation need not occur if a matching mechanism is in operation. The match allows residents to indicate their preferences with respect to multiple states of the world: "I would love to join program A, but if no position is available at A, then I'd be happy to join program B, and if not B, then C." A resident need not accept an offer from program B only because a deadline looms and the resident lacks information about the availability of positions at program A. The computer has access to complete information about all participants' preferences, allowing it to satisfy them all simultaneously.¹⁰⁴ The match therefore generates different results than the competitive process would in the absence of a match, not just because competition occurs later in the students' education than it otherwise would, but also because it can make use of all information available in the market.

104. See *supra* notes 46–53 and accompanying text (describing the matching algorithm).

The match produces different results from unconstrained competition by ensuring that participants can act upon more fully-informed preferences,¹⁰⁵ and by precluding strategic behavior that takes advantage of information gaps. The match algorithm is designed to obtain “stable” matches that reflect participants’ preferences, as expressed in their rank-orderings.¹⁰⁶ If a match is stable, no candidate or program will be matched to an unacceptable partner. The match algorithm necessarily achieves this result because it considers only matches suggested by the participants’ rank-order lists; if no acceptable match is possible, the participant will remain unmatched.¹⁰⁷ In addition, if a match is stable, there can be no scenario in which a program and a candidate would both prefer to be matched to each other rather than to the partner assigned to them by the computer. If a program and candidate did prefer each other to their assigned matches, they would have an incentive to deviate from the initial match, destabilizing it. Thus, for a matching program to be successful, the algorithm used must produce a stable match. Professor Roth has found that previous NRMP match outcomes were stable.¹⁰⁸

In other words, if the preferences as reflected in the rank-orderings are taken as given, none of the participants could successfully “contract around” the results of the match. A resident soliciting programs higher on his or her list would find that none would be interested in accepting the resident, because they would already be filled with more desirable candidates. A program soliciting residents higher on its list would find none that would be interested in joining, because the residents had

105. The presence of a matching mechanism may make candidates’ preferences less informed in one respect. In a standard market, job candidates may use the existence and/or timing of a firm offer to gauge a potential employer’s intensity of preference. See Priest, *supra* note 26, at 155–56 (discussing timing as an indicator of intensity of preference). Centralized matching mechanisms prevent this expression of intensity. If a candidate’s preference for a firm is ordinarily determined in part by the firm’s preference for a candidate, then a centralized matching mechanism deprives the candidate of relevant information. *But see infra* Part III.C.4 (arguing that offer timing is a poor signal of preference intensity).

106. In his work, Roth defines stability and analyzes the extent of its existence. See, e.g., Roth, *supra* note 57, at 998 (“A given outcome x is called *unstable* if some student or hospital receives an unacceptable assignment or if there exists a hospital program h , and student s , who each prefer the other to their assignment at x .”); *id.* at 1001 (finding the existence of a stable outcome).

107. See *supra* notes 46–53 and accompanying text (describing the matching algorithm).

108. Roth & Peranson, *supra* note 85, at 754. For a discussion of stability in simple matching mechanisms and more complex matching mechanisms such as the NRMP, see *id.* at 752–55.

already been accepted into programs they preferred. If participants' preferences were fully informed, accurately stated, and static, there would be no way to reassign residents to programs after the match has occurred that would better satisfy the preferences of some participants, without making others worse off. Because the competitive process generates initial matches based on choices made with less than complete information, the same cannot be said for unconstrained competition.

The match could be said to be procompetitive because it facilitates the acquisition of information in order to generate a result that market participants could not easily improve upon through mutual agreement. Because match participants' preferences are related to the output of residency programs, matching mechanisms will ultimately help maximize this output, including resident training, patient care, and medical research.¹⁰⁹

C. Efficiency in Matching Markets: A Response to Professor Priest

In a recently published article, *Reexamining the Market for Judicial Clerks and Other Assortative Matching Markets*, George Priest has challenged the theory that matching mechanisms benefit competition.¹¹⁰ The article provides a comprehensive analysis of unraveling markets and the effects of offer-timing restrictions on aggregate welfare.¹¹¹ Priest offers two core insights about matching markets and the unraveling phenomenon. First, he explains that for the concept of unraveling to have coherence, the point at which a market "begins" must be identifiable.¹¹² Then, by definition, a market unravels when transactions take place before this time, or more specifically, before some perceived "optimal market beginning time."¹¹³ But Priest takes issue with the notion that a market as a whole can have an optimal

109. Avery and his co-authors point out that the quantity of social benefit produced by the matching system will depend on how the benefit is created. Specifically, if the benefit produced is simply an additive function of the inputs of match participants, the match will have no effect on the total quantity of benefit produced. On the other hand, if the benefit production function is multiplicative, then it is important that better-quality participants be matched to each other. See Avery et al., *supra* note 83, at 804.

110. Priest, *supra* note 26.

111. See generally *id.* The discussion *infra* focuses on the paper's theory of market unraveling and its implications for the residency market. It thus sacrifices the nuance and detail of Priest's re-analysis of the judicial clerk market, and omits mention of the clerkship-related evidence Priest uses to support his arguments.

112. *Id.* at 151.

113. *Id.*

beginning time.¹¹⁴ He explains that typically, we assume that market participants search for contracting partners as long as the marginal benefits of the search exceed its marginal costs.¹¹⁵ Benefits of search include the improved matches made possible by superior information, while costs include the resources used to obtain that information. Each market participant makes its own decision about when to engage in a transaction, based on perceived costs and benefits. In this setting, it is not clear why unraveling should have any normative significance.¹¹⁶

Second, Priest argues that the cause of unraveling is artificially-imposed constraints on the dimensions along which market participants can compete.¹¹⁷ Many markets might be said to have natural “beginnings,” yet no unraveling occurs.¹¹⁸ So what distinguishes these entry-level markets from markets for clerkships or residencies? Priest points out that for most jobs, current salaries, future salaries, working conditions, work stability, and other characteristics vary.¹¹⁹ Employers hoping to attract more workers, a particular type of worker, or specific individual workers would adjust these characteristics accordingly; employers signal their intensity of preference through the adjustments they make. But the market for federal judicial clerks is different: job conditions are similar, salaries fixed.¹²⁰ Clerkships may vary substantially in terms of prestige, but again this characteristic is fixed, not negotiable.¹²¹ For this reason, Priest explains, a judge’s offer is by its nature a take-it-or-leave-it offer.¹²² The only dimension along which judges can

114. See *id.* at 149–53 (“As a general matter, the concept of an optimal beginning date of a market is an artifact . . .”).

115. *Id.* at 151.

116. *Id.* at 151–52.

117. *Id.* at 153–56. Note, however, that an assertion that fixed compensation (or other terms) is necessary for markets to unravel is inconsistent with economic work that has examined equilibria in early contracting markets. Wing Suen considers a market in which the quality of firms is known, the quality of workers is initially unknown but later revealed, and workers are risk averse—not an unreasonable description of the residency market. He shows that early contracting may occur despite prices that reflect supply and demand. See Wing Suen, *A Competitive Theory of Equilibrium and Disequilibrium Unravelling in Two-Sided Matching*, 31 RAND J. ECON. 101, 101–05 (2000).

118. To illustrate, Priest cites Roth’s work finding little unraveling in entry-level markets for MBA recipients, despite the fact that they return to full-time work simultaneously after completing their degrees. Priest, *supra* note 26, at 153.

119. *Id.* at 153–54.

120. *Id.*

121. *Id.* at 154–55 (observing that “some judges stand out as ‘feeders’ to the Supreme Court,” but noting that the reputation and ability of a judge are fixed at the time of application).

122. *Id.* at 155.

compete is the timing of their offers.¹²³ Judges signal the intensity of their preferences by making an offer before others make theirs, and thus “the timing of the offer becomes the currency” that clears the market.¹²⁴

Together, these arguments give reason to doubt that a computerized matching algorithm is advisable as a mechanism for equilibrating a market. First, unraveling is not necessarily problematic, as it results from individual decisions about the costs and benefits of further search, and there is no reason to presume that more opportunity to search is better. Applying this idea to the residency market, the notion that it is optimal for the residency match to occur during a student’s fourth year in medical school has no obvious economic basis. If individual programs and students jointly decide to form a match during the students’ third year, it is only because their benefits from doing so exceed their costs. Why should they be required to wait for the match? Second, any early offers in the residency market are caused by constraints on other negotiation dimensions. To the extent that programs refuse to negotiate concerning their characteristics, there is no guarantee that the market will equilibrate; timing of offers therefore becomes a basis for allocating residency positions. A prospective resident who wants an attractive position must take it early or risk its disappearance.

While Priest’s article focuses on the market for judicial clerks, Priest discusses the market for residencies as an example of a market subject to unraveling.¹²⁵ He presents Roth’s description of the early-offer and exploding-offer problems as

123. *Id.* (explaining that the judge cannot adjust a clerk’s salary, change the job description, or alter working conditions, but the judge can control the timing of the clerkship offer).

124. *Id.* at 156. See also *id.* at 153–56, for the full version of the argument outlined here; and *id.* at 182–99, for a broader presentation of the argument and an application to a variety of markets in which unraveling has been alleged to occur. In this section of his article, Priest argues that

there are two separate phenomena to which Professor Roth’s work has alerted us that explain virtually all of his examples of true unraveling. One is a market phenomenon; the second, a regulatory phenomenon. They are related because they both involve constraints in some form on the operation of market forces.

The first and, I believe the most interesting, phenomena in the Roth examples are markets in which time-of-offer is employed as a currency in the market transactions themselves. . . .

The second phenomenon . . . derives from market regulation. . . . [Specifically, it derives from efforts to evade regulations that] constrain transactions to some “optimal” time.

Id. at 183–84.

125. *Id.* at 187–90.

they existed in the 1940s, notes the adoption of the match as a solution, and shows that in 2001, the range of compensation paid to medical school graduates was much smaller than the range for law school graduates.¹²⁶ He concludes that this difference suggests an explanation for why time-of-offer becomes important for medical school graduates, but not for law school graduates.¹²⁷ Modern salary information is perhaps not the best source of evidence for a theory linking constrained terms to early offers, however. While the salary information might indicate a constraint on terms, there is no evidence of early-offer competition in the modern era, simply because the match would preclude it. However, it is possible that salaries were as compressed in the 1940s as today, and that it was this compression that led to the spate of early offers that could be eliminated only by the match. (Note that a finding that salaries were as compressed in the 1940s as they are today would tend to refute the plaintiffs' arguments that the match itself has caused salary compression. Instead, any compression would have resulted from some other type of constraint present both in the 1940s and today.)

Regardless of whether recent residency stipends offer support for Priest's theory, Priest's general insights on matching markets could apply to the residency market. Priest's paper concludes that restrictions on the terms of trade (such as salary) lead to the use of time-of-offer to clear markets, that the use of time-of-offer as a market-clearing mechanism rewards those who can recognize talent on the basis of limited information, and finally that the "introduction of restrictions on the time-of-offer currency, like any other restriction on terms of trade, will change the allocative outcomes of the market . . . but can generally be predicted to reduce aggregate welfare."¹²⁸ In the context of the residency market, this logic would suggest that constraints on negotiations over residency benefits lead to unraveling, that unraveling rewards those who make the best use of information available early in the process, and that the introduction of the timing restrictions inherent in the match changes the program-resident matching process and outcomes in such a way that aggregate welfare is reduced.

It is important to evaluate Priest's analysis because it implies that there is reason to doubt the procompetitive benefits touted by match advocates. While match supporters typically

126. *Id.*

127. *Id.* at 190.

128. *Id.* at 203–04.

claim that matching mechanisms are efficient because they reduce mismatches, Priest argues that an efficiency analysis focusing on mismatches is incomplete in three ways.¹²⁹ First, unraveling is not properly characterized as a prisoner's dilemma; it is not the case that everyone gains through a coordinated matching mechanism.¹³⁰ Second, it is important to take into account the costs and benefits of information acquisition; given the costs of information gathering, it is not necessarily the case that more is better.¹³¹ Finally, the finding of match-related welfare improvement is based on ordinal matches, and ordinal matches do not take into account differences in intensity of preferences that would ordinarily be reflected in differences in the terms of trade.¹³² Each of these three arguments offers an important insight, but as the below analysis demonstrates, none necessarily leads to the conclusion that the residency match impedes efficiency.

1. *Does the Match Solve a Prisoner's Dilemma?* Roth has characterized the early-competition problem as a prisoner's dilemma.¹³³ In the classic prisoner's dilemma, the prisoners' payoffs are structured such that making individually rational decisions (to confess) will generate a poor equilibrium outcome for both prisoners.¹³⁴ Coordination (through an agreement not to confess) will generate improved outcomes for each relative to the poor equilibrium outcome.¹³⁵ If the relevant benchmark is the benefit to the prisoners, coordination generates a more efficient result than decentralized rational decisionmaking: each is better off. If early competition is structured as a prisoner's dilemma, in which all programs and students are made better off by delaying decisionmaking, then mechanisms that delay decisionmaking will increase aggregate welfare. If the match solves a prisoner's dilemma, then it necessarily improves the welfare of match participants.

In Priest's view, however, the prisoner's dilemma label is inappropriate because some participants may not become better off through coordination.¹³⁶ Because the judges that make the

129. *Id.* at 159–61. This Article considers Priest's arguments in reverse order.

130. *Id.* at 161.

131. *Id.* at 159–61.

132. *Id.* at 159.

133. Roth, *supra* note 57, at 992.

134. THE MIT DICTIONARY OF MODERN ECONOMICS 344 (David W. Pearce ed., 4th ed. 1992).

135. *See id.*

136. Priest, *supra* note 26, at 161.

earliest offers benefit from the unraveling and would lose from coordination, the game that judges play cannot properly be characterized as a prisoner's dilemma. If some participants gain from coordination while others lose, then the only way to assess whether a matching mechanism has improved welfare is to find some way to measure the aggregate welfare of all involved, trading off one judge's gain against another's loss. The aggregate efficiency improvement from the forced delay is no longer so clear.¹³⁷

This assessment can be made more complete by differentiating between the two effects of matching mechanisms. One effect is to delay all decisions. Because more information about quality and preferences is available later, all else equal, the majority of market participants will prefer to delay their decisions. So, for example, most students and programs would likely prefer to schedule the NRMP in students' fourth year in medical school, as it is now scheduled, to scheduling the NRMP in students' second year in medical school. In this sense, coordination that delays the formation of matches does make everyone better off. The judge that makes the first offer in January of an unraveling market would prefer to make the first offer in June; the judge that makes the second offer would rather make that offer in July than in February.

But the second effect of a matching mechanism, relative to an uncoordinated market, is to alter the *way* in which matches are formed. While an early-offer judge may rather make an offer in June than in January, that judge may not want to participate in such a system if it also means that the judge can no longer make the first offer. An early-offer judge takes strategic advantage of the fact that participants are uncertain about the likelihood of other matches to improve the judge's payoff from the match. A matching mechanism forces the judge to sacrifice the strategic advantage associated with moving early, and this loss may exceed the gain from delaying the choice process. As a result, the early-offer judge may not prefer the matching mechanism outcome. Priest is thus correct to note that coordination does not necessarily benefit all participants.¹³⁸

137. For this reason, Avery and his co-authors note that the concept of Pareto efficiency (which exists when "there is no way to make one or more parties better off without making at least one person worse off") is not particularly useful in evaluating the functioning of the judicial clerk market. Avery et al., *supra* note 83, at 800. They suggest as a potential alternative method examining the "sum total of satisfaction" of clerks and judges. *See id.* at 800-04.

138. Priest, *supra* note 26, at 161 (finding that coordination does not benefit all; for example, less-favored judges or applicants might benefit from the early-offer market).

