NBER WORKING PAPER SERIES

CHOOSING YOUR POND:
MEASURING PREFERENCES FOR RELATIVE CONSUMPTION

Nicolas L. Bottan
Ricardo Perez-Truglia

Working Paper 23615
http://www.nber.org/papers/w23615

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2017, Revised October 2017

Previously circulated as "Choosing Your Pond: Revealed-Preference Estimates of Relative Income Concerns." We are thankful for comments from David Albouy, Dan Bernhardt, Jeff Brown, Mikhail Galashin, Fred Finan, Craig Fox, Ori Heffetz, Alex Imas, Edward Leamer, Erzo F.P. Luttmer, Ben Marx, Alex Rees-Jones, Romain Wacziarg, Melanie Wasserman and seminar discussants at UCLAAnderson, UIUC, Loyola Marymount University, 2017 SITE Workshop, 2017 Advances with Field Experiments Conference, 2017 GEM-BPP Workshop, Binghamton University, and the Federal Reserve Bank of New York. We thank UCLA-Anderson and its Behavioral Lab for providing funding for the experiment; and thank the support from the Robert Ferber Dissertation Award. This project was reviewed and approved in advance by the Institutional Review Board at University of California Los Angeles (IRB #16-001968). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Choosing Your Pond: Measuring Preferences for Relative Consumption
Nicolas L. Bottan and Ricardo Perez-Truglia
NBER Working Paper No. 23615
July 2017, Revised October 2017
JEL No. D62,D91,I31,Z13

ABSTRACT

We provide a unique revealed-preference test of the hypothesis that, in addition to their absolute level of consumption, individuals care about their relative consumption. We study the decisions of senior medical students participating in the National Residency Match Program (NRMP). They must choose between programs that offer similar nominal incomes, but in cities with different costs of living and income distributions. As a result, they face trade-offs between absolute consumption and relative consumption. We conducted a survey experiment with 1,100 NRMP participants. We elicited their perceptions about cost of living and income distribution in the cities that they are considering living in, as well as their rank order submissions. To assess the direction of causality, we embedded an information-provision experiment that generates exogenous variation in perceptions. We find that, holding absolute consumption constant, the average individual prefers higher relative consumption. Moreover, we find substantial and meaningful heterogeneity in relative concerns by relationship status.

Nicolas L. Bottan
Department of Economics
University of Illinois at Urbana-Champaign
1407 W Gregory Dr.
214 David Kinley Hall, Mc-707
Urbana, IL 61801
bottan2@illinois.edu

Ricardo Perez-Truglia
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza
Los Angeles, CA 90095
and NBER
ricardo.truglia@anderson.ucla.edu

1 Introduction

For a long time, economists have posited that, in addition to their absolute consumption, individuals also care about their relative consumption (Duesenberry 1949). Incorporating relative consumption concerns in economic models is of importance because it can improve the ability of models to explain behavior, and because of the significant policy implications.[1] However, identifying and quantifying preferences for relative consumption has proved challenging. The existing evidence mainly relies on happiness data (e.g., Luttmer 2005) and on decisions in laboratory experiments (e.g., Solnick and Hemenway 1998; Kuziemko et al. 2014). In this paper, we offer unique revealed-preference evidence based on a field experiment with 1,100 medical students participating in the National Residency Match Program (NRMP).

An ideal context for studying whether individuals care about relative consumption would be one in which individuals have to choose from a list of cities where they would get different combinations of relative and absolute consumption. In such a setting, one would compare how individuals make trade-offs between these two attributes. However, such datasets do not exist. Although some datasets identify individuals moving from one location to another,[2] they do not include sufficient information to estimate these preferences; for instance, they do not identify the alternative combinations of absolute and relative consumption that the individual could have gotten in other locations.

We collect this ideal data by taking advantage of a unique context. Graduating U.S. medical students have to rank their preferred residency programs for the NRMP, and this information is used to match students to the residency programs that will employ them for roughly 5 years. The students have to choose between programs that offer similar nominal income, but in cities with largely different costs of living and income distributions. As a result, these job candidates face trade-offs between the absolute consumption (defined as the nominal earnings divided by the cost of living index) and relative consumption (defined as the individual's rank in the distribution of absolute consumption in the same city).

Several features of the NRMP setting make it uniquely apt for this type of revealed-preference analysis (Benjamin, Heffetz, Kimball and Rees-Jones 2014). First, since there is a deadline to submit the rank order, there is an identifiable moment when the decision becomes irreversible. Second, it is possible to identify and specify the entire choice set faced by these individuals. Third, because of the incentive-compatible matching algorithm used by

[1] For instance, relative consumption concerns create positional externalities that one can correct for via consumption or income taxes (Boskin and Sheshinski 1978; Frank 1985a).

[2] For instance, the National Survey of Families and Households (Luttmer 2005), or the United States Postal Service's National Change of Address (Perez-Truglia 2017).

the NRMP, which students are largely aware of (Rees-Jones 2017), we can infer preferences directly, without having to estimate a complicated model. Fourth, this is a high-stakes choice to which participants devote ample time and attention—indeed, this decision is arguably one of the most important for their careers and lives.

We conducted a survey with 1,100 senior medical students participating in the 2017 main residency match. The survey asked participants to list their top two favorite programs among those they were considering. We concentrated on the rank of the top two because that is the part of the decision with the highest stakes and to which individuals pay the most attention. We elicited perceptions about two aspects of the cities in which these two programs were located: the cost of living and the earnings rank (i.e., the position in the distribution of earnings in the city). Variation in these two perceptions can identify preferences over absolute and relative consumption: with the nominal earnings held constant, decreasing the cost of living increases the expected absolute consumption without affecting the relative consumption. And with the nominal earnings held constant, increasing the earnings rank increases the expected relative consumption without affecting the absolute consumption.

After obtaining perceptions, we elicited each subject's expected rank submission. With these data on perceptions and choices, we can estimate how city differences in absolute and relative consumption affect location choices. We expect individuals to prefer higher absolute consumption. Individuals may prefer higher or lower relative consumption, depending on the mechanisms at play. On the one hand, models of status and consumption aspirations predict that individuals prefer higher relative consumption (i.e., they prefer to locate in poorer ponds). On the other hand, if individuals want to interact with richer neighbors, such as single individuals looking for a partner, we would expect them to prefer lower relative consumption (i.e., they prefer to locate in richer ponds).

One potential concern is that perceptions about relative and absolute consumption may be correlated with unobservable attributes of the options, which can generate omitted-variable biases. To deal with this concern, we embedded an information-provision experiment in the survey. Right after eliciting perceptions about cost of living and earnings rankings, we provided all individuals with statistics about these two measures, and again asked individuals for their perceptions. We randomized the value of the feedback given to the individual, in a non-deceptive way, by randomizing the data source used to compute these statistics. For instance, an individual who is considering earning \$55,000 in Champaign-Urbana, IL will receive one of two messages: that, according to data from the Current Population Survey, her earnings rank will be 71.4%; or that, according to data from the American Community Survey, her earnings rank will be 62.7%. This source-randomization experiment creates exogenous variation in perceptions. We can use that exogenous variation in instrumental variables model to estimate the causal effect of perceptions on choice.

Our baseline estimates, which use the experimental as well as the non-experimental variation in beliefs, suggest that a 1 percent increase in absolute consumption increases the probability that a program is chosen by 0.202 percentage points (i.e., a behavioral elasticity of 0.202). Indeed, this concern for absolute consumption is consistent with the fact that half of pre-med students report money to be the primary motivation for their career choice (Daniel and O'Brien 2008). Even though this preference for absolute consumption is statistically and economically significant, it is by no means the primary concern for doctors – we find that doctors care substantially more about other characteristics of the residencies, such as their prestige and career prospects.

Most important, our baseline estimates suggest that, conditional on absolute consumption, individuals also care about relative consumption: on average, a 1 percentage point increase in relative consumption increases the probability that a program is chosen by 0.185 percentage points (i.e., a behavioral elasticity of 0.185). These baseline estimates suggest that the average individual gives roughly the same importance to relative consumption as to absolute consumption.

