

Data Access Information for:
Can Marginal Rates of Substitution Be Inferred
From Happiness Data?
Evidence from Residency Choices

Daniel J. Benjamin
Cornell University and NBER

Ori Heffetz
Cornell University and NBER

Miles S. Kimball
University of Michigan and NBER

Alex Rees-Jones
Cornell University and NBER

This Draft: February 12, 2014

This document describes the procedure to request supplemental analysis of the data from “Can Marginal Rates of Substitution be Inferred from Happiness Data? Evidence from Residency Choices.” In section 1, we briefly describe the need for maintaining the confidentiality of this dataset. In section 2, we describe the options available for requesting additional analysis.

1. The need for confidentiality

In the process of preparing to collect the data for this paper, we determined that it would be necessary to provide our survey respondents with a level of confidentiality substantially higher than is standard in experimental or survey work. In our setting, this confidentiality serves two purposes. First, it ensures that we protect the respondents in our subject pool, who are being asked to share private information with a potential to be professionally damaging. Second, assuring our subjects of confidentiality was deemed to be essential to generate a reasonable survey response rate, given the strong norm of secrecy surrounding residency preference orderings. We elaborate on the rationale behind these two points below.

The successful operation of the medical residency match depends critically on the ability of students to report their preferences confidentially. If a residency is ranked relatively low on a student’s preference order, that residency might find this student’s matriculation less desirable. In the extreme, this could lead to discrimination against the student. As a result, the NRMP

actively protects students' right to keep their rank ordering confidential, and forbids residencies from compelling students to reveal this information. However, both sides of the market attempt to signal that the other is highly ranked, although such communications constitute unverifiable "cheap talk." If residencies could observe the preference order students submitted (as they potentially could from our dataset), it might (a) create the possibility for certain types of collusion, or (b) offer the opportunity to verify if the student pursued the ranking strategy they reported in private communications, and to punish the student if s/he did not. Either of these possibilities creates a potential for damaging outcomes, both to the functioning of the market as a whole, but also to the individual students participating in our survey.

While our primary dataset does not include direct identifying information, the population of medical students is small enough that certain combinations of categorical information (e.g., specialty, school, ranked programs, age, gender, etc.) could uniquely identify a respondent. As a result, we determined that the public posting of this data could present an actual risk for our survey respondents.

The potential risk of being identified was readily apparent to a number of students we interviewed in the initial stages of designing our survey. Due to the concerns above, a strong norm of secrecy has evolved in the residency match. In conversations with both medical students and medical school officials, we were told that students are regularly advised to not share their preference information with effectively anyone. These same individuals expressed serious doubts that we would get a non-trivial response rate if confidentiality were not credibly assured.

In light of both the desire to protect our survey respondents, and the need to emphasize that they might participate without worry of professional damage, we included strong assurances of confidentiality in the introduction to our survey. In our informed consent document on the first page of our web survey, we say: "Your responses will be used solely for research purposes and will be kept strictly confidential, shared only with the researchers named below." This provision limits direct data access to the authors of this paper.

2. Procedures for requesting additional results or obtaining data

We are committed to maintaining our respondents' confidentiality while still permitting reasonable efforts at replication or further analysis. Detailed descriptions of the survey design, data, and analysis are available in the paper and its accompanying survey and web appendices.

The Stata .do file used to produce the reported analysis is also publicly available on the journal's website. Should researchers want to conduct additional robustness analysis, they may modify the Stata code to produce their desired analysis and send this file to the authors. We will run Stata code provided to us and report back the Stata output, as long as this output does not reveal sensitive, individual-level identifying information. Please contact Alex Rees-Jones with any such requests.

Some supplemental analysis, available in the web appendix and survey appendix, additionally makes use of data from the U.S. News and World Report medical school and hospital rankings. Access to ranking information may be purchased directly from U.S News and World Report. If you seek to request additional analysis based on our confidential data merged with U.S. News ranking information, we ask that you purchase an individual subscription to the relevant ranking data beforehand. To seek additional information or access to these rankings, go to <http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-medical-schools>.