The degree to which this observation distinguishes the clerkship setting (or residency setting) from the prisoner's dilemma setting turns on the characterization of the counterfactual. It is true that individual judges prefer the result where they make offers before other judges to the result where all judges make offers at the same time.¹³⁹ But it is also true that prisoners prefer the result where they confess and their counterparts do not to the result of the coordinated game in which neither confesses.¹⁴⁰ In the simple version of the prisoner's dilemma, however, the outcome where one prisoner confesses and the other does not is not an equilibrium outcome. Because each prisoner would like to be the confessing one, the result is that both will confess, unless they both agree not to. It is for this reason that coordination improves the welfare of both. The reason that coordination does not benefit all judges is that the counterfactual is in effect a *dis*-equilibrium outcome, one in which one of the judges moves first. Each judge would like to move first, but the judges' choices are not dichotomous. The game is not one in which a judge can choose to be "first" or "last" in the same way that a prisoner can choose to "confess" or "not confess"; instead, the judge must choose an offer time along a continuous timeline. The result is likely that rather than all of the judges making offers "first," as in the prisoner's dilemma—the worst possible outcome in that the judges both sacrifice information and fail to gain an advantage over competitors—the judges' offers spread out along the timeline. The strategy that produces this spread resembles the one involved in the prisoner's dilemma, even if the outcome does not. A second result of this game is unraveling, which occurs when offer dates creep backward as judges annually update their predictions about when they must move in order to move first.

If the spread along the timeline were a product only of an annual guessing game, then the distribution of gains and losses associated with relative offer timing would be more a matter of chance than of choice. Judges would expect to gain little from the operation of such a mechanism, and lose much, due to the decreasing availability of information as timing creeps backward. They, like the confessing prisoners, would have reason to advocate for the adoption of a coordinated mechanism.

139. See *id.* (noting that judges who extend early offers maximize their individual interests, and are better off as a consequence).

140. See THE MIT DICTIONARY OF MODERN ECONOMICS, *supra* note 134, at 344 ("The incentive here, for the rational [prisoner], concerned only with his own survival, is to confess and let the others suffer the consequences.").

Judges' relative positions along the timeline may be determined by more than just chance, however. As Priest recognizes, judges' preferences over offer timing depend on a variety of factors.¹⁴¹ Potential factors that might influence offer timing include the judge's ability to predict the quality of candidates and the judge's disutility from selecting poor candidates. A judge particularly talented at selecting candidates based on little information, for example, will be more likely to make offers that fall on the early-offer end of the timeline. Because this judge need not benefit from a coordinated matching mechanism, competition among judges (or residency programs) does not necessarily take the form of a prisoner's dilemma. That some judges benefit disproportionately from early-offer competition, however, does not mean that such competition is efficiency-enhancing relative to competition within a matching system. The losses to other judges may equal or exceed the gains to these judges. The only way to evaluate the efficiency question is to investigate more carefully the nature of the early-offer competition.

2. *Do Early Offers Obtain Distributive or Productive Gains?*

After rejecting the prisoner's dilemma analogy, Priest turns to a different analogy, one comparing a judge to a prospector who buys mineral-rich land.¹⁴² The availability of an early-offer strategy encourages "judges to invest in techniques of prediction" to improve their matches, just as the possibility of profits encourages potential buyers to invest in the development of "specialized knowledge . . . to acquire a valuable object in advance of the revelation of its value to the wider set of market participants."¹⁴³ If we believe that markets are generally efficient, and that the market for land containing minerals is typical of efficient markets, then the implication of this analogy is that perhaps markets characterized by early offers are efficient too. But there are important differences between unraveling markets and markets for real estate that suggest that this analogy is problematic.

Unraveling markets are often characterized by the existence of some date before which the transaction—trade of a commodity,

141. See, e.g., Priest, *supra* note 26, at 161, 163 (suggesting that less-favored judges and less prominent judges may be willing to make early offers); *id.* at 151 (explaining that market timing will be determined in part by search costs, including "the value of time, the costs of obtaining information" and "the level of risk aversion deriving from expected uncertainty").

142. *Id.* at 162.

143. *Id.*

provision of a service, formation of a relationship—cannot begin.¹⁴⁴ In the case of residencies, this date is naturally determined by the completion of students' medical school education. Full-time residencies that build on students' full-time medical school training cannot possibly begin before medical school is complete.¹⁴⁵ As a result, the benefit produced by the match cannot be realized until the residency actually begins. Market efficiency is a function of this benefit. It is true that a market for mineral-rich real estate rewards the first person to recognize its value,¹⁴⁶ and that this reward helps to achieve market efficiency. But in this sort of market, the benefit starts accruing immediately after the "match" is made. From an efficiency perspective, it is better to identify the value of land earlier rather than later, because the owner can immediately extract the minerals, or inform the world of their presence and resell the land to someone who can. The more quickly information is discovered, the more quickly prices adjust, and the more quickly resources can be allocated to their most efficient use. Markets for law clerks and residents do not share this characteristic. Whether the match is made two years prior to the beginning of the position (on the basis of predictive powers) or one year prior to the beginning of the position (on the basis of transcript information), no benefit can accrue until the training actually begins. There is no efficiency justification for rewarding first movers solely because they are first.¹⁴⁷

This distinction has been captured in analysis of the contract law doctrine of unilateral mistake. Robert Cooter and Thomas Ulen argue, for example, that when information is "productive" and acquired by investment, a mistake defense based on the failure to acquire this information should be rejected and the contract enforced.¹⁴⁸ They cite "information that farmland contains valuable mineral resources" as an example of productive information, and note that "[e]fficiency demands giving people strong incentives to discover productive facts."¹⁴⁹ Those who

144. See *id.* at 151 (noting that many of the markets studied by Professor Roth "can be defined as 'beginning' at particular moments").

145. See *id.* (observing that "entry-level employment markets can be defined as 'beginning' only after the moment of graduation from a high school, college, or professional school").

146. See *id.* at 162 (noting that early investors are rewarded).

147. If the timing of offers and acceptances is highly correlated with the strength of market participants' partner preferences, and early offers produce matches that would not otherwise be formed, then there may be an efficiency benefit from early-offer competition. This possibility is discussed (and discounted) in subpart III.C.4 *infra*.

148. ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 282–83 (4th ed. 2004).

149. *Id.* at 281–82.

invest in discovery of minerals should be allowed to profit from their investments; those who do not invest should not be permitted to profit from the discoveries of those who do.

But Cooter and Ulen also identify a category of “redistributive information,” which “creates a bargaining advantage that can be used to redistribute wealth in favor of the informed party.”¹⁵⁰ A judge who expends resources in assessing candidate quality for the sole purpose of determining whether to make an early offer (i.e., to redistribute wealth from a later judge to him or herself) is investing in redistributive information. Cooter and Ulen point out that investment in redistributive information is wasteful, and further explain that it “induces defensive expenditures by people trying not to lose their wealth to better-informed people.”¹⁵¹ If one judge invests in making the predictions that are necessary for early offers, other judges must make similar investments so that they too can make early offers and thus avoid a loss of wealth to the first-mover judge.

It might be argued that investments in assessing quality are productive, as well as redistributive, because they allow higher-quality matches to be formed. From this perspective, investment should be encouraged. However, the information that is produced in an unraveling market is produced quite inefficiently. A program that must predict the quality of a resident two years into medical school must invest substantially more resources in that prediction than a program that predicts the quality of a resident two years later, because it must develop predictive models based on information that may be only indirectly connected with future performance.¹⁵² In contrast, a program making a later prediction can acquire more directly applicable information cheaply, because it is produced nearly without cost as a byproduct of the educational process (in the form of grades, for example). Because investment costs are likely a declining function of time, all else equal, a matching process that occurs later will be more efficient than a process that occurs earlier.¹⁵³

150. *Id.* at 282.

151. *Id.*

152. *Cf.* Tom Ginsburg & Jeffrey A. Wolf, *The Market for Elite Law Firm Associates*, 31 FLA. ST. U. L. REV. 909, 936 (2004) (arguing that “[b]ecause law firms hire when candidates have so little tangible, relevant information to provide, firms must spend more time in interviews to gather information”). Ginsburg and Wolf analyze the market for law firm associates, including its unraveling tendencies and the potential explanations for the failure to implement a centralized matching mechanism. *See id.* at 936–63.

153. The arguments about information presume that the information will be useful. One study has concluded that medical school grades and exam performance did not predict residency clinical performance. Stephen M. Borowitz, et al., *Information Collected During the Residency Match Process Does Not Predict Clinical Performance*, 154

Given that the time at which a specific piece of information is produced or collected has no bearing on its usefulness in assessing match quality, the main reason to make a costly investment in predictions rather than waiting for information to be revealed is to redistribute wealth from competing judges to oneself.

The idea that investment in acquiring soon-to-be naturally revealed information should be discouraged has also been explored in the context of contract law. Melvin Eisenberg argues that gains from acquiring “foreknowledge,” defined as “knowledge that ‘will, in due time, be evident to all,’” are often redistributive, and that the cost of acquiring such knowledge exceeds its social value.¹⁵⁴ He concludes that the law should not provide incentives for acquiring foreknowledge.¹⁵⁵ When a judge (or residency program) tries to predict qualities that will be naturally revealed as a candidate acquires more experience inside or outside of the classroom, the judge (or program) invests in foreknowledge. The acquisition of such information is likely inefficient, and the match eliminates this inefficiency by eliminating the reward for it (the pairing that would result from the early offer).

Priest acknowledges that the distributional gains associated with being the first to capture a limited resource may provide an incentive for socially excessive investment, but argues that this concern does not apply in the clerkship or resident contexts.¹⁵⁶ He explains that socially excessive investment arises when the resource’s value remains constant over time, but “the value of the resource to the judge”—presumably, the value of the clerk—changes over time.¹⁵⁷ As explained previously, however, the benefit to the judge (or residency program) of matching with a particular candidate is not a function of the timing of match formation; it is purely a function of the productivity of the

ARCHIVES PEDIATRICS & ADOLESCENT MED. 256, 256 (2000). While these results do cast some doubt on match supporters’ arguments, they do not undermine them. First, the study finds a correlation between performance and having been ranked in the top ten of the match list. *Id.* Second, the study is limited to sixty-nine residents in a single residency program, and may not be representative of the experience elsewhere. *Id.* Third, programs’ preferences (and hence rankings) may take into account factors other than clinical performance. Finally, even if programs cannot effectively make use of the additional information gained through a later match process, students may.

154. Melvin A. Eisenberg, *Disclosure in Contract Law*, 91 CAL. L. REV. 1645, 1664–65 (2003) (quoting Jack Hirshleifer, *The Private and Social Value of Information and the Reward to Inventive Activity*, 61 AM. ECON. REV. 561, 562 (1971)).

155. *Id.* at 1665.

156. Priest, *supra* note 26, at 160.

157. *Id.*

relationship once the relationship begins. Judges make early offers not to ensure that a given clerk or group of clerks will produce more than they otherwise would, but to ensure that the benefit of the clerk's fixed productivity accrues to them—a distributional gain, rather than a productive one.¹⁵⁸

While the benefit associated with hiring any given clerk or resident does not change over time, the cost of hiring the clerk may. But given that additional information about prospective clerks is naturally revealed as time progresses, the costs associated with hiring may actually decrease over time, suggesting that it is more efficient to hire clerks later than earlier.¹⁵⁹ In fact, the declining-cost nature of information in the market for clerks or residents is another factor that distinguishes these markets from the market for mineral-rich land. In the market for mineral rights, the cost of information acquisition may decrease over time because technology improves or because someone may eventually stumble across a tell-tale sign of the presence of minerals. But for the most part, there is no reason to believe that there is a particular advantage to waiting a few years to investigate the quality of land.

In his analysis of timing decisions, Priest emphasizes a different type of cost, the cost of “foregoing the opportunity of securing the clerk because of intervening rival offers.”¹⁶⁰ From an individual market participant's perspective, the cost of foregone opportunities is indeed a very real cost of delay, one that plays an

158. If the value of a given clerk does not change across time, is there anything else that does? Priest points to two possible considerations: the costs of the search process, which are discussed in the next paragraph and in subpart III.C.4, and “those costs associated with the effect of the passage of time on the expected value of those applicants remaining in the market.” *Id.* If early-offer judges correctly select the most productive candidates, they leave a less-productive pool of candidates behind. This is one of the “costs” associated with waiting to make an offer. But this cost is no different from that faced by a real estate investor who fails to identify a valuable property before a competing investor purchases it; the investor then is left with only less valuable properties to purchase. This cost stems from the limited availability of the resource in question (the most valuable properties), and is the same cost that generates socially inefficient overinvestment. The declining-quality phenomenon does not make early offers socially productive; on the contrary, it is the basis for the distributive gain that provides the incentive for early offers.

159. It is true, as Priest argues, that it is not always beneficial to acquire more information; it may be desirable, for example, to make an offer before all interviews are complete. *Id.* at 166 (arguing that canceling an interview or extending an early offer in an effort to minimize costs is not inefficient). For instance, if an interview costs \$100, and the second applicant is at most \$50 better or worse than the first applicant, it is more efficient to make an offer to either candidate immediately than to interview the second one. But when information can be obtained at almost no economic cost purely through delay, an efficient system would facilitate delay.

160. *Id.* at 157.

important role in the participant's cost-benefit analysis of when to make or to accept an offer.¹⁶¹ A rational participant would certainly take this cost into account. These opportunity costs differ from those of information acquisition, however, because they do not necessarily represent real economic costs to society.¹⁶² They do not involve the consumption of resources, and so would not be part of an assessment of the overall economic efficiency of the matching process. Their only impact on efficiency would be through an indirect route; they affect the nature of the offers and acceptances that are made, thereby affecting the quality of the resulting matches. The scarcity of information that exists when early offers are made gives reason to doubt that these offers would generate a superior match.

3. *Does Inefficiency Arise from Ordinal Ranking in the Match?* Priest's third argument, however, does suggest a way in which early-offer competition might generate a superior result to the residency match. He argues that it is inappropriate to infer efficiency based on ordinal, rather than cardinal, expressions of preferences.¹⁶³ By relying entirely on ordinal rankings of preferences, the residency match may achieve inefficient results.¹⁶⁴ Participants in the medical residency match indicate their first and second choices, but do not disclose how much more they value their first choice than their second.¹⁶⁵ Nor is there any basis for determining the extent to which one program's preference for its top-ranked candidate exceeds a second program's preference for its top-ranked candidate.¹⁶⁶

Imagine the following scenario: Program A would generate \$6 in surplus from candidate Y, and \$4 from candidate Z. Program B would generate \$5 from a match with Y, and \$1 from a match with Z. Both programs would therefore rank Y first and Z second. Imagine that both candidates would generate \$2 in surplus if matched with A, and \$1 if matched with B; both would therefore rank program A first and B second. Assume further

161. See *id.* (noting that the value of delaying an offer to a "candidate may well be less than the cost incurred by delaying the offer").

162. Priest implicitly acknowledges this point later in the article, when he argues that those who fail to get their preferred match because of early offers "may claim that a 'cost' has been imposed upon them, but it is not a cost that in a competitive economy commands normative significance." *Id.* at 162.

163. See *id.* at 159 (noting that the residency match relies on ordinal rankings that fail to account for differences in intensity of preferences); *id.* at 174 (arguing that "ordinal maximization does not guarantee the maximization of aggregate preferences").

164. *Id.* at 174.

165. See *supra* notes 46–53 and accompanying text (describing the match algorithm).

166. Priest, *supra* note 26, at 159.

that there is no way for any match participant to reallocate a portion of its surplus to another. If each program had one residency position available, the residency match algorithm would assign candidate Y to program A, and candidate Z to program B, generating \$10 in total surplus (\$6 for A, \$2 for Y, \$1 for B, \$1 for Z). But this is not the maximum surplus possible. If the program had assigned Z to program A and Y to program B, \$12 would have been created (\$4 for A, \$1 for Y, \$5 for B, \$2 for Z). Note that this result does not contradict the earlier observation that the residency match algorithm produces a stable match. Because program A would be uninterested in forming a match with Z (given that it has Y), and Y would be uninterested in forming a match with B (given that it has A), any effort to rearrange the results of the match under the given assumptions would be unsuccessful. And yet, all of the participants would be *collectively* better off—the match result would be more efficient—if the match had worked the other way around.