However, these average estimates mask meaningful heterogeneity. While non-single individuals (i.e., married or in a long-term relationship) prefer to be richer than their neighbors, single individuals have the opposite preferences. This difference in preferences is large in magnitude and highly statistically significant ($p = 0.001$). This heterogeneity is consistent with prior evidence from the happiness literature: Luttmer (2005) finds that the positive effects of relative income on happiness are driven entirely by non-single individuals. Our favorite explanation is that single individuals want richer neighbors because they want to date from a pool of richer individuals. In this sense, the evidence suggests that at least part of the concerns for relative consumption are instrumental rather than purely hedonic (Benabou and Tirole 2006).

We find the estimated preferences for relative consumption to be robust to a number of checks. First, these estimates are not sensitive to the inclusion of several residency and location characteristics as control variables. Second, these baseline estimates are consistent with the experimental estimates that focus on the variation generated by the information provision experiment. Third, we find that the information related to relative consumption, provided roughly one month before the submission deadline, had a long-lasting effect on the final rank order submitted to the NRMP. On the other hand, the coefficients on absolute consumption are weaker under some of the alternative specifications. As a result, if anything, our baseline estimates underestimate the importance of relative concerns.

Our paper is related to several bodies of research, including the literature on the effect of relative income on subjective well-being. Since the seminal contribution by Easterlin (1974), several studies have shown that, with own income held constant, subjective well-being increases with the relative income in the area of residence (van de Stadt, Kapteyn, and van de Geer 1985; Clark and Oswald 1996; Luttmer 2005; Ferrer-i-Carbonell 2005; Perez-Truglia 2015).[3] Our contribution to this literature is twofold: we estimate preferences for absolute and relative consumption using revealed-preference rather than well-being data, and we disentangle the direction of causality using an experimental design.

Our estimates are not directly comparable to those from the happiness literature because of differences in contexts and in model specifications. With that caveat in mind, our baseline estimates suggest that relative consumption concerns are smaller in magnitude than the relative concerns estimated with happiness data in Luttmer (2005); however, due to lack of precision, we cannot reject the hypothesis that these two estimates are equal. Additionally, our results shed light on an important aspect that is impossible to study with happiness data: even if we assume that happiness depends on relative consumption, it is unclear whether individuals anticipate the externalities brought by neighbors (Luttmer 2005). Our findings suggest that individuals anticipate at least part of these externalities.

Our paper is also related to studies that use surveys that ask individuals to choose between pairs of hypothetical scenarios that encompass trade-offs between income and status or between positional and nonpositional goods. These studies find that individuals are sometimes willing to give up absolute income in exchange for higher status (e.g., Solnick and Hemenway 1998; Johansson-Stenman, Carlsson, and Daruvala 2002; Yamada and Sato 2013; Clark, Senik, and Yamada 2017). We contribute to this literature by estimating these trade-offs in a real-world context that has high stakes.

Our study also contributes to a growing literature showing that individuals have substantial misperceptions of their own income rank (e.g., Cruces, Perez-Truglia, and Tetaz 2013; Karadja, Mollerstrom, and Seim 2017). While this literature shows that correcting these misperceptions has significant effects on stated preferences for redistribution, there is no evidence that these misperceptions have a significant effect on behavior. We fill this gap in the literature by showing that misperceptions about relative consumption can have meaningful economic consequences.

This paper also relates to a large literature studying the relevance of consumption amenities, such as cost of livings differentials, on household location decisions using a variety of methods (e.g., Albouy 2008, Holmes and Sieg 2015). Our study adds to this literature by

[3]These studies often use a slightly different specification: holding own income constant, well-being decreases with the average income in the group of reference. It must be noted that some studies find the opposite effect (Senik 2004) or mixed evidence (Clark, Kristensen and Westergård-Nielsen 2009). For an extensive review of the literature, see Clark, Frijters and Shields (2008).

directly estimating the causal effect of perceptions about certain city attributes on location choice, and by identifying relative consumption as an additional driver of location decisions. Our study is also related to a literature that studies conspicuous consumption (e.g., Heffetz 2011). And, regarding the potential mechanisms at play, this paper is related to a recent but growing literature using field experiments to study self and social image: e.g., Bursztyn and Jensen (2015), Perez-Truglia and Cruces (2017), Dellavigna, List, Malmendier and Rao (2017) and Bursztyn et al. (2017).

The rest of the paper proceeds as follows. Section 2 introduces the survey design. Section 3 presents the econometric models. Section 4 presents implementation details and descriptive statistics. Section 5 presents results on the distribution of perceptions and learning. Section 6 shows the main results on preferences over absolute and relative consumption. The last section concludes.

2 Survey Design

2.1 Timing of the Surveys

After graduating from medical school, students have to complete a residency to become a Medicinae Doctor (MD). A residency usually lasts from three to seven years, after which individuals may obtain their medical license. During the fall semester of 2016, fourth year medical school students started their participation in the residency match by submitting applications to residency programs. Later in the semester, they were interviewed and flown out by some of the programs they applied to.[4] After all interviews were completed, the students had roughly two months to decide on how to rank their favorite programs.

We follow Benjamin et al. (2014) in using this context to study preferences. They conducted a survey of medical students after the students had submitted their rankings to the NRMP. The survey measured the choice rankings submitted to the algorithm as well as the perceived rank of many aspects of the programs, such as life satisfaction, happiness, and sense of control. In that study, Benjamin et al. (2014) measure and compare the preferences inferred from rank choices to those inferred from subjective well-being. We follow the survey collection method from Benjamin et al. (2014) closely, but we change the survey itself to test a different hypothesis, that is, whether individuals make a trade-off between absolute and relative consumption. In doing so, we deviate from the survey design in an important aspect: we collect our baseline survey before subjects submit their rank choices to the NRMP.

[4] In 2015, the median number of applications submitted was 30 and the median number of interviews 16 (NRMP 2015).

We also embed an information-provision experiment, which allows us to address concerns on causality.

In the 2017 Match, the submission window for rank order lists opened on January 15 and closed on February 22. We conducted a baseline survey early in the submission period, which we describe first. We also conducted a follow-up survey after the submission window closed, which is described in Section 2.7.

2.2 General Structure of Baseline Survey

The baseline survey starts and ends with some background questions, such as the subject's medical school and marital status (see Appendix A.1 for the full questionnaire of the baseline survey). The core of the survey comprises the following group of questions, in the order listed below:

1. Choice Set: Elicit the names of the two favorite programs that the individual was considering for his or her order rank submission.

2. Prior Beliefs: Elicit perceptions about the cost of living and the earnings rank in the cities where these two programs are located.

3. Feedback: Provide subjects with feedback related to their perceptions.

4. Posterior Beliefs: Re-elicit perceptions about the cost of living and the earnings rank.

5. Rank Choice: Elicit the individual's expected rank submission (between the two programs).

The following sections provide details about each of these modules.

2.3 Choice Set

The survey asks individuals to list their top two preferred programs, in no particular order, from a user-friendly list of all the available programs organized by state and metro area. We limited the survey to two programs because otherwise it would have been too cognitively demanding. Most participants expect to be matched to one of their top-two choices: similar to previous years, 50.9% of the participants in the 2017 match were assigned to their first choice and 16.6% were assigned to their second choice. We concentrated on the participants' top two programs rather than a random pair of programs because this happen to be the part of the decision with the highest stakes and to which individuals were paying the most attention. In any case, our focus on the top two choices does not challenge the validity of our estimates: the research design would be valid with any pair (or group) of options, not only the top-two.[5]

2.4 Perceptions about Cost of Living and Earnings Rank

One important feature of the residency match process is that salaries are relatively homogeneous across the different programs, even across specialties.[6] Indeed, each program offers the same salary to all its candidates (and that salary is often publicly available on the program's website). Because incomes are quite homogeneous, and given that we cannot generate exogenous variation in incomes, exploiting income differences to identify preferences for relative and absolute consumption would not have been feasible. Instead, we exploit differences in costs of living and earnings distributions in the cities that the individual is considering living in. This approach generates variation in the absolute and relative levels of consumption that individuals expect for the duration of their residency programs.

When designing the survey, we were constrained to using metropolitan areas rather than other geographical levels of aggregation (e.g., commuting zones) because the sources of data on cost of living are not collected at a finer level than the metro area. With incomes held constant, the cost of living in a metro area affects absolute consumption without affecting relative consumption. If the cost of living decreases in an area, it increases one's absolute consumption level because one can afford more goods with the same nominal income. However, it also increases the absolute consumption level of everyone else in the city, leaving one's relative consumption rank unchanged. In contrast, with the cost of living held constant, a change in the distribution of the earnings in a metro area affects one's relative consumption, but it does not affect one's absolute consumption.