One way to alter this result is to permit the reallocation of surplus through a Coasian side deal.¹⁶⁷ So, for example, if program B sweetened the surplus its residents would get by offering a \$3 bonus, Y would have ranked program B first, and the algorithm would have assigned Z to A and Y to B and generated the \$12 surplus (\$4 for A, \$4 for Y, \$2 for B, \$2 for Z). Alternatively, after the match has been made, B would be willing to pay A and Y to switch partners. Paying A \$2.25 and Y \$1.25, for example, would improve B's welfare by 50 cents, Z's welfare by \$1, and the welfare of both A and Y by 25 cents each, resulting in an aggregate gain of \$2. If the match generates inefficient results, the parties have an incentive to renegotiate. But if the match is based strictly on ordinal rankings, the only way this match reallocation can occur is if the parties are allowed to express their preferences through payments of bonuses or bribes. To the extent that Priest is correct in asserting that one characteristic of markets with a tendency to unravel is

167. SAMUELSON & NORDHAUS, *supra* note 62, at 379. The Coase theorem suggests that “[w]here property rights are well defined and transaction costs are low, the affected parties can get together and negotiate an efficient solution.” *Id.*

inflexibility in the terms of trade (a questionable assertion),¹⁶⁸ then this efficiency-enhancing reassignment cannot occur.¹⁶⁹

Thus, the degree of inefficiency arising from an ordinal match depends both on the initial preferences of the participants and the extent to which reallocation of surplus is possible. If programs are all of the same quality, and students are all of the same quality, and the participants' benefits from the match depend only on quality, then any assigned pairing would be efficient. (But of course, in this case there would be no need for a match in the first place.) This is an unlikely description of the conditions surrounding the residency match. Still unlikely, but less so, is that programs differ in quality, and candidates differ in quality, but a given program produces the same benefit for all candidates, and a given candidate produces the same benefit for all programs. If the numbers of the above example were altered to reflect preferences based on these assumptions, the match algorithm should generate the efficient result without any need for reallocation of surplus. In order for the ordinal preference-based rankings to be problematic in the above example, there

168. While constraints on terms of trade may contribute to unraveling, they may not be a precondition of unraveling. See *supra* note 117. Constraints on terms will almost certainly affect market equilibria. If compensation is artificially constrained, for example, it will not reflect the scarcity of positions (or of candidates) with desirable attributes. Participants cannot make up for deficiencies in their attributes by offering a higher wage or accepting a lower one. But even if compensation is unconstrained, it may not lead to equilibrium in markets characterized by scarcity, incomplete information, and an inability to renegotiate. From a residency program's perspective, for example, only a few candidates have completed the educational prerequisites necessary to fill a position, and only a subset of this group is likely to have the attributes the program seeks. If program A could lure away a candidate at any time from program B by offering better terms, and if transaction costs barred program B from seeking compensatory damages for the candidate's departure, there would be little gain to program B from making an early match. The unraveling problem arises when that initial match is binding and decisions are made without full information about others' preferences; these characteristics help to create the strategic incentives to move early. Even if a candidate can negotiate terms with the first program that makes her an offer, the program has an incentive to make an offer early so that she cannot be enticed away by a better offer later. It is the constraints on renegotiation, then, that contribute to the unraveling problem. The plaintiffs in the NRMP suit alleged that program rules restrict movement from one program to another. Complaint, *supra* note 1, ¶ 88(b). Restrictions need not be artificial in nature, however, to result in unraveling. The costs associated with changing positions, such as moving costs or the costs of gaining familiarity with a new environment, may be enough to impede later movement. The higher these transactions costs, the more likely unraveling is to occur.

169. In general, from an efficiency standpoint it is important that mismatches, whether due to misinformation, changing preferences, or other factors, be corrected. Rules that impede movement from one residency to another, either outright or by raising the costs of movement, may therefore contribute to inefficiency. The plaintiffs in the NRMP suit alleged that "the ACGME imposes substantial obstacles to the ability of a resident physician to transfer employment from one employer to another during the period of residency." Complaint, *supra* note 1, ¶ 88(b).

must be a difference in the relative valuations programs assign the candidates, or in the relative valuations candidates assign the programs. In the above example, it is because program B's degree of preference of Y to Z is stronger than program A's degree of preference of Y to Z that it is more efficient for Y to join program B.

In the presence of initial preferences such as those in the above example, reallocation of surplus among the participating parties is necessary to achieve the efficient result. The example was posed in terms of an initial allocation of surplus, and then reallocation in the form of a bonus. In the real world, however, students do not think in terms of surpluses and bonuses, but in terms of program attributes and compensation. Students' preferences for a program are a function of the quality of the training and educational opportunities the program provides, as well as the program's geographical location, culture, and prestige.¹⁷⁰ Their preferences will also be a function of the program's immediate monetary benefits, including its stipend.¹⁷¹ Programs' preferences for students will be based on the benefits the programs anticipate the students will generate, whether these benefits take the form of providing valuable patient care services, assisting in research, or enhancing the teaching physicians' work environment. The ability of programs to recruit successfully the students they most prefer depends on the degree to which they can influence these students' choices. One way to influence these choices is to increase the students' calculations of expected "surplus" by increasing stipends. So, for example, the \$3 bonus that B offered might in reality take the form of a higher stipend. This higher compensation will increase the prospective residents' anticipated surplus, which in turn will increase their ranking of the program, which in turn will affect the match that is assigned, thereby affecting the total surplus generated by the

170. Cf. David A. Thomas, *The Law School Rankings Are Harmful Deceptions: A Response to Those Who Praise the Rankings and Suggestions for a Better Approach to Evaluating Law Schools*, 40 HOUS. L. REV. 419, 432-33 (2003) (listing location, prestige, rankings, financial incentives, and quality of education as some of the reasons prospective law students prefer certain schools over others).

171. The relative influence of training and monetary compensation on prospective residents' choices is an empirical question. Quality of training has certainly long been an important criterion. Positions in teaching hospitals (hospitals with a medical-school affiliation) were much more sought after than those in regular community hospitals, both before and after World War II. LUDMERER, *supra* note 39, at 93, 185. Ludmerer also states that community "hospitals often offered financial inducements to prospective interns, but usually to no avail, since students would choose internships on the basis of perceived educational benefits, not on the size of the stipend." *Id.* at 95. This statement lends support to the observation that residents value education and training highly, perhaps even more highly than hospitals value resident services.

final match.¹⁷² The higher stipend obviates any need for subsequent reallocation.

The implication of this analysis is that the initial example that generated the inefficient matching might never arise in the first place. It does not describe a market equilibrium. Programs will compete for the most productive candidates by adjusting the compensation that they offer. As long as programs are permitted to set their compensation, the nature of the competitive process will generate surpluses for the candidates such that the ordinal nature of the rankings will not ultimately substantially impede the efficiency of the match.¹⁷³

In a model of the matching process, Bulow and Levin have indeed shown that if programs are able to make *personalized* offers of compensation, a stable assignment of candidates to programs will achieve an efficient result.¹⁷⁴ While commentators have implied that there is some individual negotiation over residency terms,¹⁷⁵ Bulow and Levin acknowledge that personalized offers are unusual.¹⁷⁶ Instead, programs tend to offer a single compensation level to all potential residents. This means that the match outcome may not be completely efficient. Bulow and Levin's results suggest that in a market in which programs make offers that do not individually vary, there will be some inefficiency, but that the nature of the competition that occurs limits the magnitude of inefficiency.¹⁷⁷ Ultimately, the level of efficiency that an ordinal match will achieve depends on the structure of the participants' underlying preferences and the nature of the competitive process that determines the salaries

172. Sanders Chae notes that programs may refuse to increase their salaries "because they cannot be guaranteed that the students whom they really want (and whom they believe are worth a higher price) will be matched with them." Chae, *supra* note 1, at 353. It is true that program B may be nervous about increasing its salary offer by \$3; if student Y does not rank them first, they will end up being matched to student Z at a loss of \$3 rather than a gain of \$2. But the point of the compensation increase is to solicit Y's top ranking, and under the assumptions, a rational Y will rank program B first. Program B will be guaranteed to get the students it wants, if it has sufficient knowledge of the preferences of the participants in the matching process, and if it has enough money to offer them. See Bulow & Levin, *supra* note 82, for a more thorough model of matching mechanisms that takes into account the problems posed by making uniform offers to all potential residents.

173. See generally Bulow & Levin, *supra* note 82 (analyzing competitive equilibria in matching markets). Bulow and Levin model the types of strategies hospitals are likely to use in setting compensation in response to competitive pressure from other hospitals. *Id.*

174. *Id.* at 26–27.

175. See *supra* note 78.

176. Bulow & Levin, *supra* note 82, at 27 ("[P]articipants in the residency match report that personalized contingent offers are quite uncommon.").

177. See *id.* at 3, 19–22 (explaining and demonstrating that the inefficiency is minimal due to the local nature of the competition among "similar opponents").

offered.¹⁷⁸ In the context that Priest considers, the federal judicial clerk matching process, salaries are set by the federal government and do not necessarily reflect the workings of a competitive process.¹⁷⁹ This makes it much more difficult to obtain the differentiation necessary to generate an efficient match, regardless of whether there is a formal ordinal matching process. To the extent compensation in residency programs is permitted to vary, it is more likely that the residency match will generate efficient results.

4. *Summary: Factors in an Efficiency Assessment of the Match.* Ultimately, an efficiency analysis of the residency match must include a comparison of the aggregate benefits for all participants in a market characterized by early-offer competition (unraveling) to a market characterized by a formal matching mechanism. For an individual participant, the most important determinant of benefit is likely the quality of the partnership produced. The quality of the partnership produced in each setting will depend on the nature of competition in each setting.

Priest argues that when negotiation over terms is constrained, the timing of offers becomes the mechanism used to express the preferences of market participants: “In the context of the clerkship market where . . . price signals are suppressed, time-of-offer serves as a market-clearing mechanism to allow judges with higher intensities of preference for clerks to bid them away from judges whose intensities of preference are lower.”¹⁸⁰ This description is reasonable in an unraveling market where the initial match is binding.¹⁸¹ The stronger the preference of a participant for a particular partner relative to other potential

178. This conclusion is evident from Bulow and Levin’s careful theoretical evaluation of the efficiency properties of matches. In analyses not described here, they investigate the effects of a variety of assumptions. See *id.* for more details.

179. Priest, *supra* note 26, at 154–55. In Priest’s view, it is this rigidity of terms that generates the unraveling phenomenon. See *id.* at 155–58.

180. *Id.* at 174.

181. If the initial match were not binding, and if formation of a new match were costless, then unraveling would not occur. There would be no benefit to forming an early match because a partner offered a preferable match later would simply abandon the early one. This process would ensure an efficient result. Any barriers to movement, however, would impede this efficient result. This is one justification for the Department of Justice’s challenge to residency program “guidelines” that limited the ability of programs to recruit residents away from other programs. See Complaint ¶ 16, *United States v. Ass’n of Family Practice Residency Dirs.*, No. 96-575-CV-W-2 (W.D. Mo. Aug. 15, 1996), available at <http://www.usdoj.gov/atr/cases/f0700/0774.pdf> (“The Guidelines embody an agreement among the member family practice residency program directors to limit . . . competition among themselves.”); see also *supra* notes 168–69 (explaining how constraints on negotiation and movement contribute to unraveling and inefficient results).

candidates, the more quickly that participant will want to form a match to avoid the risk of losing that partner. In this way, an early offer can be a signal of preferences. If the early offer is an accurate signal, and if participants respond by forming matches based on these signals (perhaps because they benefit from matching to partners who prefer them), early offers will generate productive gains and an efficient result.

Unfortunately, early offers are quite noisy signals of preferences for match partners. First, they are the product of participant preferences at the time the match is formed, and not their preferences at the time of the residency itself. As discussed in subpart III.B, participants' partner preferences may change over time, both as their own needs change and as their information base grows. Because the benefits of the match can accrue only after the residency begins, it should be participants' preferences at this later time that matter for the calculation of aggregate social welfare.

In addition, early offers reflect not only the participants' desire for a particular match, but also the level of uncertainty that the participants face about the likelihood of other offers materializing and the participants' level of risk aversion.¹⁸² When uncertainty exists, early offers may be preferable to late offers. Individuals who are uncertain about their own qualifications, for example, cannot possibly accurately forecast how well they might fare if the matching process is delayed until more information is revealed. Rather than face a broad distribution of potential payoffs, they may prefer to establish a match early. In other words, the match provides insurance value. This insurance value complicates efforts to infer preferences about match outcomes from offer timing. The decision to make or accept an early offer is as much a signal of risk aversion as it is a signal of preference for the other party. In general, early offers and acceptances may provide some indication of market participants' partner preferences, but because they are the product of limited information and risk aversion, they are not likely to produce

182. See, e.g., Hao Li & Wing Suen, *Risk Sharing, Sorting, and Early Contracting*, 108 J. POL. ECON. 1058, 1061 (2000) (“[E]arly contracting is positively related to the degree of risk aversion of workers and to the degree of uncertainty regarding ability.”). The importance of uncertainty and insurance value for equilibria in matching markets has been explored mathematically in several papers on early contracting. See *id.* (modeling competitive early contracting as a function of risk aversion); Hao Li & Wing Suen, *Self-Fulfilling Early-Contracting Rush*, 45 INT’L ECON. REV. 301, 302 (2004) (recognizing that “[i]ncentives to sign early contracts in a competitive market can be understood in terms of the trade-off between the insurance benefits and the sorting inefficiencies generated by early contracts”); Suen, *supra* note 117, at 103–05 (demonstrating that unraveling is related to risk aversion).

matches of as high quality as those produced by matching mechanisms.

Note, however, that the insurance value associated with forming a partnership based on an early offer provides a real benefit to a risk-averse prospective resident or clerk.¹⁸³ While some risks could be completely avoided through the use of a matching mechanism—for example, the risk that a candidate will be displaced by a less-qualified candidate because of information failures—others, such as the risk that the candidate will prove to be less qualified than he or she hopes, cannot. In an assessment of total welfare, the insurance value of an early-offer system should be taken into account along with the quality of the resulting match. So too should the fact that early knowledge of a match may facilitate planning; an applicant may want to know the location of a future residency so that his or her spouse can begin searching for employment in the same location, for example. Because early-offer competition gives those who experience high disutility from uncertainty the flexibility to accept early offers, it may generate a higher level of benefit than a coordinated match for at least some participants.

Just as uncoordinated and coordinated markets may produce different levels of benefit, they may also entail different levels of cost. The differences are likely to be insubstantial, however. Costs associated with any matching process, coordinated or uncoordinated, include the costs of collecting data about potential match partners and of processing this data to determine one's rankings of potential partners. Establishing a relationship early in the process allows participants to discontinue their search, saving costs such as those associated with subsequent interviews.¹⁸⁴ But because it is important to establish a match early, participants in the early-offer competition process must prepare by evaluating potential partners in the same way that participants of a coordinated match process would. Participants must identify their preferred partners before the process begins, so that they can arrange interviews sufficiently early to preserve the possibility of an offer.¹⁸⁵ Savings from having less information

183. See, e.g., Suen, *supra* note 117, at 117 (stating that “assignment markets tend to produce a distribution of income that is more dispersed than the distribution of ability” and that this dispersion creates a demand for insurance among market participants that can be met through early contracting).

184. Chae observes that the match may reduce transaction costs by streamlining the interview and bargaining process (as does Crall, *supra* note 1, at 270), but also argues that it may escalate costs by increasing the number of interviews conducted by participants fearful of failing to match. Chae, *supra* note 1, at 354.

185. Priest notes that in an unraveling market, participants are likely to invest in

to evaluate at an early time would likely be offset by the additional processing costs involved in trying to predict match quality based on this information.¹⁸⁶

Evaluating the relative welfare effects of early-offer competition and a coordinated match requires weighing the aggregate benefits and costs produced by each. If the costs of the two competitive systems are indeed similar, the welfare comparison turns on the question of whether the benefits associated with the superior outcome of a coordinated match exceed the benefits associated with flexible timing. The current timing of the NRMP, in the spring of students' fourth year, may already reflect an implicit weighing of the gains from superior matches resulting from the later timing of the match against the gains from earlier resolution of uncertainty. If held later, the match would give residents little advance notice of their future assignments. If held earlier, preparations for the match would need to begin in students' third year, when they have just begun their clinical rotations, depriving them of experience that may be important for determining their desired career path. It is likely that the timing of the match is satisfactory for the majority of participants. Early-offer competition would allow those with a preference for earlier matches to satisfy their preference, but only at the cost of producing mismatches among those who, for strategic reasons, must make or accept offers before they otherwise would. There is no reason to expect that early-offer competition would produce higher aggregate welfare for market participants than would the match.

This welfare comparison of early-offer competition and the residency match for the most part omits discussion of one of match critics' core complaints: the potential effect of the residency match on compensation.¹⁸⁷ In the medical residency market, unlike the clerkship market, individual residency programs are at least arguably empowered to compete with one another by setting their own compensation levels.¹⁸⁸ If the

pre-market research and sorting in order to obtain successful matches. Priest, *supra* note 26, at 163.

186. On the other hand, early offer competition might conserve costs because it permits credible offers, which can serve as a low-cost source of information for other parties. See *supra* note 66 and accompanying text (explaining that potential employers who observe credible offers made by other employers can free ride off of these employers' investment in assessing the quality of a candidate).

187. See, e.g., Complaint, *supra* note 1, ¶ 83 (alleging that the "matching program has the purpose and effect of depressing, standardizing and stabilizing compensation").

188. Salaries for federal clerks are determined in accordance with a federal formula based on experience, bar membership, and locality. See Federal Law Clerk Information System, Law Clerk Employment Information, <https://lawclerks.ao.uscourts.gov/>

residency match mechanism affects not only the identities of partners matched to each other, but also wage levels, then the welfare analysis presented thus far is incomplete. It takes the identities of the market participants as given, and uses the benefits and costs accruing to these participants as the measure of welfare. But distorted wages can affect the quantity and identity of market participants, necessitating a broader welfare analysis.

D. The Role of Compensation in an Efficiency Analysis

The plaintiffs argue that the NRMP's actions in operating the match have depressed compensation.¹⁸⁹ To the extent that the plaintiffs' argument is correct, it raises two concerns. The first concern is distributional: Any mechanism that depresses compensation favors residency programs and disfavors residents. Not-for-profit teaching hospitals that operate residency programs cannot redistribute gains to owners or other affiliated individuals, so any savings resulting from lower compensation must be put to other uses.¹⁹⁰ They could be passed through to payers such as the federal government, which would then not need to pay as much to support resident training.¹⁹¹ Alternatively, they could be reallocated toward other aspects of the hospital's mission, such as conducting medical research or providing care

employinfo.htm (last visited Oct. 1, 2005) (describing criteria determining law clerk salaries).

189. Complaint, *supra* note 1, ¶ 92.

190. See THE MIT DICTIONARY OF MODERN ECONOMICS, *supra* note 134, at 308 (defining nonprofit institutions as firms that "either explicitly or implicitly do not exist to make a profit"). Hospitals that wish to maintain tax exemption under § 501(c)(3) of the Internal Revenue Code must ensure that no part of their net earnings "inures to the benefit of any private shareholder or individual." 26 U.S.C. § 501 (c)(3) (2000).

191. Mark Pauly has similarly pointed out that if an insurer acts as an agent for consumers, consumers can gain from the insurer's exercise of monopsony power in its purchase of medical services. Mark V. Pauly, *Managed Care, Market Power, and Monopsony*, 33 HEALTH SERVS. RES. 1439, 1449 (1998). In contrast, a for-profit institution exercising monopsony power is unlikely to pass on its price cut to customers. See *id.* at 1449–50 (noting that "all of the lower price paid to providers by monopsonistic insurers might not get translated into lower final product prices for insureds"). Blair and Harrison explain that if the monopsonist sells into a competitive market, it will charge the competitive price; if it sells into a market in which it has monopoly power, it will charge a higher price. ROGER D. BLAIR & JEFFREY L. HARRISON, *MONOPSONY: ANTITRUST LAW AND ECONOMICS* 39–42 (1993); see also Marius Schwartz, *Econ. Dir. of Enforcement*, U.S. Dep't of Justice, Address at the Fifth Annual Health Care Antitrust Forum at the Northwestern University School of Law: Buyer Power Concerns and the Aetna-Prudential Merger, Part II.B (Oct. 20, 1999), available at <http://www.usdoj.gov/atr/public/speeches/3924.pdf> (suggesting that monopsony power that lowers supplier prices may not ultimately benefit consumers).

for the indigent.¹⁹² Some might consider such a redistribution appropriate in light of physicians' future incomes or their status as medical professionals, particularly when residency programs' savings are redirected toward the production of public goods. Others, however, would undoubtedly argue that it is unfair to place the financial burden of supporting such functions on residents already struggling with the burden of considerable debt.¹⁹³ While this fairness concern is not ordinarily part of an antitrust analysis, Congress might choose to take it into account when evaluating the merits of an antitrust exemption.

The second concern, however, is one arguably at the core of antitrust law: the association between depressed compensation and inefficiency. It is natural to think that lower resident wages must translate to lower prices of care for patients, thus increasing their consumption of care and, all else equal, improving their welfare,¹⁹⁴ but this is not necessarily the case. To understand the relationship between low wages and efficiency in a conventional market, first consider the relationship between high prices and efficiency. Economic theory suggests that a seller of a good who operates in a perfectly competitive market will price the good at its marginal cost of production.¹⁹⁵ If the seller prices the good above cost, competitors will offer the same good at a lower price and the seller will lose its customers. When goods are priced at marginal cost, all consumers who value the good at a level higher than the cost of production will purchase it. This is an economically efficient outcome. A failure to consume a good valued higher than its cost of production would impose a loss on

192. Skeptical observers might speculate that teaching hospitals would instead allocate gains from savings toward administrators, other employees, or teaching physicians, or squander the gains through inefficiency. An examination of this issue is beyond the scope of this Article.

193. In 2003, the median amount of medical school debt for the 85% of public medical school students who graduated with debt was \$100,000. Paul Jolly, *Medical School Tuition and Young Physicians' Indebtedness*, 24 HEALTH AFF. 527, 528 (2005).

194. Chae suggests, for example, that "[i]t is also hard to believe that the Match reduces output by decreasing the amount of health care that teaching hospitals provide. Indeed, low salaries might increase output by enabling teaching hospitals to hire more residents, leading to more health care." Chae, *supra* note 1, at 354. Interestingly, such a scenario would imply that residents' stipends have not been depressed below competitive levels. Specifically, if wages are depressed, hospitals should not be able to find additional residents willing to work at the offered wages. If a hospital is able to attract additional residents of comparable quality to its current residents after lowering its salary, then its previous salary exceeded competitive levels.

195. See STEVEN E. LANDSBURG, PRICE THEORY AND APPLICATIONS 188–89 (5th ed. 2002) (explaining that a firm operating in a competitive environment will produce goods until the price of those goods equals the marginal cost of production).

society. Marginal cost pricing achieves an efficient result because it ensures that this sacrifice does not occur.¹⁹⁶

By definition, companies with monopoly power can successfully raise prices above marginal cost.¹⁹⁷ The lone producer who erects barriers to entry, or multiple producers who collude, can successfully increase prices because there are no competitors threatening to lure customers away with lower prices. When producers with monopoly power raise prices, they will sell less output relative to the perfectly competitive equilibrium.¹⁹⁸ They will end up selling their products only to the customers who value the product the most.¹⁹⁹ Customers who value the good above its marginal cost, but below the monopoly price, will refuse to purchase the good.²⁰⁰ Potential benefit to society will be sacrificed as a result. Customers whose valuation falls between marginal cost and the monopoly price will spend their money on goods they value less, resulting in allocative inefficiency. The harm to consumers from supracompetitive prices exceeds the additional profits gained by producers.²⁰¹ While some of the surplus associated with consumption of the product is transferred from consumer to producer through the higher price, the remainder simply disappears because less of the product is consumed. Monopolies are associated with supracompetitive prices, which result in suboptimal quantities consumed and therefore inefficiency.²⁰²

Reducing wages below competitive levels may similarly result in allocative inefficiency.²⁰³ A seller who can raise output

196. More specifically, it is efficient because it ensures that neither sacrifice nor waste occurs. If price were for some reason lower than cost, then consumers who value the product less than its production cost may nonetheless consume it. The result is economic waste.

197. See SAMUELSON & NORDHAUS, *supra* note 62, at 161 (“When a firm has market power in a particular market . . . the firm can raise the price of its product above its marginal cost.”).

198. Or alternatively, producers with monopoly power are able to sustain high prices by restricting the quantity of output sold. See LANDSBURG, *supra* note 195, at 344–45 (noting that monopolies restrict output to sustain higher prices and generate higher profits).

199. See *id.* at 345 exhibit 10.2 (illustrating graphically that only those consumers who value a product the most will purchase goods at monopoly price).

200. *Id.*

201. See *id.* at 344–45 (demonstrating that there is a net social welfare loss associated with monopoly pricing because the consumers’ losses exceed the producer’s gain).

202. See *id.*

203. See, e.g., BLAIR & HARRISON, *supra* note 191, at 36–39 (explaining how monopsony results in social welfare loss); Pauly, *supra* note 191, at 1445–51 (exploring the circumstances under which monopsony can result in inefficiency); Schwartz, *supra* note 191, Part II.A (describing the effects of monopsony).

prices above the competitive level is said to have monopoly power; a buyer who can depress input prices (such as wages) below the competitive level is said to have monopsony power.²⁰⁴ While in the monopoly case a higher price is associated with a lower quantity sold, in the monopsony case a lower price is associated with a lower quantity bought.²⁰⁵ More specifically, when a monopolist raises the price it charges for its output, some of the customers who would have purchased output at the competitive price will refuse to purchase output at the higher price.²⁰⁶ When a monopsonist lowers the compensation it offers to suppliers (such as workers), some of the suppliers who would otherwise have been willing to supply a good or service (such as labor) will refuse to do so.²⁰⁷ In the monopoly case, there is a relationship between the higher output price charged and the lower quantity sold; in the monopsony case, there is a relationship between the lower input price offered and the lower quantity of input bought.²⁰⁸ Monopolies result in allocative inefficiency when consumers buy less valuable products; monopsonies result in allocative inefficiency when workers divert their time to leisure or accept employment in positions in which they generate less value. Monopolies sacrifice gains from trade between producers and consumers; monopsonies sacrifice gains from trade between producers and suppliers (such as workers).²⁰⁹ Lower input prices achieved through the exercise of buyers' market power may thus decrease social welfare.

One question to be asked in light of the resident antitrust suit's compensation claims is whether teaching hospitals exercise monopsony power through the use of the match, thus decreasing social welfare. If the reasoning of match critics were based on this standard monopsony framework, it would go something like this: hospitals collude through the match system in order to keep the wages they pay low. These low wages cause inefficiency because too few residents will work; residents who were willing

204. See THE MIT DICTIONARY OF MODERN ECONOMICS, *supra* note 134, at 290–91 (defining monopoly power); see also Schwartz, *supra* note 191, Part II.A (defining monopsony power).

205. See Schwartz, *supra* note 191, Part II.A (explaining that the “monopolist raises prices above the competitive level by restricting the quantity it sells” and that the “monopsonist depresses the input price it pays below the competitive level by reducing the input quantity it purchases”).

206. See *id.* (explaining that “at higher prices less is demanded”).

207. Alternatively, the monopsonist is able to sustain low input prices by restricting the quantity of input purchased. Schwartz, *supra* note 191, Part II.A.

208. *Id.*

209. See BLAIR & HARRISON, *supra* note 191, at 39 (noting that there are “unrealized gains from further trade” between firms and workers in monopsonies).

to work for the competitive wage now direct their efforts elsewhere. If the match is associated with lower wages and the hiring of fewer residents, the match likely decreases social welfare by lowering total output.²¹⁰

Match proponents would likely object to this characterization of the match. Because the match dictates only the mechanism for assigning residents to programs, and not wage or employment levels, its rules do not directly facilitate the creation of a collusive monopsony.²¹¹ The match itself does not provide an enforcement mechanism that limits the number of residents that teaching hospitals can seek, either individually or collectively.²¹² An individual program might decide unilaterally to

210. The consumer welfare effects of a change of resident output depend critically on the nature of the market for the services provided by residents. For example, if competing nonteaching hospitals offer the very same services through nurse practitioners, nurses, technicians, or nonresident physicians, and if they operate on a flat long-run marginal cost curve, they can simply expand the services they provide by the same amount that the teaching hospitals have decreased the services provided by residents. If total output is unaffected, consumers remain unharmed by teaching hospitals' actions. Since the "main concern of the antitrust laws" is the consumer, the possibility that consumers might be unaffected by the exercise of monopsony power raises the question of whether such exercise should be viewed as an antitrust violation at all. HERBERT HOVENKAMP, 12 ANTITRUST LAW ¶ 2011b1 (1999). If there is no output reduction, is there any antitrust concern? Hovenkamp considers an example of bid rigging in the context of an auction. *Id.* ¶ 2011c. When all goods are sold, there is no reduction in output, just "mere wealth transfer" from the seller to the buyer "that the antitrust laws were not designed to remedy." *Id.* Nonetheless, Hovenkamp suggests that the bid rigging should be a per se violation of antitrust laws in part because of the long-run reduction in sellers' incentive to sell. *Id.* The Supreme Court has held that price-fixing by purchasers violates the Sherman Act "even though . . . the persons specially injured under the treble damage claim are sellers, not customers or consumers." *Mandeville Island Farms, Inc. v. Am. Crystal Sugar Co.*, 334 U.S. 219, 235 (1948). Seizing on this language, the Tenth Circuit rejected a defendant's argument that a monopsony should not be actionable unless it injures consumers. *Telecom Commc'ns, Inc. v. Southwestern Bell Tel. Co.*, 305 F.3d 1124, 1133-36 (10th Cir. 2002) (holding that suppliers can claim the protection of antitrust laws without a showing of harm to consumers and dismissing language in two district court opinions that implied otherwise).

211. See *supra* Part II (discussing the general rules and nature of the matching process).

212. Other sorts of enforcement mechanisms may exist, however. In the years before the match, medical educators limited the number of specialists by limiting the number of residency positions available. LUDMERER, *supra* note 39, at 99. Today, the ACGME's Residency Review Committees may limit the number of residency positions available. ACGME, Common Program Requirements § IV.B, available at http://www.acgme.org/acWebsite/dutyHours/dh_dutyHoursCommonPR.pdf (last visited Oct. 1, 2005) ("The RRC will approve the number of residents based upon established written criteria that include the adequacy of resources for resident education (e.g., the quality and volume of patients and related clinical material available for education), faculty-resident ratio, institutional funding, and the quality of faculty teaching."). Teaching hospitals deemed to have too many residents have lost their accreditation. See, e.g., Charles Ornstein, *King/Drew Medical Center States Its Case in Plea to Accrediting Council*, L.A. TIMES, Sept. 25, 2003, at B3 (reporting that a residency program lost its accreditation due to training more fifth-

limit the number of residents it hires, but other programs may then respond by hiring more residents. Similarly, the match rules are completely silent on the issue of compensation, and a unilateral effort to decrease the compensation offered is likely to be unsuccessful.

Express restrictions on wages, however, are not the only means of limiting compensation. The match-imposed limitations on bargaining for higher salaries of the sort discussed in subpart III.A, for example, may depress compensation without any express agreement. The technical theoretical models of Kamecke and Levin and Bulow described in subpart III.A also demonstrate that under certain assumptions, the very structure of the match can depress compensation;²¹³ neither model depends on coordinated limits on compensation or hiring to achieve its results.²¹⁴ If compensation drops below competitive levels, prospective residents may seek other positions, resulting in inefficiency and the loss of social welfare.²¹⁵

year residents than permitted). The antitrust suit challenged such limits. Complaint, *supra* note 1, ¶ 88(a). But there is a procompetitive justification for such limits. In an environment where it is difficult to assess the quality of care, patients may prefer to rely on organizations such as the ACGME to ensure that physicians have received adequate training. By enforcing resource-based limits on the number of residents, the ACGME may be able to assure training quality more easily, thus promoting the sale of resident services. See *infra* note 295 and accompanying text for a discussion of the effects of limiting residency positions.