We asked two questions about cost of living (one for each metro area) and two questions about the earnings rank (one for each city), in that order. For the cost of living question, we provided the following brief introduction: "You probably noticed that the average prices of goods and services are different across different cities. As a result, with the same income, you would be able to buy more things in some cities and less in other cities." After this introduction, we asked individual how much more or less expensive each metro area was, relative to the U.S. average. To make answering the question easier, we split it in two

[5]When individuals were listing the second program, we required respondents to make a selection from a different metro area because otherwise no differences would be present in absolute and relative consumption across choices. Our survey data indicates that no more than 4% of individuals tried to select the same metro area. For those subjects, the comparison was between two of their top programs but not necessarily the top two.

[6]Even though there are no large income differences in residency salaries, there can be large differences in post-residency salaries, especially across specialties.

questions. The first question was: “Imagine that you chose to work in the [Metro Name] metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?” The respondents could choose either “cheaper” or “more expensive.” The second part of the question was: “How much [cheaper/more expensive] is the [Metro Name] metro area than the U.S. average?” Respondents could answer this second question with a drop-down menu ranging from 0% to 50%, in 1 percentage point increments.

We also provided an introduction for the question about earnings rank: “Now we want to ask you about your expected earnings rank. This rank is defined as the share of the working individuals of a city who earn less than you. You probably noticed that the distribution of earnings is different across different cities. As a result, with the same earnings, you may be relatively rich in some cities but relatively poor in other cities.” After this introduction, we asked the following question for each city: “Imagine that you chose to work in [Metro Name]. With your individual annual earnings of $[Salary], you would be richer than what percentage of [Metro Name]'s individual earners?” Respondents could select their answer from a drop-down menu that ranges from “Richer than 1% of individual earners” to “Richer than 100% of individual earners,” in 1% increments.

2.5 Information-Provision Experiment

One limitation with using perceptions is the potential for omitted-variable bias. For instance, conditional on income and perceptions about cost of living, perceptions about relative consumption may happen to be correlated with perceptions about other characteristics of the area, such as the crime rate, amenities, public goods, and so forth. To address this concern, we generate exogenous variation in the perceptions about cost of living and earnings rank by embedding an information-provision experiment in the survey.

Immediately after respondents provided their prior beliefs on both measures, they were shown two messages: one page with statistics about the cost of living in the two cities being considered and a second page with statistics about the earnings rank in each of the two cities. The following message is a sample of the feedback page for cost of living: “Los Angeles-Long Beach-Anaheim, CA metro area is 17.0% more expensive than the U.S. average. The Champaign-Urbana, IL metro area is 6.6% cheaper than the U.S. average.” The following message is a sample of the feedback page for earnings rank: “With your individual annual earnings of $62,000, you would be richer than 64.6% of Los Angeles-Long Beach-Anaheim, CA's population. With your individual annual earnings of $60,000, you would be richer than 61.2% of Champaign-Urbana, IL's population.” In both of these feedback pages, individuals were asked to take a moment to review the information carefully and were alerted that the information was only going to be shown once. We did not allow respondents to continue to the next page until at least 10 seconds had elapsed.[7]

After individuals finished reviewing the feedback, we re-elicited their perceptions about cost of living and earnings rank, which we denote as the posterior beliefs. Given that our feedback entailed many figures for participants to remember and process, we wanted to make it easier for individuals to compare the options. Therefore, after eliciting respondent's posterior beliefs, we gave subjects a third page of feedback based on their posterior beliefs. The following is a sample of that feedback page: "We understand this is a lot of information to process, so we will help you make the comparison simpler. According to your final answers about incomes, cost of living and earnings rank: If you chose to live in Los Angeles-Long Beach-Anaheim, CA, you would be able to afford 17.9% less than if you chose to live in Champaign-Urbana, IL. If you chose to live in Los Angeles-Long Beach-Anaheim, CA, your earnings rank would be 6.6% higher than if you chose to live in Champaign-Urbana, IL."[8]

We computed the statistics shown to the subjects using two alternative data sources, and we cross-randomized which of the two sources were shown to each individual. The sources were randomized between individuals; that is, we used the same cost of living source for the two cities being considered by each individual, and the same earnings data source for the two cities. As a result, individuals were randomly assigned to one of four treatment groups. For cost of living estimates, the two sources used were the Regional Price Parity (RPP) data by the Bureau of Economic Analysis and the Cost of Living Index (COLI) data compiled by the Council for Community and Economic Research. For the earnings rank feedback, the two sources used were the American Community Survey (ACS) and the Current Population Survey, both conducted by the U.S. Census Bureau.[9]

This source randomization created a substantial amount of exogenous variation in signals.[10] For instance, the correlation of the pairwise difference in cost of living shown to the respondents versus the corresponding pairwise difference from the alternative source is 0.656; the corresponding correlation for the earnings rank is 0.649. These differences across sources arise from a combination of several factors, most notably sampling variation and data definitions. For instance, the cost of living data is subject to sampling variation because it tracks the prices of a limited number of goods and services, and earnings rank data is subject to sampling variation because the estimates are based on a limited number of survey respon-

[7]The median length of time spent in the feedback page was 18.5 seconds.

[8]The difference in absolute consumption was calculated as $100 \cdot \left(\frac{w_1}{w_2}\frac{COL_2}{COL_1} - 1\right)$, where w_i is the nominal wage for city i and COL_i is their posterior belief about cost of living (from 50 to 150). The difference in earnings rank was calculated as $100 \cdot \left(\frac{ER_1}{ER_2} - 1\right)$, where ER_i is the posterior belief about earnings rank in city i. As with the other feedback pages, 10 seconds had to elapse before respondents could move to the next page. The median duration on the post feedback page was 19.5 seconds.

[9]For more details, see Appendix C.

[10]For more details, see Appendix C.

dents. The variation in definitions arise because different cost of living indices give different weights to expenditure categories, and because the earnings rank measures are based on surveys with significant differences in the survey method and the phrasing of the questions used to elicit total annual earnings.

For the sake of transparency and to ensure the validity of the information, the individuals were debriefed in the feedback messages on the name of the source of the information that they received. We would not expect the source name to have an effect in and of itself, given that the individuals did not have expertise on the data, and even experts may have only a weak preference on which source is more trustworthy. Indeed, we find that the reaction of individuals to the information was orthogonal to the name of the information source.[11]

2.6 Rank Submission Choices

The survey asked respondents to indicate which program they expected to rank higher in the NRMP submission: "As of this moment: of the two programs discussed so far, which one would you expect to rank higher for the NRMP?" Individuals could indicate their ranking on a 6-point scale ranging from "Very likely [Program 1] (in [Metro 1])" on one side to "Very likely [Program 2] (in [Metro 2])" on the other. In the baseline results we look at the binary choice of whether they expect to rank Program 1 over Program 2 because a comparison with the ex post submission choices is more straightforward. Nevertheless, results are similar when using the full likelihood scale.[12]

The algorithm used by NRMP was designed by Roth and Peranson (1999) to be 100% resistant to attempts of "strategic behavior," meaning that it is a weakly dominant strategy for students to submit their true preferences (i.e., it is optimal regardless of the behavior of the other applicants). Students receive training from the NRMP that makes it explicit that it is in their best interest to submit truthful ranks. Indeed, survey data indicates that only 5% of participants attempt to misreport their true preferences with a strategic motive (Benjamin et al. 2014, Rees-Jones 2017).[13] Furthermore, almost all NRMP participants receive a match,[14] and backing out from a match entails serious sanctions.[15] As a result, it

[11] Results reported in Appendix D.3.

[12] Results reported in Appendix D.6.

[13] These results are consistent with other surveys (NRMP 2015). Given the small share of individuals attempting to manipulate rankings, we decided not to include questions about this. Relatedly, Rees-Jones and Skowronek (2017) provide complementary behavioral evidence that NRMP participants may fail to fully optimize, including a discussion of the source of those frictions.

[14] For instance, in 2017, 95% of the 27,048 U.S. graduating medical students received a successful match.

[15] For example, applicants with confirmed violations of NRMP policies are subject to a one year bar from accepting or starting a position in any program sponsored by a Match-participating institution, from one year to a lifetime bar from participation in future NRMP Matches, and from one year to a lifetime identification in the matching system as a match violator (Source: http://www.nrmp.org/policies/the-match-commitment/).

is a good approximation to assume that the rank choices are direct representations of the individuals' true preferences.

Since most of the evidence on relative concerns is based on the happiness literature (e.g., Luttmer 2005), we want to compare preferences inferred from choice data with respect to the preferences inferred from happiness data, in the spirit of Benjamin et al. (2012, 2014). For this purpose, we included the following question about happiness rank: "If assigned to it, in which of the two programs would you expect to live a happier life?" Responses used the same likelihood scale as for rank.

2.7 Follow-Up Survey

Shortly after the NRMP rank submission window closed, we conducted a follow-up survey with the subjects that responded to the baseline survey. Appendix A.2 shows the full questionnaire of the follow-up survey.

Most importantly, we collected data on the final rank orders submitted to the algorithm, at the very beginning of the survey. Additionally, we took the opportunity to ask individuals for some additional information. We elicited the perceptions about cost of living and earnings rank, which allows us to measure the persistence of the information learned in the information-provision experiment. Also, we measured additional characteristics of the subjects, such as the places where they grew up and measures of materialism (Richins and Dawson 1992) and competitiveness (Smither and Houston 1992). We did not measure these secondary characteristics in the baseline survey due to space and time constraints.

3 Econometric Model

3.1 Baseline Model

With data about perceived cost of living, perceived earnings rank and rank submission we can estimate preferences over absolute and relative consumption. In this baseline model, we exploit all the variation in perceptions, which includes the experimental variation induced by our information-provision as well as the remaining non-experimental variation.

Let i index subjects and $j \in \{1, 2\}$ denote the two programs being considered by the subject. We define $ER_j^{i,posterior}$ and $COL_j^{i,posterior}$ as the posterior beliefs for earnings rank and cost of living for program j in the baseline survey. Let $ER_{1,2}^{i,posterior} = ER_1^{i,posterior} - ER_2^{i,posterior}$ be the perceived difference in earnings rank between the two programs. Similarly,

Additionally, the NRMP has established rules prohibiting programs from contacting candidates to ask or coordinate their rank orders.

let $COL_{1,2}^{i,posterior} = COL_1^{i,posterior} - COL_2^{i,posterior}$ be the perceived difference in cost of living between the two programs. Let $Program_1 \succ_i \; Program_2$ denote that individual ranks program 1 over program 2, and let $I(\cdot)$ be an indicator function. The regression specification is:

$$I(Program_1 \succ_i \; Program_2) = I\Big(\beta^{rel} \cdot ER_{1,2}^{i,posterior} - \beta^{abs} \cdot COL_{1,2}^{i,posterior} + \\ + \theta X^i + \varepsilon_i \geq 0\Big), \quad (1)$$

where X^i is a vector of control variables and θ is the corresponding vector of coefficients. We always include a constant and the log-difference of nominal residency wages as control variables. In the baseline specification, we include an additional set of controls consisting of pairwise differences in some residency and location characteristics: residency program rank (from Doximity), quality of life inferred from compensating differentials (Albouy 2016), population size, population density, share of African-American residents, and share of Democrat residents.[16] In any case, we present results with alternative sets of control variables. Last, ε_i is the error term, which in the Probit specification is assumed to be normally distributed.

The two key parameters of interest are β^{rel} and β^{abs}. Conditional on nominal wages and $COL_{1,2}^{i,posterior}$, a higher $ER_{1,2}^{i,posterior}$ increases the individual's relative consumption in option 1 relative to option 2, while leaving absolute consumption unchanged. Thus, the parameter β^{rel} measures preferences for relative consumption during the duration of the residency. Depending on the mechanism at play, we may expect β^{rel} to be positive or negative. For instance, the status models predict that $\beta^{rel} > 0$: i.e., individuals want to be higher up in the distribution of consumption or, equivalently, they want to choose poorer ponds.

Conditional on nominal wages and $ER_{1,2}^{i,posterior}$, a higher $COL_{1,2}^{i,posterior}$ decreases the individual's absolute consumption in option 1 relative to option 2, while leaving the relative consumption unchanged. On the other hand, a higher value of $COL_{1,2}^{i,posterior}$ implies lower absolute consumption in option 1 relative to option 2, while leaving relative consumption unchanged. Note that the model specification lets β^{abs} multiply $-COL_{1,2}^{i,posterior}$ instead of $COL_{1,2}^{i,posterior}$. As a result, the coefficient β^{abs} measures preferences for absolute consumption during the residency. We expect $\beta^{abs} > 0$: i.e., individuals prefer higher absolute consumption.

β^{rel} and β^{abs} correspond to preferences over relative and absolute consumption during the residency. Relative and absolute consumption *after* the end of the residency would be part of the error term.[17] The duration of a residency depends on the specialty, ranging from three

[16]The source for the demographic characteristics is the 2011-2014 American Community Survey. For the share of Democrat residents, we use the share of Obama voters between all voters in the 2008 Presidential Elections.

[17]For a minority of subjects who may expect to continue living in the same city after the residency, the

to seven years.

3.2 Instrumental Variables Model

The second model exploits the variation in beliefs induced by the source-randomization experiment to estimate the causal effects of perceptions on choice. Let $ER_{1,2}^{i,shown}$ be the information randomly chosen to be shown to the individual, and $ER_{1,2}^{i,alt}$ be the alternative information that could have been shown to the individual but was not chosen. Let $\Delta ER_{1,2}^{i} = ER_{1,2}^{i,shown} - ER_{1,2}^{i,alt}$ be the difference between the information shown and the alternative information that could have been shown. We estimate an IV-Probit model that uses $\Delta ER_{1,2}^{i}$ and $\Delta COL_{1,2}^{i}$ as instrumental variables. In other words, this model uses the variation introduced by the random assignment of sources to estimate the effect of perceptions on choice:

$$
\begin{aligned}
I(Program_1 \succ_i Program_2) = I(\beta^{rel} \cdot ER_{1,2}^{i,posterior} &- \beta^{abs} \cdot COL_{1,2}^{i,posterior} \\
&+ \lambda_1 \cdot ER_{1,2}^{i,alt} + \lambda_2 \cdot COL_{1,2}^{i,alt} + \theta X^i + \varepsilon_i \geq 0)
\end{aligned}
$$

$$
ER_{1,2}^{i,posterior} = \gamma_1^{ER} \cdot \Delta ER_{1,2}^{i} + \gamma_2^{ER} \cdot \Delta COL_{1,2}^{i} + \gamma_3^{ER} \cdot ER_{1,2}^{i,alt} + \gamma_4^{ER} \cdot COL_{2,1}^{i,alt} + \gamma_5^{ER} X^i + \epsilon_{1,i}
$$

$$
COL_{1,2}^{i,posterior} = \gamma_1^{COL} \cdot \Delta ER_{1,2}^{i} + \gamma_2^{COL} \cdot \Delta COL_{1,2}^{i} + \gamma_3^{COL} \cdot ER_{1,2}^{i,alt} + \gamma_4^{COL} \cdot COL_{2,1}^{i,alt} + \gamma_5^{COL} X^i + \epsilon_{2,i}
$$

There is a simple way to understand the intuition behind this instrumental variables approach. In a deceptive design, we would have shown subjects the statistic from a certain source, add a random noise to this statistic, and then only exploit the variation in beliefs generated by the random noise. In our context, $\Delta ER_{1,2}^{i}$ and $\Delta COL_{1,2}^{i}$ play the role of the random noise added to the feedback, only that they are generated in a non-deceptive manner.