213. See *supra* notes 80, 82 and accompanying text.

214. See *supra* notes 80, 82 and accompanying text (describing the models).

215. This statement assumes that the more resident services provided, the higher consumer welfare. If physicians lacking patients can induce their own demand by taking advantage of patients' lack of information about the appropriateness of treatment, however, then social welfare might be improved by having fewer physicians. See Richard A. Cooper & Linda H. Aiken, *Human Inputs: The Health Care Workforce and Medical Markets*, 26 J. HEALTH POL. POL'Y & L. 925, 933–34 (2001) (considering the proposition that “supplier-induced demand implies the provision of excessive or inappropriate care” but concluding that “physician-induced demand is of small magnitude”). Peter Hammer has suggested that “antitrust law should recognize a defense for private acts that restrain ‘competition’ under the traditional antitrust analysis but advance total welfare.” Peter J. Hammer, *Antitrust Beyond Competition: Market Failures, Total Welfare, and the Challenge of Intramarket Second-Best Tradeoffs*, 98 MICH. L. REV. 849, 851 (2000). Those who believe that the match reduces the quantity of physician services but also believe that there are “too many” physicians could characterize the match as a private act that advances total welfare. It is not clear, however, that match-induced downward wage pressure would satisfy all of Hammer's criteria for the exception. See *id.* (advancing four requirements for a welfare defense). In addition, the debate over whether we have too many or too few physicians largely revolves around efforts to predict consumer demand; predictions of “too many” physicians have recently been displaced by predictions of “too few” physicians. See Cooper & Aiken, *supra*, at 933 (stating that “shortages are developing in some of the specialties”). See generally David Blumenthal, *New Steam from an Old Cauldron—The Physician-Supply Debate*, 350 NEW ENG. J. MED. 1780 (2004) (“Just yesterday, it seems, the conventional wisdom was a confident prediction that we faced a worrisome surplus of physicians. But today, a swelling chorus of experts contends

The extent to which inefficiency might arise in the residency market depends on a variety of institutional features. It is possible that even with lower levels of compensation, the total number of residents working would remain the same. There is little evidence that substantial numbers of American-trained medical graduates go unhired due to residency program limitations; each year the vast majority of U.S. seniors are successfully matched through the NRMP.²¹⁶ The problematic monopsony effects outlined previously assume that total resident supply is responsive to the compensation offered, but it is possible that resident supply is fixed over the relevant wage range. If the resident supply is unresponsive, the programs would still be able to attract a full complement of residents, despite the lower wages, and there would be no immediate efficiency loss.

Even if in the short term all medical graduates are hired, there may be a concern about the long-term effects of lower resident compensation.²¹⁷ Lower resident compensation may translate into lower lifetime income for physicians, decreasing the incentive to enter medical school.²¹⁸ A substantial decrease in wages would be required, however, for lower medical school application rates to translate into lower numbers of U.S.-educated residents. Over the last twenty years, the number of applicants to U.S. medical schools has fluctuated, from 35,720 in

that we may face an equally worrisome shortage of physicians.”).

216. See *supra* note 54 and accompanying text (discussing placement in residency programs). While most seniors in American medical schools find residency positions, many international medical graduates do not. NAT'L RESIDENT MATCHING PROGRAM, RESULTS AND DATA: 2003 MATCH 5 (2003). The total number of residency position applicants, including international medical graduates, exceeds the number of positions listed in the match. *Id.* at 6–7. For example, the ratio of positions per U.S. senior was approximately 1.5 in 2003, while the ratio of positions per applicant, including international medical graduates, was 0.87. *Id.* at 7. The addition of international medical graduates to the resident market will tend to shift the resident supply curve outward, decreasing compensation for existing residents, but improving social welfare. See THOMAS A. PUGEL, INTERNATIONAL ECONOMICS 361–63 (12th ed. 2004) (showing graphically that addition of migrants to a labor market decreases equilibrium wages but produces gains for the market's employers that exceed the losses to the market's existing employees).

217. See BLAIR & HARRISON, *supra* note 191, at 72 (explaining that although a collusive monopsony in a market with a fixed supply of wheat has only distributive consequences in the short run, “[a]s producer profits are reduced by the collusion, their incentives to plant durum wheat are reduced and they may reduce supply in the future [which has] adverse consequences for consumer welfare in the future”).

218. Lower resident compensation translates into lower lifetime income for physicians, assuming that there is no corresponding increase in income during post-residency years. For an analysis suggesting that actions that decrease the number of residents may increase long-term physician income, see *infra* note 295.

1982-83 to 46,965 in 1996-97, and back to 33,625 in 2002-03.²¹⁹ But the number of accepted applicants has not varied nearly as much. The number of applicants accepted was 17,294 in 1982-83, 17,385 in 1996-97, and 17,592 in 2002-03.²²⁰ The ratios of applicants to accepted applicants were therefore 2.1:1.0, 2.7:1.0, and 1.9:1.0, respectively.²²¹ Far more people desire to become physicians than actually become physicians. Lower stipends might discourage some students from entering the profession, but the high ratio of applications to acceptances indicates that others would be willing to take their places. Lower stipends are unlikely to have a significant dynamic effect on the quantity of U.S. residents.²²²

They can, however, have a dynamic effect on the *quality* of U.S. residents. Talented and productive college graduates facing lucrative alternative career paths may choose to pursue those opportunities rather than enrolling in medical school because the bargaining power of the residency programs depresses wages just enough to make a difference. These potential applicants' medical school positions, and later, residency positions, may be taken over by lower-quality applicants.²²³ If medical schools apply an absolute quality screen to applicants so that they refuse to admit applicants of lower quality, lower wages may be associated with fewer U.S. residents; if medical schools apply a relative quality screen to applicants so that they have a set number of positions available and will fill them with lower-quality applicants if necessary, lower wages will be associated with lower-quality residents.²²⁴ The degree to which the quality of residents would decline depends on the extent to which the students who fill the

219. Barbara Barzansky & Sylvia I. Etzel, *Educational Programs in US Medical Schools, 2002-2003*, 290 JAMA 1190, 1192 tbl.3 (2003).

220. *Id.*

221. *Id.*

222. This may change if the number of medical school positions increases. In 2005, the AAMC recommended a 15% increase in medical school enrollment by 2015. Press Release, AAMC, AAMC Calls for Modest Increase in Medical School Enrollment (Feb. 22, 2005), available at <http://www.aamc.org/newsroom/pressrel/2005/050222.htm>.

223. *Cf.* Pauly, *supra* note 191, at 1451 (noting that in health insurance, monopsony may be associated with "lower quality, rather than . . . fewer people insured").

224. In 1989, when the ratio of first-time applicants to medical school places fell to 1.2:1.0, some medical schools declined to fill their classes. Cooper & Aiken, *supra* note 215, at 931. It appears, then, that in 1989 medical schools perceived the quality of rejected applicants to be lower than that of their accepted counterparts. *See id.* (reasoning that "as a means of maintaining quality" the schools did not fill their classes). On the other hand, the AAMC has recently advocated an expansion of medical school enrollment, and the AAMC's president argues that this expansion would not decrease quality. *See* Myrle Croasdale, *Physician Shortage?: Push Is on for More Med Students*, AM. MED. NEWS, Mar. 14, 2005, at 1.

vacated medical school slots can, with training, become residents of comparable quality to the students they replace.

The sensitivity of prospective residents—or prospective medical school applicants—to resident compensation levels is ultimately an empirical question. While reductions in potential lifetime income will affect the quantity or quality of resident supply on the margin, institutional features of the graduate medical education process are likely to limit the magnitude of the effects. The role of the residency position as the sole gateway to a future career in medicine will tend to limit the responsiveness of the resident supply to resident compensation. State licensure laws generally require that physicians receive training in a residency program as a precondition to practicing independently.²²⁵ Students who decide not to pursue a residency therefore also forgo any income that would be derived from professional practice. Because for many students the average income of medical practice is likely to exceed that of any other career path, they have a financial incentive to join a residency program, even if the compensation during residency years is low.²²⁶ Because the residency market is linked to the market for physician services, and state legislatures impose a barrier to entry in the market for physician services, students may join residency programs even if the initial compensation is not as high as it could otherwise be. The higher compensation earned by physicians insulated by state licensure laws against competition tends to offset the negative incentive effects that would otherwise result from lower resident compensation.²²⁷

While the discussion so far has focused on the relationship between potential residents' demand for residency positions and efficiency, Kamecke focuses on the flip side of the analysis: the relationship between teaching hospitals' demand for residents and efficiency.²²⁸ In the Kamecke model, inefficiency arises from

225. See, e.g., N.Y. State Educ. Dep't, *supra* note 37 (listing the requirements for becoming a licensed physician in the state of New York).

226. One author suggests that the net present value of an investment in medical school education is over one million dollars. Jolly, *supra* note 193, at 534.

227. This argument assumes that potential future physicians consider lifetime income, rather than just the income during residency years, when deciding whether to join the profession. If students expect their financial needs to exceed their income during their residency years and also expect to have insufficient access to credit, they may choose another career. See *id.* at 531 (reporting survey results finding that the cost of medical school was a significant deterrent to prospective medical school applicants). Educators have long been concerned that low resident compensation might deter students from pursuing training; such concerns were articulated even before the match began. LUDMERER, *supra* note 39, at 424 n.93.

228. Kamecke, *supra* note 80, at 33.

the possibility that the low compensation levels assured by competitive restraints associated with the match will lead hospitals to employ inefficient numbers of residents relative to their substitutes (such as nonresident physicians, nurses, or technicians).²²⁹ In other words, the match introduces distortions of wages that affect programs' decisions about whether to hire residents in the first place. These distortions can lead to inefficient levels of hiring.

The Bulow-Levin model focuses on yet another compensation-related source of inefficiency: mismatches. Even if the residency match does not affect the total number of residents hired, it can reduce welfare if it assigns candidates to the "wrong" positions, ones in which they do not achieve their maximum productivity. In the Bulow-Levin model, when the programs offer a uniform compensation rate to all of their residents, they choose a compensation strategy that is rational but nonetheless results in some mismatches.²³⁰ In particular, the equilibrium compensation strategy of each participating program will be "mixed" in the sense that the program will randomize compensation offers over a particular range.²³¹ Sometimes (i.e., with some probability), the chosen compensation will be sufficiently high for the program to beat out its competitors in the race to attract the most productive match, and the efficient match will therefore result; other times, however, the chosen compensation will be too low, and the program's desired resident will go to a competitor—an inefficient result.²³²

Similarly, a resident not able to use a firm offer to negotiate a higher salary because of residency match constraints on pre-match contracting might argue that this prohibition results in inefficient matches. More specifically, the firm offer helps to convey information to the market about the value of a particular resident.²³³ Without this information, a residency program may underestimate the value of an individual applicant, while at the same time perhaps overestimating the value of others. As a result, the program will offer a low stipend, and the applicant will choose another program that may ultimately generate a

229. *Id.* at 47–48.

230. Bulow & Levin, *supra* note 82, at 18–19 ("Because there is mixing, firms and workers may not be efficiently matched.").

231. *Id.* at 6–11.

232. *See generally id.* (computing the probabilities of efficient results when offers are randomized).

233. *See supra* Part III.A.

lower value from that resident. Again, the result is an inefficient mismatch.

To summarize, the reason that low compensation is problematic from an efficiency perspective is that it results in an equilibrium where potential residents do not take the positions in which they would be the most productive. Low compensation levels send the wrong signals. A prospective medical student might choose another profession; a teaching hospital, an inefficient mix of employees; a prospective resident, the wrong residency program. In each case, resources are wasted.

The magnitude of this waste, however, is likely to be limited. As previously discussed, the nature of the market for physician services (for example, the role of resident training as a prerequisite for physician licensure) will tend to constrain the magnitude of compensation-related distortions in the resident market. In addition, while Bulow and Levin do find some inefficiencies in the presence of a match, the magnitude of the distortion is small.²³⁴ The reason the distortion is small is that the compensation range over which each program randomizes overlaps with that of its closest competitors.²³⁵ The overlapping compensation ranges of programs of similar quality mean that when a resident is assigned to the “wrong” program from the standpoint of efficiency (because the resident ranked more highly the program that happened to offer the higher compensation), the “wrong” program is close in quality to the “right” program. The localized nature of the compensation competition ensures that the inefficiency that arises from the mismatches will be small.²³⁶ A numerical calculation suggests that the more firms and workers involved in a matching process, the lower the level of inefficiency.²³⁷ With 1,000 firms and 1,000 workers participating, for example, the calculation shows only a 0.04% loss in efficiency.²³⁸

Moreover, this decrease in efficiency is measured relative to the competitive ideal—the perfect match. Because of the lack of

234. See Bulow & Levin, *supra* note 82, at 3 (finding that “the inefficiency in the market will be small”).

235. See *id.* at 6–10, 19 (explaining the small loss of efficiency associated with randomizing offers).

236. *Id.* at 3, 19 (suggesting that inefficiency is limited when “firms compete ‘locally’ against similar opponents”).

237. See *id.* at 22 (showing a decrease in inefficiency as the number of firms participating in a matching program increases).

238. See *id.* The 0.04% loss in efficiency is calculated by dividing the inefficiency associated with 1000 firms by the surplus realized by generating an efficient match. *Id.* tbl.2.

information that characterizes unraveling markets, there is no reason to expect that a residency market without a matching mechanism will achieve the perfect match.²³⁹ This observation is also relevant to an assessment of the concerns raised by Kamecke's article. Kamecke himself acknowledges the importance of comparing the results of the matching mechanism to the appropriate counterfactual:

One has to keep in mind that the centralized procedure could solve the 'unraveling' phenomenon described by Roth . . . which can lead to serious distortions inside the matching market. As long as one cannot control these problems in a decentralized system, it is not useful to abolish a workable centralized mechanism just because some distortions may remain.²⁴⁰

Thus, gains from implementing a centralized matching mechanism may exceed losses associated with any match-related compensation effects.

Subpart III.B has already demonstrated how unraveling markets generate subpar matches that can reduce benefits to programs, residents, and the public (if a poor match reduces the quantity of medical research, for example). Unraveling markets may also give rise to monopsony power issues of their own. One recognized source of monopsony power is ignorance of the supplier of the purchased goods or services about other potential buyers.²⁴¹ If early offers or exploding offers do not give participants sufficient time to solicit alternative offers, teaching programs will have considerable market power with respect to individual applicants.²⁴² While it would be ideal for residents to search for positions and to use an offer by one firm to increase the stipend offered by another, a resident with an exploding offer may not have the opportunity to do so.²⁴³ An empirical study of the gastroenterologist fellowship market further suggests that the scope of labor markets is smaller in decentralized matching

239. See, e.g., Roth & Xing, *supra* note 91, at 1034–35 (“Unraveling may be ex ante as well as ex post inefficient . . .”).

240. Kamecke, *supra* note 80, at 48.

241. See ALAN MANNING, MONOPSONY IN MOTION: IMPERFECT COMPETITION IN LABOR MARKETS 360 (2003) (describing “ignorance among workers about labor market opportunities” as a source of friction that gives employers monopsony power).

242. See *supra* text accompanying notes 97–101 (discussing exploding offers to residents and law clerks).

243. See *supra* text accompanying note 100 (reasoning that a candidate with an exploding offer may not want to allow it to explode, given the candidate's uncertainty about future offers).

markets than in centralized matching markets.²⁴⁴ The fewer employers effectively participating in the market, the more market power they are likely to have in setting wages. Abandoning the match may limit some forms of market power exercised by residency training programs, but it may allow for the introduction of others.

The possibility that the match may under some circumstances depress wages relative to the competitive ideal thus tells us little about its social desirability. The compensation-related effects of the match depend on multiple variables, including the sensitivity of residents to compensation, and any effects are likely to be outweighed by the benefits to residents, programs, and society generated through improved matching.

IV. THE EMPIRICAL EVIDENCE ON RESIDENT COMPENSATION

One way to investigate the net effects of the match on resident compensation is to examine the empirical evidence. In their complaint, the plaintiffs cite data that first-year residents earned a salary of about \$35,700 during the 2000-01 employment year in support of their claim that anticompetitive restraints, including the match, have depressed compensation.²⁴⁵ The mean first-year resident compensation for an institution offering residency programs during the academic year 2003-04 was \$39,809.²⁴⁶ To determine whether this compensation is below

244. See generally Muriel Niederle & Alvin E. Roth, *Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and Without a Centralized Match*, 111 J. POL. ECON. 1342 (2003) (comparing the entry-level market for gastroenterologists in the years in which it was centralized with the same market in the years prior and subsequent to centralization). When no centralized matching mechanism was in operation, fellows tended to be matched more often at the same hospitals and in the same cities where they worked as residents. *Id.* at 1343-44. This phenomenon could result from the need to rely on informal networks to share information about market participants. *Id.* at 1344.

245. Complaint, *supra* note 1, ¶ 93.

246. ASS'N OF AM. MED. COLLS., *supra* note 18, at 6. In real terms, this mean salary represents an increase of 0.8% over the 2002-03 year, a 2.2% increase over the 2000-01 year, a 6.4% increase over the 1990-91 year, and an 8.2% increase over the 1980-81 year, but only a 1.2% increase over the 1970-71 year. *Id.* In their complaint, the plaintiffs cite stability of wages as one of the effects of the allegedly anticompetitive restraints. Complaint, *supra* note 1, ¶ 93. It is difficult to determine what the appropriate benchmark should be, given that supply and demand conditions can vary tremendously across markets. Statistics for similar markets, however, suggest that wage growth might indeed be low. Hospitals' employment costs overall for civilian workers increased 9% between March of 1990 and 2003; the similar figure for private employers of professional specialty and technical occupations was 11%. BUREAU OF LABOR STATISTICS, U.S. DEPT OF LABOR, EMPLOYMENT COST INDEX, CONSTANT DOLLAR, HISTORICAL LISTING 22 tbl.3, 69 tbl.5a (Apr. 29, 2005), available at <http://www.bls.gov/web/econst.pdf>.

competitive levels, it is important to first identify the appropriate benchmark.

One possible comparison is to practicing nonresident physicians; residents make significantly less. One survey found that the average net income from medical practice in 1999 was \$187,000, for example.²⁴⁷ But residents, particularly first-year residents, will not be as productive as more experienced physicians,²⁴⁸ so it is not surprising that their compensation is significantly lower. Compensation does increase with the number of years in the program: The resident-weighted mean stipend in 2004-05 was \$40,552 for a first-year resident, \$44,122 for a third-year resident, and \$47,783 for a fifth-year resident.²⁴⁹ On the other hand, even the highest amount does not approach the mean earnings for nonresident physicians.²⁵⁰ This discrepancy is particularly glaring since the residency programs that require five years of training tend to be for the more highly compensated specialties, such as surgery; the average compensation after expenses for surgeons in 2002 was approximately \$255,000.²⁵¹

Resident compensation might also be compared to that of other health care providers. The plaintiffs allege that “[r]esidents generally earn less, on both an annualized and hourly basis, than other hospital employees such as nurse practitioners and physician assistants.”²⁵² In 2002, the median annual income for

247. MARIE REED & PAUL B. GINSBURG, CENTER FOR STUDYING HEALTH SYSTEM CHANGE, DATA BULL. NO. 24, BEHIND THE TIMES: PHYSICIAN INCOME, 1995-99 (2003), available at <http://www.hschange.org/CONTENT/544>.

248. Cf. Robert S. Huckman & Jason R. Barro, *Cohort Turnover and Productivity: The July Phenomenon in Teaching Hospitals* 30-31 (Nat'l Bureau of Econ. Research, Working Paper No. 11182, 2005), available at <http://papers.nber.org/papers/w11182.pdf> (finding that the annual turnover of residents at a hospital reduces hospital productivity).

249. ASS'N OF AM. MED. COLLS., *supra* note 18, at 9. The relatively low rate of increase of stipends over program years and the sudden jump in physician salaries after residency completion suggest that stipend levels reflect factors other than resident productivity. One possibility is that given the transaction costs associated with moving between programs, residents and programs both view the residency program as a multi-year relationship, with competition occurring at the point of entry into the program. See *supra* notes 168-69 (discussing impediments to moving between programs). If so, programs may choose to smooth compensation over the life of a program by paying residents high stipends (relative to productivity) at the beginning of the program, and lower stipends toward the end. Other explanations of stipend levels are considered *infra*, notes 256-76 and accompanying text.

250. See REED & GINSBURG, *supra* note 247 (reporting that the average net income for all patient care physicians in 1999 was \$187,000).

251. According to the Medical Group Management Association, the median total compensation for family practice physicians in 2002 was \$150,267; for internists, it was \$155,530; and for surgeons, it was \$255,438. BUREAU OF LABOR STATISTICS, U.S. DEP'T OF LABOR, OCCUPATIONAL OUTLOOK HANDBOOK (2004-05), available at <http://www.bls.gov/oco/pdf/ocos074.pdf>.

252. Complaint, *supra* note 1, ¶ 93.

physician assistants employed in hospitals was \$65,910; median income for first-year physician assistants in all settings was \$64,670.²⁵³ The annual compensation of residents is therefore less than the income of not only physicians who have completed their training, but also other professionals who have fewer years of formal training and who share some of the residents' duties.²⁵⁴ A hospital that must substitute other workers for residents is likely to find that its salary expenses increase. Some have estimated that the ACGME's recently-imposed limits on resident work hours will result in increased hiring of nurse practitioners and others at a cost that could reach into the millions of dollars for large academic medical centers.²⁵⁵

Comparisons of monetary compensation alone are misleading, however, because in addition to their stipends, residents receive compensation in the form of training.²⁵⁶ While the plaintiffs characterize the market being restrained as a market for resident services, the market in question could just as easily have been characterized as a market for resident education.²⁵⁷ Residency programs generally involve formal

253. BUREAU OF LABOR STATISTICS, *supra* note 251. For more detailed descriptions of 1997 earnings data, see American Academy of Physician Assistants, PA Incomes: Results from the 1997 AAPA Physician Assistant Census, available at <http://www.aapa.org/research/97income.html> (last visited Oct. 1, 2005).

254. While the market for resident services overlaps with the market for the services of physician assistants and other hospital employees, the markets are not coextensive. Teaching hospitals compete mainly among themselves for the services of residents, while they must compete with community hospitals, physician practices, and other potential employers for the services of other types of health care workers. The differences in demand for these types of labor may contribute to compensation differentials. In particular, some hospitals may prefer not to hire residents because of the training costs that would be entailed.

255. See Katherine Vogt, *Hospitals Count Up Cost of Reduced Resident Hours*, AM. MED. NEWS, Aug. 11, 2003, at 19, 19.

256. Others have noted that the exchange of educational services for patient services complicates the assessment of compensation and the antitrust case. See, e.g., Peter J. Hammer & William M. Sage, *Critical Issues in Hospital Antitrust Law*, 22 HEALTH AFF. 88, 93 (2003) (noting that a "twist" to the NRMP litigation is that "two markets are involved simultaneously: teaching hospitals acting as buyers of residents' clinical services, and residents acting as buyers of graduate medical education"); Miller & Greaney, *supra* note 1, at 915 (recognizing that "medical residents are buyers of educational services from seller residency programs" as well as sellers of labor to the buyer hospitals). The training and service aspects of residency positions and the tensions between them have long been recognized. See, e.g., LUDMERER, *supra* note 39, at 92 (describing the many duties of residents and stating that "[a]chieving the proper balance between education and service would perplex medical educators throughout the twentieth century").

257. This characterization may reflect a strategic decision by the plaintiffs. Courts have previously recognized that monopsonistic practices in employment are subject to antitrust laws. See, e.g., HOVENKAMP, *supra* note 210, ¶ 2012c ("[A] naked agreement among employers limiting salaries or wages is unlawful per se."). The Third Circuit has stated, however, that "[i]t may be that institutions of higher education 'require that a

training sessions and direct supervision by faculty physicians.²⁵⁸ When they provide their services, residents are also gaining experience. More advanced residents may have limited interaction with faculty, but still practice in a teaching hospital environment, which will generally expose them to a broader range of learning opportunities (such as more difficult or unusual cases) than is available to most physicians. The compensation that residents receive for their efforts therefore includes not just a stipend and other employment benefits, but also the valuable training that will serve as a foundation upon which to build their careers. The fact that training can be very valuable is evidenced by physicians who pursue fellowship opportunities after their residencies, sacrificing the income of independent practice in favor of stipends averaging in the \$40,000 range.²⁵⁹

One way to incorporate the value of training into an assessment of resident compensation is to reconceptualize the meager residency stipend as a much higher stipend that has been reduced implicitly by tuition. Undergraduate students typically do not produce significant value for their institutions through their labor and so must pay substantial tuition for the education they receive. Graduate students typically do produce value for their institutions, through research or teaching assistantships.²⁶⁰ Because this value offsets the tuition they would otherwise pay, graduate students often pay only a small amount of tuition, or pay no tuition at all and receive a small stipend.²⁶¹ Residents produce more value for their institutions than most graduate

particular practice, which could properly be viewed as a violation of the Sherman Act in another context, be treated differently.” *United States v. Brown Univ.*, 5 F.3d 658, 678 (3d Cir. 1993) (quoting *Goldfarb v. Virginia*, 421 U.S. 773, 788 n.17 (1975)). For a discussion of the potential applicability of the court’s analysis in *Brown University* to the residency match case, see Miller & Greaney, *supra* note 1, at 917–18, and Chae, *supra* note 1, at 353–54.

258. AM. MED. ASS’N, *supra* note 32.

259. See Muriel Niederle & Alvin E. Roth, Letter to the Editor, *Relationship Between Wages and Presence of a Match in Medical Fellowships*, 290 JAMA 1153, 1153 (2003) (documenting fellowship compensation in various subspecialties). For further evidence of the value of training, see *supra* note 171 and accompanying text (describing the relative importance of training and compensation in the choice of residency positions).

260. See, e.g., Univ. of Houston, Graduate Catalog Online, University Policies for Graduate Student Assistantships, http://www.uh.edu/grad_catalog/grad_assistantship/uh_policy.html (last visited Oct. 1, 2005) (listing five graduate student appointment categories: “teaching fellow, teaching assistant, instructional assistant, research assistant, and graduate assistant”).

261. See, e.g., Univ. of Houston, Graduate Catalog Online, Graduate Assistant Tuition Fellowship Policy for Students in TA/TF/IA/RA/GA Appointment Categories, http://www.uh.edu/grad_catalog/grad_assistantship/gatf.html (last visited Oct. 1, 2005) (allowing for nine resident credit hours of support per regular semester to be awarded to eligible graduate students).

students, and receive a higher stipend than graduate students.²⁶² But where tuition may be explicit for some graduate students, it is implicit for medical residents: They pay it in the form of lower stipends than they might otherwise receive.²⁶³ Residents, like other students, are willing to pay this tuition because of the long-term benefits their training provides. In fact, one study has suggested that some residents would be willing to pay teaching hospitals for residency positions in dermatology, general surgery, orthopedic surgery, and radiology because physicians working in these areas receive high financial returns from their training.²⁶⁴ Thus, while resident compensation may appear low compared to that of experienced physicians and other health care professionals, it is difficult to know whether this data demonstrates an anticompetitive effect because these comparison groups are not receiving comparable training.²⁶⁵ In other words,

262. See ASS'N OF AM. MED. COLLS., *supra* note 18, at 3, 5 tbl.1 (reporting that the average stipend of first-year residents in 2004-05 was \$40,788).

263. MEDICARE PAYMENT ADVISORY COMM'N, RETHINKING MEDICARE'S PAYMENT POLICIES FOR GRADUATE MEDICAL EDUCATION AND TEACHING HOSPITALS 7 (1999).

264. Sean Nicholson, *Barriers to Entering Medical Specialties* (Nat'l Bureau of Econ. Research, Working Paper No. 9649, 2003), available at <http://www.nber.org/papers/w9649>.

265. Some analysts might protest that most employment has a training component to it, and yet significant salaries are still paid. Recent law school graduates, for example, often earn significantly more than recent medical school graduates, despite also receiving on-the-job training. See Priest, *supra* note 26, at 188 tbl.7 (reporting, for example, that the seventy-fifth percentile starting salary for medical school graduates in the South was \$35,300, while the comparable figure for law graduates in Georgia was \$100,000). Is there any explanation for the compensation difference, beyond anticompetitive restraints in the residency market? One possibility is a difference in demand facing the employer; law firm clients may have a higher willingness to pay for law firm associates' services than government and other purchasers are prepared to pay for residents' services. A second possibility is that physicians must have residency training to practice independently, while lawyers can practice independently without on-the-job training; physicians will therefore be less sensitive to low compensation levels in the short term. A third possibility is that law firms' decisions about compensation, at least traditionally, have been made in the context of a potential long-term relationship. Recruiting and training a new lawyer may cost upward of \$200,000. See Brenda Sandberg, *Who'll Stop the Raids?*, AM. LAW., June 2000, at 20. For this reason it may take two to three years before a law firm associate begins to financially benefit a firm. Joel A. Rose, *Adjusting Compensation for Today's Economics*, PENN. L. WKLY., Oct. 14, 2002, at F6, F6 ("Most [hiring partners] acknowledged that during the first two or three years with the firm, while in training to learn how to practice law, the great majority of associates are not worth the current market price."). Law firms may be willing to provide training and high salaries because they anticipate that their efforts will not only attract top-notch associates to the firm, but also encourage them to remain; top-notch associates will eventually become profit-generating firm partners. See Ginsburg & Wolf, *supra* note 152, at 953-54 (explaining that while "a residency director picks a future co-worker[,] . . . a hiring partner picks a potential future co-owner" who may affect that "partner's personal financial and reputational equity over years to come"). The trajectory for residents is different. While hospitals could conceivably make long-term arrangements with some of their residents (and some residents' programs already last longer than three years), the nature of

the total benefit received by medical residents may not actually be low relative to other providers when both monetary and nonmonetary forms of compensation are included in the calculation.²⁶⁶

Ultimately, competitive compensation levels are determined by supply and demand, and supply and demand conditions differ for each of the groups that might be compared to residents, for reasons unrelated to the match. It is therefore unsurprising that compensation levels would differ. As previously explained, employers will likely be willing to offer higher compensation to experienced physicians than to residents, and residents are more likely to accept lower compensation than physician assistants, because of the on-the-job training the residents receive.²⁶⁷ But compensation levels may differ for other supply- and demand-related reasons as well.

For example, in addition to providing training, residency positions also provide the credentials necessary for a future career as a physician.²⁶⁸ Participation in a residency program is a precondition both for future licensure as a physician and for board certification.²⁶⁹ Given these benefits of residency programs, it is not surprising that demand for these positions is high or that residents are willing to work for less monetary compensation than other professionals (such as physician assistants).²⁷⁰ In 2003, there were 23,965 active applicants—including 14,332 U.S. senior students, 1,987 U.S. foreign graduates, and 5,029 non-U.S.

physicians' work often changes after their training is complete. Many physicians develop office-based rather than hospital-based practices; while such physicians may on occasion care for hospitalized patients, they are often not hospital employees and are not required to admit their patients to any particular hospital. See BUREAU OF LABOR STATISTICS, *supra* note 251 (noting that less than one quarter of physicians are employed by hospitals). Thus, hospitals may not ultimately receive the long-run returns that justify law firms' high level of investment in and compensation of their associates. See Chae, *supra* note 1, at 355 (noting hospitals' disincentive to train because "laborers can leave and take their skills elsewhere").

266. Students benefiting from training provide their services at low cost in other professions as well. Nursing students, for example, have also traditionally served as a source of cheap labor. See LUDMERER, *supra* note 39, at 422 n.61 (citing SUSAN M. REVERBY, *ORDERED TO CARE: THE DILEMMA OF AMERICAN NURSING, 1850-1945* (1987)).

267. See *supra* note 171 and accompanying text (discussing the value residents place on the educational benefit of residency programs).