4 Implementation Details and Summary Statistics

Our recruitment strategy is similar to that of Benjamin et al. (2014). During December 2016 we contacted the Associate Dean of Student Affairs at all 135 accredited medical schools in the United States by email to ask for permission to invite fourth year students participating in the 2017 Main Residency Match to take part in our study – a sample of the invitation email is shown in Appendix A.3. Our goal was to recruit as many respondents as possible, so we followed up, by email and phone, with all the deans who showed interest. Of the 79 schools that answered our invitation, 27 agreed to participate. The main reason given by the

cost of living and earnings ranking may also be relevant for post-residency consumption.

schools that declined to participate was school policy restricting external surveys, in place to avoid survey fatigue. Our sample of participating schools includes 22 of the 50 U.S. states, and it is quite representative of the whole sample of 135 accredited medical schools–we do not find statistically significant differences in observable characteristics such as total enrollment, average MCAT scores, undergraduate GPA at admission, acceptance rate, and U.S. News rank.[18]

For confidentiality reasons, we were not given email lists to directly invite students to participate in our study. Instead, the deans agreed to forward our invitation email containing the link to the survey to eligible students (i.e., senior medical students participating in the NRMP). This email invitation, a sample of which is shown in Appendix A.4, asked students to participate in a confidential survey about the Main Residency Match for a study on how medical students select residency programs. The message mentioned that the survey would take less than 10 minutes to complete and respondents would be sent a $10 Amazon gift card by email as a token of appreciation. Finally, the email stressed the eligibility criteria for participating in the survey: being a graduating medical student participating in the Main Residency Match who has not yet submitted his or her rank to the NRMP.[19]

The only reason why we excluded individuals who had previously submitted their ranks was because we wanted individuals who were still deciding and thus prone to using the signals from the information-provision experiment. However, this concern is not important in the sense that submissions can be modified anytime before February 22. Even if some students had already submitted their rank at the time of responding to the survey, they would still be able to log back in and modify their rank. In any case, the vast majority of our subjects responded to the baseline survey quite early in the submission period.

We took several measures to minimize the chance that non eligible students would participate in the survey. First, deans were asked to carefully forward the invitation to senior students participating in the Main Residency Match. This request was not an issue since such a mailing lists already existed; targeted announcements were already being sent to this group during the semester regarding the Match. Second, individuals were reminded of these restrictions in the invitation email and on the consent page of the survey. Third, the first questions of the surveys acted as filters; we asked what match the respondent was participating in and whether they had already submitted their ranks. If they responded with a match other than the Main Residency or "yes" to already submitting their rank, the survey ended there, and they were excluded from taking the survey again.[20]

[18]For details, see Appendix B.

[19]There are a number of alternative matches for some specialties that have different deadlines than the Main Residency Match.

[20]The survey platform blocks users from taking the survey again by using their I.P. address and cookies,

Last, at the end of the survey, respondents were required to submit their university email address to "sign" a statement claiming that they were medical students participating in the NRMP and they understood that we reserved the right to verify their status before making a payment. We asked some schools to verify the list of survey respondents, and they confirmed the validity of 100% of the respondents. Given all the measures taken and the evidence obtained, we are confident that the survey data are of high quality.

The invitation emails were sent to students in a staggered way, with the first round of invitations sent on January 6, 2017, and the last round of invitations and reminders sent on February 7, 2017. We estimated that the student invitations were forwarded to around 3,676 students in total, with 1,080 finishing the baseline survey, implying an overall response rate of 29.38%. The median survey completion time was almost 9 minutes. At the end of the baseline survey we included an attention check question that was passed by 96.4% of respondents. For the sake of transparency, we do not drop the group that did not pass the attention check – indeed, we do not drop any other group from the baseline sample.[21]

On February 23, 2017, the day after the NRMP rank submission deadline, we sent respondents who participated in our baseline survey an invitation to participate in the follow-up survey. We offered participants an additional $5 Amazon gift card for participating in this shorter follow-up survey. We closed the follow-up survey on March 12, one day before Match Week started (i.e., the time when the students find out where they are matched). The response rate to the follow-up survey was 90.62%. Moreover, the characteristics of the individuals who responded to the follow-up survey are similar to the characteristics of individuals who did not respond to the follow-up.[22]

Figure 1 presents the distribution of dates when subjects responded to the baseline survey, when they responded to the follow-up survey, and when they submitted their ranks to the NRMP (for those who provided this information in the follow-up survey). On average, students responded to the baseline survey 24.5 days (sd 12.9) before submitting their ranks, and responded to the follow-up survey 13.9 days (sd 11.8) after submitting their ranks.

Figure 2 shows the geographic distribution of the metropolitan areas in which students' top-two programs are located. This figure shows that there is a broad geographical coverage of the U.S. territory.

Table 1 provides descriptive statistics for the key variables used in the analysis. Column (1) corresponds to all respondents to the baseline survey. Around 48% of respondents were

although students could circumvent this restriction by opening the survey link from a different device.

[21]Upon inspection of the data, the 3.6% of respondents who failed the attention check seem to have answered the survey as consistently as everyone else. And, as reported in Appendix Table D.5, the results are virtually the same if we drop this 3.6% of the sample.

[22]Results presented in Appendix Table B.

male, the average age was 27 years, 35.33% of respondents were single, 23.92% were married, and 40.75% were in a long-term relationship. On average, students were offered a salary of $54,000 for the first year of their residency – this salary would make them richer than 56% of earners in the average metro area. Of course, this sample is not representative of the general U.S. population of adults: most notably, our subject pool is younger and more educated. Nevertheless, our subject pool is close to the U.S. average in terms of nominal wages and gender composition.[23]

To verify that the randomization was successful, Table 1 breaks down the descriptive statistics by each treatment group. This table also reports the p-value for the test of the alternative hypothesis that at least one mean is different across the four treatment groups. First, this table shows that the number of respondents was almost identical number of respondents across all groups. Second, this table shows that the differences in individual characteristics are economically small and statistically insignificant across the treatment groups, thus confirming that the random assignment was successful.

5 Results: Distribution of Perceptions and Learning

5.1 Variation in Nominal Income, Cost of Living and Earnings Rank

We can check whether there is enough variation in cost of living and earnings rank to allow for the estimation of the key parameters. Given that we do not observe the "true" cost of living or earnings rank, but imperfect estimates based on different data sources, the following results use our "baseline" estimates: the RPP measure of cost of living and the ACS measure of earnings rank (using the alternative data sources yield similar results).

Figure 3 shows a scatterplot of the pairwise differences in cost of living vs. the differences in earnings rank. This figure shows three facts that are crucial for estimating the preferences for absolute and relative consumption. First, the substantial dispersion in the x-axis suggests that there are large differences in cost of living across the pairs of cities that the individuals must choose from. Second, the substantial dispersion in the y-axis suggests that there are large differences in earnings rank across the pairs of cities that the individuals must choose from.[24] Third, the $R^2 = 0.22$ indicates that, even though the two are correlated,[25] substantial

[23]For more details, see Appendix Table B.4.

[24]Furthermore, the vast majority of these differences in cost of living and earnings rank are orthogonal to differences in nominal income – see Appendix D for details.

[25]The slope of -0.664 suggests that, on average, more expensive cities tend to have a higher distribution of nominal earnings.

orthogonal variation exists between absolute and relative consumption.

5.2 Distribution of Prior Beliefs

To the best of our knowledge, ours is the first paper to measure perceptions about cost of living and earnings ranks across different U.S. cities. To get a sense of how informed individuals are about these aspects of their decision-making, we start by comparing their prior beliefs (that is, perceptions prior to the feedback) to the baseline statistics: the RPP measure of cost of living and the ACS measure of earnings rank.

Respondents seem to have a relatively good idea of the cost of living in the cities they are considering. Figure 4.a shows respondents' prior beliefs about cost of living along with the corresponding RPP estimates. If answers were completely accurate, we would expect to see all responses on the 45 degree line. On average, prior beliefs overestimate the baseline estimate by just 4 percentage points; and the prior belief and RPP estimates are positively correlated, with an R^2 of 0.550.

However, individuals are substantially less well informed about their earnings ranks. Figure 4.b plots prior beliefs about earnings rank against the ACS estimates. On average, individuals underestimate earnings ranks by almost 16 percentage points; and the prior belief and ACS estimates are positively correlated, but with an R^2 of just 0.029. Because we are ultimately interested in relative differences for their decision making, we repeat this exercise using pairwise differences instead of levels. It seems that respondents have a better understanding of relative differences in earnings rank, though they still remain far less accurate than perceptions over cost of living.[26] This finding suggests that, while prior evidence suggests that individuals have significant biases when assessing their position in the national income distribution (Cruces, Perez-Truglia, and Tetaz 2013; Karadja, Mollerstrom, and Seim 2017), these biases are even more substantial when individuals try to predict their position in places where they are not currently living.

5.3 Learning from Statistics

We next examine whether respondents learned from the information we provided. To do this, we examine the relationship between the initial perception gap of respondents (i.e., the signal received minus the prior belief) and the extent to which they revise their responses (the posterior belief minus the prior belief). If respondents learn from the information provided, we would expect a positive relation between their perception gaps and their revisions; that is,

[26] Detailed results reported in Appendix D.2.

respondents who originally overestimated would revise their beliefs downwards, while those who underestimated would revise in the opposite direction.