268. See *supra* note 39 and accompanying text (discussing the residency requirement).

269. See *supra* note 39 and accompanying text.

270. See Crall, *supra* note 1, at 270 (noting the "steep nature of the demand curve for medical residencies" arising from the gateway role of medical residencies). For further discussion of the gateway effect, see *supra* subpart III.D (explaining that low residency compensation might not deter potential residents because of the long-run benefits of a medical career).

foreign graduates—for 20,908 positions.²⁷¹ While supply and demand conditions for each program vary, the higher the number of qualified applicants, the less programs will need to pay to attract candidates, all else equal.²⁷²

Teaching hospitals' demand for residents (and thus, the supply of residency positions) will depend on the costs and benefits of employing residents, costs and benefits that are likely to differ from those associated with employing other health care professionals. Costs of employing residents include the direct costs of training, such as the cost of physician supervision.²⁷³ Other costs are indirect. An analysis based on 1981 data indicated that inpatient hospital costs "increase[d] by about 5.8 percent for every 10 percent increase in the resident-to-bed ratio" of the hospital.²⁷⁴ In part, these higher costs may result from inefficiencies in residents' provision of care; residents may, for example, order more tests than more experienced physicians would.²⁷⁵ To the extent that these costs are connected with the employment of residents, rather than other professionals, they may affect the demand for residents and thus affect equilibrium resident compensation.²⁷⁶

Teaching hospitals' demand for residents is also determined by the benefit residents generate, which in turn is derived in part from the demand of payers for resident and teaching hospital services. Historical evidence suggests that resident compensation increases in the post-World War II (and post-match) era were tied to reimbursement by private insurers and the creation of Medicare and Medicaid, programs that provided financial support for graduate medical education.²⁷⁷ Today, Medicare

271. See NAT'L RESIDENT MATCHING PROGRAM, *supra* note 216, at 5, 7.

272. See HERBERT HOVENKAMP, ECONOMICS AND FEDERAL ANTITRUST LAW 6–7 (1985) (discussing the basic economic principles of supply and demand).

273. MEDICARE PAYMENT ADVISORY COMM'N, REPORT TO THE CONGRESS: MEDICARE PAYMENT POLICY 53 (2003).

274. *Id.* at 55.

275. See Huckman & Barro, *supra* note 248, at 11 (describing studies finding that inexperienced physicians tend to provide more costly care). Higher costs may also reflect teaching hospitals' tendency to offer sophisticated technological services. Such services are expensive to provide and also attract more severely ill patients than nonteaching hospitals typically serve. MEDICARE PAYMENT ADVISORY COMM'N, *supra* note 263, at 5.

276. One commentator suggests that the cost of training medical residents in fact exceeds the value these residents generate until the residents' third year of residency. Christine Wiebe, *Federal Cuts for Medical Training—Bad Medicine or Chance to Improve?*, ACP OBSERVER, July 1996, at 1, available at <http://www.acponline.org/journals/news/jul96/fedcuts.htm>.

277. See LUDMERER, *supra* note 39, at 192 (arguing that interns and residents received salaries for the first time as a result of private insurers' policy to reimburse hospitals for their services, and stating that in the four years after the passage of

provides the bulk of financial support tied explicitly to the employment of residents.²⁷⁸ When a resident provides care to a Medicare beneficiary, the resident is not permitted to bill Medicare for his or her physician services as a nonresident physician would.²⁷⁹ Medicare instead compensates the hospital for the resident's services through its "direct" funding of the Graduate Medical Education (GME) program.²⁸⁰ In 1998, Medicare direct GME payments totaled over two billion dollars.²⁸¹

To compensate for the higher costs of operating a teaching hospital, Medicare also applies an "indirect" medical education adjustment to the formula it would otherwise use to pay for the hospital care supplied to Medicare patients.²⁸² This formula, like the direct GME formula, depends in part on the number of residents participating in an institution's GME program.²⁸³ Specifically, it increases the amount a teaching hospital would otherwise receive by a factor based on the resident-to-bed ratio of the hospital.²⁸⁴ In 1998, for example, the formula mandated that hospitals receive a seven percent increase in reimbursement for every ten percent increase in the resident-to-bed ratio.²⁸⁵ In that year, Medicare paid nearly five billion dollars in indirect medical education subsidies.²⁸⁶ Although there is no requirement that the direct or indirect medical education subsidies be passed on to

Medicare and Medicaid, the median house staff salary doubled).

278. See Wiebe, *supra* note 276 ("Through Medicare's hospital trust fund[,] . . . the federal government foots the bill for much of this training by including payment for direct and indirect teaching costs in hospital reimbursements.").

279. 42 C.F.R. § 415.200 (2004). Medicare may, however, be billed for physicians' services if a teaching physician personally furnishes services to a patient, or if the resident furnishes the service in the presence of a teaching physician. § 415.170. Medicare will also pay for physicians' services provided by residents who are fully licensed and are "moonlighting" by providing services outside of their residency program (in a hospital emergency room, for example). § 415.208. The antitrust suit plaintiffs offered as evidence of anticompetitive behavior the fact that hourly resident compensation was "substantially lower than what those same physicians can (and sometimes do) earn providing services to a hospital during off-duty hours ('moonlighting')." Complaint, *supra* note 1, ¶ 93. The fact that Medicare reimburses for physician services provided by residents while moonlighting, but not for services provided through the residency program, may contribute to this compensation differential.

280. See generally 42 C.F.R. §§ 413.75-.83.

281. COUNCIL ON GRADUATE MED. EDUC., FIFTEENTH REPORT, FINANCING GRADUATE MEDICAL EDUCATION IN A CHANGING HEALTH CARE ENVIRONMENT 23 tbl.1 (2000), available at <http://www.cogme.gov/15.pdf>.

282. MEDICARE PAYMENT ADVISORY COMM'N, *supra* note 273, at 53.

283. See *id.*

284. *Id.*

285. *Id.* at 55. The IME adjustment for 2003 was 5.5%. *Id.*

286. COUNCIL ON GRADUATE MED. EDUC., *supra* note 281, at 23 tbl.1.

residents, Medicare financing helps to support resident compensation at its current levels.²⁸⁷

On the other hand, the lower the overall reimbursement levels for resident and teaching hospital services—and the higher the costs of resident training—the less teaching hospitals will be willing and able to pay to obtain resident services, all else equal. Medicare provides significant financial support for the provision of medical education, but not so much that every hospital chooses to employ residents.²⁸⁸ While private payers historically subsidized education-related costs through higher payments to teaching hospitals, these subsidies have diminished in recent years.²⁸⁹ In general, it is clear that the relationship between the employment of a resident and financial benefit to a teaching hospital is complex; the economics of employing a resident are quite different from those of employing a nonresident physician or a physician assistant. Compensation levels for residents are therefore likely to differ from those for other medical professionals.²⁹⁰

Instead of comparing residents to nonresidents, the ideal empirical analysis would compare residents to residents. More specifically, it would compare the compensation of residents in a market with a matching mechanism to the compensation of the same residents in the same market without such a mechanism. Unfortunately, since a market cannot simultaneously have a match and not have a match, such a comparison is impossible to observe in practice. An alternative methodology would involve comparing compensation before and after a market adopts a match. While no published study appears to have undertaken a systematic analysis of this sort, historical research has shown

287. Several scholars have noted the direct connection between residency education subsidies and the demand for residents, although not necessarily the indirect connection between subsidies and the wages paid to residents. *See, e.g.*, Fitzhugh Mullan, *The Case for More U.S. Medical Students*, 343 *NEW ENG. J. MED.* 213, 215 (2000) (“[T]he current limitation on the number of Medicare-reimbursed residency positions provides a strong disincentive to expand training programs.”). Sean Nicholson and David Song found in an empirical study that a \$1,000 increase in indirect medical education payment per resident increased the number of residents hired by 0.12%. They speculated that one possible reason for the small increase was the constrained supply of residents. Sean Nicholson & David Song, *The Incentive Effects of the Medicare Indirect Medical Education Policy*, 20 *J. HEALTH ECON.* 909, 928 (2001).

288. *See* MEDICARE PAYMENT ADVISORY COMM’N, *supra* note 273, at 49 tbl.2A-6 (showing that in 2001 there were 1122 teaching hospitals and 3166 nonteaching hospitals).

289. COUNCIL ON GRADUATE MED. EDUC., *supra* note 281, at 2.

290. The fact that government intervention in the medical education market may complicate the evaluation of antitrust claims is noted by Hammer & Sage, *supra* note 256, at 93.

that before World War II—and thus before the match—first-year residents worked in return for room and board with minimal additional compensation.²⁹¹ The presence of low stipends before the match cannot conclusively prove that the match does not depress salaries, because it is possible that post-World War II salaries would have been even higher had the match not existed. It does strongly suggest, however, that factors other than the match—such as the cost and value of training—may be the source of low stipends.

A second methodology would involve comparing compensation in markets with a match to similar markets that exist at the same time but do not have a match. Professors Muriel Niederle and Alvin Roth have shown that internal medicine subspecialties that use matching mechanisms to fill fellowship positions (positions that offer training beyond that provided by residency programs) do not offer salaries significantly below those of programs that do not use the match.²⁹² One concern about this methodology is that programs that choose to adopt the match may differ systematically from those that do not. For example, if programs believe that the match facilitates wage suppression, programs subject to upward pressure on wages would be more likely to adopt the match; if the programs' belief were correct, then the match might successfully lower wages to the levels existing in programs without matches. The result would be similar wages in both programs, despite the effect of the match. However, to the extent that the adoption of matching mechanisms is a result of concerns unconnected to wages, such as concerns about the problematic consequences of early offers, this evidence shows that the residency match itself does not depress compensation.²⁹³

291. See Frank Michota, *Do Today's Residents Really Have It Better?*, 64 CLEVELAND CLINIC J. MED. 457 (1997), available at <http://www.cejm.org/1minuteconsults/oct7com.htm>. Ludmerer reports that "house officers received token compensation: from nothing to \$10 a month for interns, and \$10 to \$25 a month for residents." LUDMERER, *supra* note 39, at 96.

292. Niederle & Roth, *supra* note 259, at 1153. While concluding that "eliminating the resident match would not necessarily increase residents' wages," the authors also note the significant differences between fellowship and residency programs. *Id.* These differences may preclude the application of insights about fellowship wages to the residency wage context. See Sanders H. Chae, Correspondence, *Is the Match Illegal?*, 348 NEW ENG. J. MED. 2255, 2261 (2003) (arguing in response to a letter to the editor that a labor surplus might produce low salaries in the fellowship market, while there is a labor shortage in the residency market for American medical graduates). For further research on gastroenterology fellowships, see generally Muriel Niederle & Alvin E. Roth, *The Gastroenterology Fellowship Match: How It Failed and Why It Could Succeed Once Again*, 127 GASTROENTEROLOGY 658 (2004).

293. See *supra* Part III.B (discussing Roth's research on the problem of market

Thus, while empirical evidence suggests that resident compensation is indeed low relative to that of nonresident physicians and other health care professionals, these differences may result from factors unrelated to the match, including differences in productivity, nonmonetary compensation in the form of medical training, and variations in supply and demand conditions. (The lawsuit plaintiffs might also argue that the differences result from the defendants' sharing of stipend information²⁹⁴ and various activities of the ACGME, including limits on the number of residency positions,²⁹⁵ but a full analysis of these issues is beyond the scope of this Article.) Furthermore, while not conclusive, empirical evidence comparing compensation in markets with matching mechanisms to compensation in markets without matching mechanisms tends to suggest that such mechanisms do not have a substantial effect on compensation.²⁹⁶

V. THE ANTITRUST EXEMPTION AND OTHER REGULATORY RESPONSES

The procompetitive justifications for the match and the lack of empirical evidence of an anticompetitive effect together suggest that the match should be preserved. The likelihood that the match is procompetitive, Congress's concern about the

"unraveling"). Another methodology would involve a comparison of compensation trends over time in programs adopting or abandoning the match to trends in programs that have not. No published study appears to have undertaken such an analysis, perhaps because of data limitations.

294. For example, the AAMC conducts an annual wage survey that may enable collusion on employment terms by facilitating uniform wage-setting and the monitoring of deviations from uniformity. *See* Complaint, *supra* note 1, ¶¶ 73–82. On the other hand, the collection of wage information may also benefit competition by giving both residents and programs the information they need to assess whether their compensation levels remain competitive.

295. *See supra* note 2 (describing plaintiff allegations with respect to accreditation activities); *supra* note 212 (describing ACGME limitations on residency positions). If the ACGME limits do constrain the number of residents trained within a specialty, they will also constrain the number of physicians ultimately practicing within the specialty. Chae, *supra* note 1, at 355. The result is that although resident compensation might be less than the competitive level, residents will face fewer competitors upon completion of the program and may therefore ultimately obtain compensation at higher levels as independently practicing physicians. *Id.* While teaching hospitals may benefit from lower resident wages in the short run, they may be harmed by higher physician salaries in the long run. Higher physician salaries will tend to increase hospital expenses if the hospitals employ physicians. Even if hospitals do not employ physicians, high salaries may harm hospitals by decreasing demand for physician services that are complementary to hospital services. Thus, while resident limits may at first seem to benefit teaching programs and hospitals, it is not clear that in the long run the limits would have such an effect.

296. Niederle & Roth, *supra* note 259, at 1153.

“burden and expense of defending against litigation,”²⁹⁷ and its finding that the match has “effectively served the interests of medical students, teaching hospitals, and patients for over half a century”²⁹⁸ justify the grant of an antitrust exemption.

As Senator Kohl observed, however, it is often bad policy to provide exemptions to antitrust laws.²⁹⁹ The antitrust laws protect competition because competition produces many societal benefits. Concern about undermining these benefits is one reason why exceptions to antitrust law are so limited. Statutory exemptions are rare.³⁰⁰ The state action doctrine exempts state government regulation from antitrust law, but only if “[t]he challenged activity is authorized by a ‘clearly articulated’ state regulatory policy,” and “[a]ny private conduct authorized by the state policy is ‘actively supervised’ by an appropriate governmental agency.”³⁰¹ Given our reluctance to grant broad exceptions to the antitrust laws, it is reasonable to ask whether the antitrust exemption for the residency match is appropriately limited.

By attaching to the exemption legislation the label “Confirmation of Antitrust Status,”³⁰² the authors of the legislation were in effect asserting that the exemption was consistent with the purposes of antitrust law and presumably with the outcome of a properly decided antitrust suit. Consistency would require, however, that the antitrust exemption be limited in scope. Consider the series of arguments that might be raised if a suit focusing on the match were fully litigated. The plaintiffs would allege that teaching hospitals across the country acted together through the residency match to depress compensation in violation of the Sherman Act. The defendants would argue in response that the potential of the match to promote competition by facilitating fully informed decisionmaking suggests that it should be analyzed under the rule of reason.³⁰³ But the defendants’ ability to offer a

297. 15 U.S.C.A. § 37b(a)(2)(B) (West Supp. 2005). By early 2005, more than 100 attorneys had become involved in the case, and the “teaching medical establishment” was said to have incurred more than \$20 million in litigation-related costs. Croasdale, *supra* note 13.

298. 15 U.S.C.A. § 37b(a)(1)(E).

299. 150 CONG. REC. S3979 (daily ed. Apr. 8, 2004) (statement of Sen. Kohl).

300. For a list of such exceptions, see Memorandum from the Immunities & Exemptions Working Group to All Comm’rs, *supra* note 16, at 3–5.

301. HERBERT HOVENKAMP, FEDERAL ANTITRUST POLICY § 20.3, at 727–28 (2d ed. 1999).

302. 15 U.S.C.A. § 37b.

303. For a more thorough antitrust analysis of the case, see Miller & Greaney, *supra* note 1, at 915–16. For a discussion of the application of the rule of reason in general, see

procompetitive rationale for the operation of the residency match—the fact that the match may increase “output” by improving match quality—would not automatically result in a favorable verdict. Under the standard rule-of-reason analysis, a seemingly procompetitive restraint may still be deemed an antitrust violation if the plaintiff can demonstrate that it is not the least-restrictive alternative.³⁰⁴ One question that might be asked, then, is whether the procompetitive objectives of the match might be achieved with a mechanism that is less restrictive. If so, the exemption should be narrowed to encompass only the less restrictive alternative.