Indeed, the slope between the perception gaps and revisions can be used to quantify the degree of learning from information.[27] Let b_k^{prior} denote the mean of the prior belief k, b_k^{signal} the signal about k, and $b_k^{posterior}$ the mean of the corresponding posterior belief. When priors and signals are normally distributed, Bayesian learning implies that the mean of the posterior belief should be a weighted average between the signal and the mean of the prior belief:

$$b_k^{posterior} = \alpha_k \cdot b_k^{signal} + (1 - \alpha_k) \cdot b_k^{prior}$$

The degree of learning can be summarized by the weight parameter α_k. This parameter can take values from 0 (individuals ignore the signal) to 1 (individuals fully adjust to the signal). We can rearrange the previous equation:

$$b_k^{posterior} - b_k^{prior} = \alpha_k \cdot \left(b_k^{signal} - b_k^{prior}\right)$$

Which implies that we can estimate the learning rate (α_k) by estimating a regression of the revision ($b_k^{posterior} - b_k^{prior}$) on the perception gap between the prior and the signal ($b_k^{signal} - b_k^{prior}$).

Respondents strongly updated their beliefs after being provided with feedback. Figures 5 and 6 present the reduced-form effects of information for cost of living and earnings rank, respectively. Figures 5.a and 6.a present the short-term effect, that is, the revision made by respondents directly after being given the information. The short-term learning rates, given by the slopes reported in these figures, are 0.879 (se 0.010) for the cost of living and 0.873 (se 0.011) for the earnings rank. These two learning rates are statistically significant, precisely estimated, and we cannot reject the null hypothesis that they are equal to each other (p-value 0.754). These learning rates are remarkably close to 1, meaning that respondents almost fully reacted to the signals.

One limitation with this evidence is that individuals may have revised their beliefs towards the truth regardless of the feedback we provided. For instance, they may have taken extra time to think about the question, leading to a more accurate response. The source experiment was designed to test this specific hypothesis. We construct two variables: the information actually shown and the "alternative" information that could have been shown. If the alternative information had any effect beyond the information shown, that would be evidence that part of the revisions were due to reversion to the truth rather than reversion

[27] See for example Armantier, Nelson, Topa, van der Klaauw and Zafar (2016) and Cavallo, Cruces and Perez-Truglia (2017).

to the information provided. Figures 5.b and 6.b show the relation between the alternative information and the revision adjusted for the information actually shown. The alternative information indeed has no effect: the coefficients are close to zero (-0.006 for cost of living and 0.001 for earnings ranking) and precisely estimated.

Another concern is that respondents may use the feedback on cost of living to update beliefs about earnings rank, or vice versa. To examine this hypothesis, Figures 5.c and 6.c show the relation between the perception gap for earnings rank (cost of living) and the revision for cost of living (earnings rank). These spillovers are close to zero (-0.015 for cost of living and -0.019 for earnings ranking), statistically insignificant, and precisely estimated.

In survey experiments, one main concern is that instead of inducing genuine learning, the information provided in the experiment may elicit spurious reactions. For instance, if an individual is told that the cost of living in a city is "10% more expensive than the U.S. average" and then later asked about the cost of living in the same city, he or she may report a cost of living that is closer to "10% more expensive than the U.S. average" for spurious reasons, such as unconscious numerical anchoring (Kahneman and Tversky 1972). Under the assumption that these effects are temporary, we can disentangle genuine from spurious learning by looking at the reaction to the information provided in the experiment that persisted over time (Cavallo, Cruces and Perez-Truglia 2017).

We look at the persistence of the effect of feedback between the time participants responded to the baseline and follow-up surveys, which was 38.4 days on average. Figures 5.d and 6.d show the relation between the initial perception gap and the long-term revision based on beliefs reported in the follow-up survey (i.e., $b_k^{posterior,LT} - b_k^{prior}$). There is substantial persistence the effects of the feedback: the estimated slope for the initial perception gap and the long-term revision (i.e., the difference between long-term belief and the initial prior belief) for cost of living is 0.752 (se 0.016), while for earnings rank it is 0.626 (se 0.020). These longer-term revisions are slightly weaker than the short-term revisions, but that result is expected given that individuals must have gathered some additional information in the time between the two surveys.

6 Results: Preferences for Relative Consumption

6.1 Average Preferences

We first explore the baseline estimates of the effects of relative and absolute consumption. The baseline specification uses the Probit model from Section 3, with the expected rank submission as dependent variable. This specification exploits all the variation in perceptions, which includes the experimental variation induced by our information-provision as well as the remaining non-experimental variation. We introduce the experimental estimates later in this section.

Respondents prefer higher absolute consumption. The Probit coefficients are presented in Table 2. Column (1) presents the results for the full sample, while columns (2) through (7) present results by demographic subgroups. The estimated β^{abs} from column (1) is positive and statistically significant (p-value=0.026), suggesting that the average individual prefers programs with higher absolute consumption. To better understand the magnitude of these Probit coefficients, we can transform them into the corresponding marginal effects.[28] According to that metric, increasing absolute consumption by 1 percentage point at a program's location increases the probability of choosing that program by 0.202 percentage points (which can be interpreted as a behavioral elasticity of 0.202).

The fact that medical students care about consumption during the residency is consistent with the view that money is a primary motivation for doctors. For instance, according to a 2008 survey, 49% of pre-med students self-reported being primarily motivated by money in their career choice, in comparison to 71% of pre-law students (Daniel and O'Brien 2008).[29] Even though β^{abs} is statistically and economically significant, that does not imply that absolute consumption during residency is the main feature that medical students pay attention to. Intuitively, if doctors care about their post-residency consumption, they should choose residencies that offer better post-residency job prospects. Indeed, in complementary analysis, we find that doctors care substantially more about the prestige and career prospects than about the absolute consumption during their residency.[30]

Most important, respondents also prefer higher relative consumption on average: the estimated β^{rel} from column (1) is positive and statistically significant (p-value=0.067). This coefficient suggests that the average individual prefers to live in a city where, holding her absolute consumption constant, she gets to consume more than her neighbors. The corresponding marginal effect indicates that increasing the relative consumption at a program's location by 1 percentage point increases the probability of choosing that program by 0.185 percentage points (for a behavioral elasticity of 0.185). The elasticity for absolute consumption (0.202) is similar in magnitude to the elasticity for relative consumption (0.185) – indeed, their difference is statistically insignificant. This finding suggests that relative consumption may be as important as absolute consumption in driving location choices.[31]

[28]These marginal effects are reported in Appendix Table D.1.

[29]This survey results are based on responses from 461 takers of the Kaplan MCAT test in February 2008 and 453 takers of the Kaplan LSAT test in February 2008.

[30]Results presented in Appendix D.7.

[31]The ratio $\frac{\beta^{relative}}{\beta^{absolute}} = 0.916$ represents the marginal rate of substitution between relative and absolute

The evidence suggests that individuals take relative consumption into account when they make their location decisions. This preference could be the product of a combination of multiple mechanisms, pushing in different directions. The positive sign of β^{rel} suggests that the dominant mechanism is consistent with models in which richer neighbors impose a negative externality, as in Boskin and Sheshinski (1978), Frank (1985a), Cole et al. (1992) and Luttmer (2005), among others. For now, we focus on identifying the preference parameters, and we provide a discussion of the mechanisms in Section 6.6 below.

6.2 Heterogeneity by Relationship Status

The average preferences could potentially mask substantial heterogeneity. For instance, Luttmer (2005) finds that the effect of relative income on happiness is driven entirely by the sample of non-single individuals. Furthermore, evidence from the urban economics literature indicates that single and non-single individuals have different locational preferences (e.g., Couture and Handbury 2015). To explore heterogeneity in preferences, columns (2) through (7) of Table 2 present estimates broken down by the basic demographic groups measured in the baseline survey: relationship status, gender and (post-residency) income.

To explore heterogeneity by relationship status, we elicited the relationship status using the same categories as in Luttmer (2005). Column (2) of Table 2 shows the effect for non-single individuals (i.e., the 65% of the sample who are married or in a long-term relationship) and column (3) for the sample of single individuals (35% of the sample).[32] It is important to note that by non-single we only refer to their relationship status, not to whether the respondent participates in a dual match, which is a special regime used by roughly 7% of subjects—indeed, the results are similar if we drop subjects with dual matches.[33]

Comparing columns (2) and (3) indicate large differences in β^{rel} with respect to relationship status. For non-single individuals, the estimated β^{rel} (2.195) is positive and statistically significant at the 1% level. For the sample of single individuals, β^{rel} (-1.527) is negative and statistically significant at the 10% level. The direction of the difference in relative concerns between singles and non-singles is consistent with the evidence from Luttmer (2005).

The difference in β^{rel} between non-singles and singles is highly statistically significant (p-value=0.001). Moreover, to address spurious results from multiple hypothesis testing,

consumption. This estimated value indicates that the average individual would be indifferent between an increase of 0.916 percentage points in absolute consumption or a 1 percentage points increase in relative consumption. We must take this ratio with a grain of salt, however, because it is imprecisely estimated: the 90% confidence interval is [-0.198, 2.029].

[32] Appendix Table D.4 report results breaking down the non-single individuals into married and in a long-term relationship. The relative concerns are similar between these two groups.

[33] See Appendix Table D.5 for more details.

for each p-value reported in the table we also report the corresponding q-value based on Benjamini and Yekutieli (2001). The q-value indicates the minimum false discovery rate (i.e., the expected proportion of rejected null hypotheses that are actually true) at which the null hypothesis would be rejected for that test given all tests reported in the same table. The difference in β^{rel} between singles and non-singles has a q-value of 0.038, which indicates that this heterogeneity is unlikely to be spurious.

Contrary to the case of preferences for relative consumption, the relationship status does not seem to affect the preferences for absolute consumption. According to columns (2) and (3) of Table 2, the estimated β^{abs} is 1.095 for non-singles and 1.042 for singles, with the difference being statistically insignificant (p-value=0.957).

These estimates suggest that while non-single individuals prefer to live in poorer ponds, single individuals would rather live in richer ponds. While the preferences of non-single individuals can be rationalized by status models, the preferences of single individuals cannot be rationalized by such models. One potential explanation for the preferences of single individuals lies in their local social interactions, such as in the dating market. The single individuals in our subject pool are at their prime dating age, and thus they are likely to be looking for long-term partners during the duration of the residency. Given the evidence that individuals prefer marrying richer individuals (Hitsch et al. 2010), these single individuals may prefer richer ponds because of the more affluent pool of partners.[34] Although these individuals make close to the median U.S. salary during the duration of their residency, their permanent incomes will be at the top of the U.S. income distribution once they finish the residency. Thus, it may be specially important to choose a rich pond if they are looking to find partners who can match these high levels of post-residency incomes.

Gender is also a potential driver for heterogeneity in relative concerns. For example, Fisman, Iyengar, Kamenica and Simonson (2006) and Bertrand, Kamenica and Pan (2015) find significant gender differences in preferences for ambition and income in a partner. Columns (4) and (5) present heterogeneity by gender. The gender differences are small: β^{rel} is similar for females (1.034) and males (0.898), and β^{abs} is also similar for females (0.961) and males (1.440). Moreover, neither of these two differences are statistically significant (p-values of 0.635 and 0.900, respectively).

Last, even though all these subjects receive a similar income during the duration of the residency, they can get substantially different incomes after they finish their residencies. It is possible than individuals who selected high-earning specialties may be more concerned about

[34] Additionally, Appendix D.3 presents suggestive evidence that this negative effect of relative consumption among the single individuals may be disproportionally driven by single women, which is consistent with prior evidence that a higher relative income is detrimental to the marriage prospects of women (Bertrand, Kamenica and Pan 2015; Bursztyn, Fujiwara and Pallais 2017).

absolute or relative consumption. The test this hypothesis, columns (6) and (7) split the sample in specialties with above and below median post-residency average salaries. Again, the differences in coefficients are statistically insignificant: β^{rel} is 1.411 for below-median specialties and 0.780 for above-median specialties, and β^{abs} is 0.691 for below-median specialties and 1.232 for above-median specialties, with neither of those differences being statistically significant (p-values of 0.559 and 0.580, respectively).

We also computed heterogeneity by other characteristics measured in the follow-up survey. None of these dimensions are nearly as important as relationship status for predicting heterogeneity in preferences for relative consumption, both in terms of magnitude and statistical significance.[35] Because of the magnitude of the heterogeneity by relationship status, in the remainder of the paper, we report estimates for non-single and single respondents, in addition to estimates for the entire sample.

6.3 Comparison to Estimates from Happiness Literature

Assuming that the mechanisms have to do with externalities from neighbors, as in Luttmer (2005), it is then useful to compare the magnitude of our coefficients to those from Luttmer (2005).[36] Luttmer's main specification, which is estimated on the sample of non-single individuals, imply that most of the utility from consumption goes through relative consumption rather than through absolute consumption: non-single individuals would be willing to give up 1 percent of absolute consumption in order to decrease the median consumption of neighbors by 0.22%.[37] According to our baseline estimates from column (2) of Table 2, non-single individuals are willing to give up 1 percent of their absolute consumption in order to decrease the median consumption of their peers by 0.88% (90% confidence interval: [−0.36%, 2.14%]).[38] Relative to Luttmer (2005), our baseline estimates suggest a weaker role for relative concerns; however, this difference is not statistically significant.[39] If we assumed that Luttmer (2005) measures the true extent to which people care about relative concerns, our estimates would

[35] Results reported in Appendix Table D.4.

[36] We focus on Luttmer (2005) because it uses data for the United States and is then the most comparable sample. The results are similar when we compare our estimates to estimates from other papers using subjective data (reported in Appendix D.12).

[37] Appendix D.12 provides the details for this calculation.

[38] For the average individual in the sample, we would need to decrease the median earnings in the area by 0.88% to allow the individual to climb up 0.50 (= 1/2.004) percentage points in the earnings rank.

[39] Of course, part of the difference may be due to differences in the subject pools: i.e., senior medical students having weaker preferences for relative concerns than the average U.S. resident. Also, any differences in the trade-offs measures with happiness and choice data would not imply that one of the two results are wrong: e.g., it is possible that the happiness estimates reflect the true extent to which people care about relative concerns, but when deciding where to live, individuals under-estimate how much their well-being will depend on relative consumption. Last, we must note that our experimental estimates are closer to the results from Luttmer (2005).

suggest that individuals anticipate, at least partially, the negative externalities from richer neighbors.

For a more direct comparison between happiness and choice data, we can also exploit the survey responses on expected happiness (Benjamin et al. 2014). We find that the marginal rates of substitution inferred by happiness are statistically indistinguishable from the marginal rates of substitution inferred by choice; however, due the lack of precision of the happiness estimates, we cannot rule out large discrepancies.[40]

6.4 Robustness Check: Controlling for Other Observable Characteristics

One potential concern with the baseline specification is that of omitted-variable biases. For instance, if places where an individual expects higher earnings rank (i.e., poorer metro areas) are systematically worse in terms of quality of life, then failing to account for quality of life would introduce a negative bias in β^{rel}, thus making relative concerns look weaker than they actually are.

We present the baseline estimates using alternative sets of control variables in Table 3. Each row corresponds to a different regression, with a different set of control variables. The first row presents results for our baseline specification, but without including any control variables for the characteristics of the program or the metro area. The second row corresponds to the baseline specification from Table 2, which includes the five baseline controls listed in Section 3.1. The results in the first two rows of Table 3 indicate that β^{rel} and β^{abs} are qualitatively and quantitatively similar between the baseline specification and the specification without controls.

The third through last rows of Table 3 include different sets of additional controls. These sets of controls were selected based on attributes that could potentially be relevant for the options of the subjects and at the same time may be correlated to the earnings rank. For instance, we may want to control for crime rates: living in a poorer city may be desirable for medical students interested in certain specialties where they must learn to treat injuries that are more common in high-crime areas, such as gunshot wounds. Also, we may want to account for place of origin: medical students, who tend to grow up in affluent areas, may want to remain in the the same areas where they grew up (Agarwal 2015).

We examine the following groups of attributes: demographic characteristics (population, population density, percentage urban population, percentage same gender, percentage age 25 to 34, share of college graduates, share foreign, share Hispanic, and share black); amenities

[40]Results presented in Appendix Table D.9.

(quality of life from Albouy 2016, per capita spending on local public goods, per capita spending on education and health, overall crime rate and violent crime rate, share of registered Democrat voters in the 2012 election); geography (distance of program to city where they grew up, distance of program to current medical school); economic factors (estimated income taxes, federal and state income taxes, local sales tax, rent prices, and the Gini coefficient); a set of state dummies; objective program characteristics (residency program rank from Doximity, dummies for university hospitals and for community hospitals), and subjective program characteristics (the subjective rank in prestige, purpose, and prospect, as reported in the follow-up survey).