One argument along these lines is that while the matching algorithm itself should be subject to the protection of the exemption, some of the rules that support it should not be. One might argue, for example, that the rules of the match should be loosened to permit more contracting outside of it. As discussed in Parts II and III, match rules can be quite stringent. For example, residency programs participating in the residency match agree to hire seniors in U.S. medical schools only through the match.³⁰⁵ Violations of the match, including refusals to accept a match assignment, are subject to sanctions, including a prohibition against match participation and reporting of the violation.³⁰⁶ The antitrust suit plaintiffs objected to many of these rules.³⁰⁷ Supporters of the plaintiffs might argue that even if the exemption precludes an antitrust challenge to the operation of the matching algorithm, it should allow a challenge to some of the more restrictive rules of the match. The problem with this argument is that loosening the rules would likely destroy the residency match. The more positions allocated outside of the match, the smaller the benefit of the match to match participants. The more positions allocated outside of the match, the more pressure participants will feel to consider these positions, and the more the process will resemble the unraveling market the match is designed to avoid.

A second frequently articulated less-restrictive-alternative argument is that the unraveling problem could be avoided entirely through the imposition of rules related to offer timing.³⁰⁸

generally 7 PHILLIP E. AREEDA, ANTITRUST LAW ¶ 1511 (1986).

304. See AREEDA, *supra* note 303, ¶ 1511, at 429.

305. NRMP, *supra* note 68, § 4.2.

306. See NRMP, *supra* note 19, § 7.0 for a discussion of violation sanctions. Match rules do include some flexibility with respect to participants' needs. See *id.* §§ 2.5, 3.4 (contemplating the possibility of waiver of match commitments).

307. Complaint, *supra* note 1, ¶¶ 85–86.

308. See Crall, *supra* note 1, at 271–72 (suggesting that coordinated offer timing

Specifically, the imposition of a date before which programs would be prohibited from making offers could address the early-offer problem, while a mandatory minimum period for holding offers open could address the exploding-offer problem.³⁰⁹ Advocates of such rules might argue, in fact, that the antitrust exemption should be completely eliminated because these rules would achieve all of the benefits of the match in a less restrictive way.³¹⁰ Historically, however, such rules have failed.³¹¹ As discussed in subpart III.B, it was the failure of precisely these sorts of rules that led to the creation of residency match.

While strict enforcement of these rules would ensure that both programs and residents would have the opportunity to evaluate potential partners, it would fail to address an important contributor to the unraveling phenomenon. Because programs still would face self-imposed or ACGME-imposed resource-based constraints on the number of positions available,³¹² they would still have to make offers sequentially, waiting for a definitive negative response before making another offer. A hospital that cannot afford to finance residency positions from its own funds, for example, would not want to risk making offers (and receiving acceptances) for more positions than Medicare is willing to subsidize.³¹³ This means that programs would have to target their initial offers carefully in order to avoid losing their favored candidates to other programs. They might choose to make an offer only to their second-choice candidate, for example, because they suspect that their first-choice candidate will eventually take another offer. Residents, meanwhile, would face a choice between accepting the first position offered them or declining and hoping that a preferred offer would arrive later. These are once again the problems the match is designed to avoid. Imposition of deadlines alone could not achieve the procompetitive benefits of the match.

On the other hand, it may be possible to enhance the procompetitive benefits of the match by increasing competition *inside* of it. First, programs should be required to disclose

would be a less restrictive means of “keeping transaction costs low, avoiding informational problems, and eliminating externalities”).

309. See, e.g., *id.* (suggesting that hospitals be required to “make their offers by a certain date” and “hold offers open until a certain later date”).

310. *Id.* at 271–74 (arguing that regulation of offer dates and acceptance deadlines is a less restrictive means of achieving the benefits of the match).

311. See Roth, *supra* note 57, at 993–95.

312. See *supra* note 212 and accompanying text (describing residency position limits).

313. See *supra* Part IV (discussing the financing of medical residency positions, including the cap on the number of Medicare-subsidized positions).

compensation levels and other terms of employment before the ranking deadline. If programs disclose compensation only after the matching process, competitive pressure is limited: Once matched, residents are contractually committed to join the assigned residency program, their second-choice programs are likely full, and other possibilities (such as waiting for next year's match, or pursuing another profession) are undesirable. If programs always disclosed their compensation packages before the ranking deadline, on the other hand, prospective residents could easily compare these packages across programs and take this information into account when compiling their rank-order lists. Programs currently are free to provide compensation terms before the ranking deadline, and most likely do.³¹⁴ A disclosure requirement would ensure that all programs provided such information, however, and not just upon the request of individual applicants who might be reluctant to ask. In recognition of the merits of such a policy, the NRMP recently adopted a rule that requires programs to provide a sample of the contract applicants would be expected to sign upon being matched to the residency program.³¹⁵

Because such disclosure is critical to the effective functioning of the matching mechanism, Congress should incorporate into the antitrust exemption statute a similar requirement that programs provide all interviewed applicants a copy of the actual contract that they would be expected to sign. If structured as a condition for the exemption, the requirement would give an additional incentive for programs to provide compensation information. If the programs failed to comply, thus reducing the procompetitive benefits of the match, plaintiffs could then bring an antitrust action similar to the one brought by former residents in 2002.³¹⁶ Alternatively, Congress could create an independent mandate of disclosure, and give antitrust enforcement agencies (or the Department of Health and Human Services) the power to enforce it through fines or other measures.

314. Many disclose this information on program websites or through the Fellowship and Residency Electronic Interactive Database (FREIDA). *See, e.g.*, Am. Med. Ass'n, FREIDA Online, <http://www.ama-assn.org/ama/pub/category/2997.html> (last visited Oct. 1, 2005).

315. *See supra* note 65 (describing the current NRMP rule on sample contracts); *see also* Press Release, Nat'l Residency Matching Program, NRMP Requires Medical Residency Programs to Show Contracts to Applicants (Oct. 31, 2003), *available at* <http://www.nrmp.org/contractpr.pdf> (reporting that the new rule requiring the provision of sample contracts would be effective in 2005).

316. *See generally* Complaint, *supra* note 1.

Either alternative would permit enforcement approaches more expansive than those currently available to the NRMP.

To achieve the maximum possible competitive benefit, it is imperative that program disclosures be accurate. In other words, the “sample contract” should not just be a representative example of the type of contract that the resident would be expected to sign, but instead an actual offer of a contingent contract. By submitting a rank-order listing that contains a particular program’s name, the resident would accept the program’s offer, forming a contingent contract. If the parties are successfully matched, they would then be bound to comply with the terms of the contingent contract. If a resident’s submission of rankings indeed constitutes an acceptance of all contingent contract offers, the resident will necessarily be able to compile rankings based on full and accurate information about program terms. The match’s competitive potential will be realized.

Until recently, however, the creation of contingent contracts appears to have been prohibited. Match rules expressly prohibited the formation of “any verbal or written contract” before the match;³¹⁷ presumably, this broad prohibition would have applied to contingent contracts. Prohibitions on contingent contracting, however, are not necessary to achieve the procompetitive benefits of the residency match;³¹⁸ elimination of the prohibition would thus be consistent with antitrust doctrine requiring that a procompetitive restraint be the least restrictive alternative. The recent NRMP requirement for disclosure of sample contracts, if interpreted to mandate adherence to sample contract terms, would seem to have eliminated this prohibition.³¹⁹ The federal antitrust exemption should reinforce this result by mandating that programs enter contracts contingent on the outcome of the match.

Another way to increase competition without undermining the match would be to encourage the creation of *individualized* contingent contracts. Rather than posting a sample contract with terms that would apply to all residents, programs could

317. NRMP, *supra* note 68, § 8.0.

318. See Hammer & Sage, *supra* note 256, at 93 (“[S]pecific details of the NRMP, such as prohibitions on negotiating terms of employment prior to matching, are troubling because they go beyond the type of ancillary restriction necessary to achieve the program’s legitimate goals.”).

319. An NRMP policy that did not mandate program adherence to sample contract terms would likely be less effective in increasing competition. Concerns about developing a reputation for deviating from sample contract terms may still encourage programs to adopt the sample contract as the actual contract, however, thus preserving the benefits of competition even in the absence of a mandate.

announce minimum stipends or stipend ranges, but then provide individualized sample contracts to each interviewed applicant. To the extent that these individualized sample contracts also contained individualized stipends, they would facilitate the competitive process. For example, the compensation offered would provide some evidence of the program's intensity of preference for a particular resident.³²⁰ In circumstances in which residents' preferences depend on programs' preferences for them, this information would help residents make more fully informed decisions. Residents could also use the information contained in individualized contracts as a tool to negotiate better offers from other programs. Admittedly, contingent contracts are less than ideal for both of these purposes because they are not necessarily credible. Programs' rank-orders need not correspond to the level of compensation offered; indeed, programs could offer excellent employment terms to all applicants, but then exclude some applicants from their ranking lists. To the extent that programs are concerned about filling their slots, however, they will include as many acceptable candidates as possible on their rank-order lists. In addition, transaction costs associated with individualizing contracts may discourage programs from providing contracts to candidates they do not intend to rank. In either case, the individualized contracts would facilitate competition.

Individualized compensation would mitigate a number of the concerns that have been raised in connection with the residency match. Both of the theoretical economic studies showing the relationship between matching mechanisms and wage depression focus on systems under which programs cannot distinguish

320. Because compensation offers, like rank-order lists, provide information about participants' preferences and intentions, it is reasonable to ask why the same rules should not apply to both. In particular, pre-match solicitation of rank-orderings from match participants is prohibited. *See supra* note 71 and accompanying text. Why should compensation information be treated differently? There are at least two justifications for prohibiting solicitation of ranking information. First, if the ranking information provided is incorrect or vague when given, or becomes incorrect due to later modification, then the prohibition prevents unethical behavior and ensures that participants are not misled. Second, if the ranking information is correct, then the prohibition prevents unraveling. If programs are permitted to make binding commitments to a particular ranking, then candidates must be sure to interview before the program makes commitments to other interviewees. The result is the same chain of events the match is intended to prevent. Neither of these two concerns applies to contingent offers of compensation. First, because the contract is binding if a match is made, there is no risk of a party being misled about compensation. Second, because there is no necessary connection between salary offers and rankings, because programs' financial exposure is limited by the number of residency slots offered through the match, and because budgets may be somewhat flexible, one high contingent offer of compensation need not preclude another. Thus, disclosure of compensation information need not induce unraveling.

among candidates in setting their wages.³²¹ But when compensation is allowed to vary in Bulow and Levin's model, the matching mechanism achieves the efficient result.³²² Similarly, individualized wages address the ordinal ranking concern raised by Professor Priest by allowing cardinal expressions of preferences.³²³ Programs can indicate the strength of their preferences through individualized compensation offers in addition to rankings, thus preventing the match from making inefficient assignments of residents to programs.

It is not clear that residency programs would choose to individualize employment terms. From an efficiency perspective, if all residents selected by a program had similar characteristics, there would be little reason to vary employment terms. Lack of proven variation in productive capabilities may be one reason that residents within a program are generally offered the same compensation. On the other hand, it cannot explain why compensation often does not vary *across* programs within an institution. Third-year residents in surgery programs may receive the same compensation as third-year residents in internal medicine programs, despite significant differences in the incomes of independent physicians in these specialties.³²⁴ The tendency of similarities in compensation to persist despite differences in skill sets may instead be due to expectations of "fair" treatment within an institution. Administering compensation packages may be less costly when they are consistent across employees, in part because the need to monitor for favoritism is reduced.³²⁵ Many employers offer similar compensation levels to their workers, despite potentially

321. See *supra* note 82 and accompanying text. Kamecke suggests that when programs do not offer individualized wages, there may be no wage that equates supply and demand; the result is "disorderly behavior" that is then addressed by a matching mechanism. See Kamecke, *supra* note 80, at 34. The Kamecke model assumes that wages within programs are uniform, and states that "it is the major objective of the paper to discuss the consequences of this restriction on the wage formation in an organized entry-level job market." See *id.* at 37.

322. See *supra* text accompanying note 174.

323. See *supra* Part III.C.3 (explaining Priest's argument that matching mechanisms based only on ordinal expressions of preferences can result in inefficiency).

324. For example, the Baylor College of Medicine Program pays \$41,517 to third-year residents in both the internal medicine and surgery programs; the University of Minnesota Program pays \$44,630 to third-year residents in both programs. See Am. Med. Ass'n, *supra* note 314 (providing detailed program information about the characteristics of individual residency programs).

325. See Susan J. Stabile, *Motivating Executives: Does Performance-Based Compensation Positively Affect Managerial Performance?*, 2 U. PA. J. LAB. & EMP. L. 227, 259 (1999).

observable differences in worker quality.³²⁶ For these reasons, it is not obvious that residency programs would, or even necessarily should, choose to individualize employment terms.

Given the potential benefits of individualized compensation for competition, however, regulators should facilitate individualized contracting for programs inclined to engage in it. In addition to potentially increasing efficiency, individualized contracting would make both express and tacit collusion much more difficult relative to a system under which institutions offered the same contract to all residents participating in the institutions' programs. Variations in compensation levels would complicate any efforts by independent institutions to agree upon a single level of compensation, or to enforce adherence to such an agreement.

A contingent contract mandate might itself encourage individualized contracting by encouraging open discussion of contract terms before rankings are submitted. Mandating that the contingent contracts be individualized—requiring that they include the applicant's name and signature, for example—might further facilitate the tailoring of compensation packages. While programs might still choose to offer the same compensation package to each of their applicants, the mandate would provide a natural opening for more compensation negotiation than would otherwise exist. The mandate would serve as an opportunity for programs to depart from any norms that might inefficiently discourage individual negotiation. Programs could publicize minimum compensation or compensation ranges on their websites or through the FREIDA database, so that prospective applicants would have general information about compensation before they apply, but then supplement this information with the individualized sample contract. It would not be appropriate, though, to actually mandate tailored compensation. Such a mandate would not likely be administrable,³²⁷ and individualized compensation might in some cases be determined to be inefficient or unfair. A mandate to provide individualized contingent contracts, however, might eventually increase the frequency of

326. See, e.g., Stephen Machin & Alan Manning, *A Test of Competitive Labor Market Theory: The Wage Structure Among Care Assistants in the South of England*, 57 *INDUS. & LAB. REL. REV.* 371, 371–74 (2004) (finding “surprisingly little wage dispersion within firms” despite the effective lack of “external constraints on the wage-setting process”). For discussions of the practice of uniform compensation and its potential causes, see, for example, Bulow & Levin, *supra* note 82, at 27, and MANNING, *supra* note 241, at 134–36.

327. For instance, how much would resident compensation packages be required to vary?

2005]

THE RESIDENCY MATCH

835

negotiation over compensation packages and thus improve competition within the residency match system.

VI. CONCLUSION

The antitrust suit filed by former medical residents against organizations that operate and participate in the residency match raised important questions about the process of graduate medical education in the United States. Residency training demands multiple years of work for long hours at compensation lower than that of other medical professionals. By preventing the creation of credible firm offers to individual residents, the residency match weakens individual residents' ability to negotiate better compensation. Economic theory suggests that under certain assumptions matching mechanisms may depress compensation and impede the efficient functioning of markets relative to a competitive ideal.

On the other hand, market imperfections and market failures almost certainly would prevent the realization of this ideal. The residency match is a market intervention that promotes competition by assuring market participants a wide range of potential partners and an opportunity for fully informed decisionmaking. The residency match helps to achieve a more efficient pairing of residents and programs than would be generated by an unraveling market, perhaps contributing to higher resident compensation than would otherwise exist, and likely increasing gains to society as a whole. Given these arguments and the considerable cost of litigation, the congressional grant of an antitrust exemption is justified.

Having sacrificed the protections provided by the antitrust laws, however, Congress has now taken on the responsibility for protecting the public against anticompetitive behavior. It is unfortunate that the antitrust exemption was enacted without a full airing of the arguments on both sides of the residency debate. Much remains to be understood about the nature and magnitude of the effects of the match. When Congress created a retroactive antitrust exemption, it likely precluded an exploration of these issues through litigation. It should therefore now actively study the effects of the match, and work to promote competition through measures such as a legislative requirement for pre-ranking deadline offers of contingent contracts to individual match participants.

In its roles as a payer for medical education, a purchaser of care, and a protector of the public interest, Congress should consider the effect of market interventions such as the match on

the welfare of not just residents and residency programs, but also the public as a whole. In a market characterized by failures and imperfections, market interventions may improve social welfare. Continued vigilance is required to ensure that they actually do so.