Comparing the results across rows of β^{rel} and β^{abs} of Table 3 suggests that these estimates are robust to the choice of control variables, both in terms of statistical significance and economic significance. Of course, small differences occur in the point estimates across specifications. For instance, relative to the baseline β^{rel} of 1.022, the β^{rel} ranges from a minimum of 0.873 without controls to a maximum of 1.199 with demographic controls.[41] However, all of these differences are statistically insignificant.

6.5 Robustness Checks: Experimental Estimates and Long-Term Effects

In this section, we present results from two robustness checks. The first check addresses concerns about omitted-variable bias by exploiting the exogenous variation in beliefs generated by the source-randomization experiment. The second is intended to address potential concerns about spurious effects of the information-provision experiment, by comparing the short-term effects to the long-term effects. To make these estimates directly comparable to the long-term effects, in this section we restrict the sample to individuals who responded to the follow-up survey.

Panel A of Table 4 presents the results for β^{rel}. The first row presents the baseline specification, while the second row presents the experimental estimates. The experimental estimates are less precisely estimated than the baseline estimates because they only use a portion of the variation in beliefs. For each of the subgroups of single and non-single respondents, shown in columns (2) and (3), the estimated β^{rel} is qualitatively consistent across the baseline and experimental specifications. For non-singles, the coefficient is 2.337 (p-value=0.001) in the baseline specification vs. 2.955 (p-value=0.026) in the experimental

[41] According to the pseudo-R^2 reported in panel C of Table 2, including these variables increases the explanatory power of our model to some degree. For the full sample, the pseudo-R^2 increases from 0.015 in the specification with no additional controls to a minimum of 0.018 with controls for objective program characteristics and a maximum of 0.123 with controls for subjective programs characteristics.

specification. And for singles, the coefficients are -1.666 (p-value=0.094) in the baseline specification vs. -4.984 (p-value=0.011) in the experimental specification.

Column (1) shows that, for the entire sample, β^{rel} is slightly lower in the second row (0.858) than in the first row (1.130) and, due to the lower precision, becomes statistically insignificant in the second row. However, we must take this finding with a grain of salt. First, due to the precision of the experimental coefficient, this difference between the first and second rows is statistically insignificant. Second, the reduction in the average β^{rel} is driven primarily by the fact that the coefficient becomes more negative for singles – in other words, even though the average coefficient becomes slightly smaller, the experimental estimates imply that relative consumption plays a more important role for location choices.

Panel B of Table 4 presents the results for β^{abs}. The results from the baseline specification (first row) are qualitatively different from the results in the experimental specification (second row). All the coefficients (for the entire sample, singles and non-singles) move downwards, become negative, are imprecisely estimated, and are statistically insignificant. We must take this evidence with a grain of salt: since the experimental estimates are not precisely estimated, we cannot rule out large positive values for β^{abs}, and in most cases we cannot reject that the experimental coefficients are equal to those from the baseline specification. Also, the coefficients from the first and second row should not be expected to be equal, to the extent that the experimental coefficients identify local average preferences instead of average preferences.[42] However, these findings are at least suggestive evidence that the baseline estimates may overestimate the importance of absolute consumption.

As discussed above, the treatment groups were balanced in observable characteristics, suggesting that the randomization was indeed successful. As an additional robustness check, we re-estimate the instrumental variables model but, instead of the rank order, we use the list order as dependent variable (i.e., the order in which the individual listed the residency programs at the beginning of the survey). Because it takes place before the provision of feedback, the feedback should not have any effect on the list order. We present results for this falsification in the fourth row of Table 4. As expected, the estimated values of β^{rel} and β^{abs} are close to zero and statistically insignificant, in the full sample as well as in the sub-samples of non-singles and singles.

The Appendix presents some additional results. In all the instrumental variable specifications, we strongly reject the null hypothesis of weak instruments. Also, the learning rates implied by the first-stage coefficients are always close to 1, and for that reason the

[42]For instance, it is plausible that the information-provision disproportionally affected individuals who were the most unsure about their priors beliefs about cost of living, who likely are those who care the least about cost of living.

instrumental variables estimates are similar to the reduced form estimates.[43]

The second robustness test is intended to address potential concerns about spurious effects of the information-provision experiment such as salience and experimenter-demand effects. For example, by asking individuals questions about the cost of living and earnings rank, the baseline survey makes those aspects more salient, which may make individuals overweight them in their expected choice. However, this salience effect may not necessarily exaggerate the importance of relative consumption, because they would be expected to inflate both β^{rel} and β^{abs}.

We should also be concerned about potential experimenter-demand effects: by providing individuals with information about cost of living and earnings rank, the experimenter may be putting pressure on the subjects to use this information in their expected choice. Again, this source of bias would not necessarily exaggerate the importance of relative consumption: since most individuals do not want to reveal to others that they care about status (Shigeoka and Yamada 2016), the experimenter-demand bias would probably shrink β^{rel} towards zero.[44]

To address these remaining concerns, we estimate the effects of the information provision on the final rank submission, which takes place an average of 38.4 days after the information provision. This can be achieved by using the same instrumental variable model, but using the final submission rank (elicited in the follow-up survey) instead of the expected submission rank (elicited in the baseline survey) as the dependent variable. If the effects were spurious due to salience or experimenter-demand effect, we would expect that the information provided in the experiment would not have any effect on the final submission choice. In other words, we test whether the information provided a month before the submission deadline had a long-lasting effect on the final ranks submitted.

The third row of Table 4 presents the experimental estimates based on the long-term effects of the experiment. By comparing the coefficients in the third row to those from the second row, we can compare the short-term and long-term experimental effects. Panel A of Table 4 presents the results for β^{rel}, while panel B corresponds to β^{abs}. The long-term experimental coefficients are somewhat different from the short-term experimental coefficients, but those differences are mostly statistically insignificant. Most important, the coefficient on β^{rel} is still positive (1.946) and statistically significant for non-singles, and negative (-5.279) and statistically significant for singles.

[43]Reduced-form and first-stage estimates are presented in Appendix Table D.7.

[44]Also, our survey was conducted confidentially and online, which reduces the scope for experimenter-demand effects (van Gelder et al. 2010). Additionally, it would be difficult to reconcile the experimenter-demand channel with the finding that the earnings rank had a positive effect on non-singles and a negative effect on singles.

6.6 Discussion

We presented unique evidence that individuals take relative consumption into account when they make their location decisions. However, it is more difficult to disentangle the mechanisms at play. The direction of effects for non-single individuals indicate that preferences for relative consumption could be explained by positional externalities as in Luttmer (2005) and others. Indeed, this is our favorite interpretation of the evidence. On the other hand, these positional externalities could not explain the preferences of single individuals.

There are several possible interpretations for the positional externalities, which are not mutually exclusive. One interpretation of these externalities is that individuals with higher relative consumption are more likely to get non-market goods and services (Cole et al. 1992): e.g., a higher relative consumption may increase one's probability of being invited on a date, a business venture, or to a club. Indeed, there is direct evidence that richer individuals get preferential treatment in a variety of interactions.[45] Another interpretation is that individuals anticipate that their consumption aspirations will be affected by the consumption of peers: e.g., individuals may want to avoid rich neighborhoods because they will need to spend more to meet the consumption expectations of their neighbors (e.g., Frank 1985a). Last, a more extreme interpretation of these positional externalities is that the income of peers enters directly into the utility function: e.g., individuals could get a boost in happiness from looking around and seeing that they are doing better than their neighbors (e.g., Boskin and Sheshinski 1978).

It is also possible that some individuals care indirectly about relative consumption. For instance, if an individual finds out that a particular location is poorer than expected, she may change her rank orderchoice because she may update beliefs about other characteristics of the location such as the quality of public goods and crime rates. However, if individuals become more pessimistic about public good provision and crime rates, they should be less likely to choose that location. As a result, this indirect learning is a potential explanation for the negative coefficient of β^{rel} among singles, but could not explain is not a likely explanation for the positive value of β^{rel} among non-singles. If anything, this mechanism would lead to an under-estimation of the positional externalities for non-singles.

Additionally, there are at least two reasons that make the indirect learning channel unlikely. First, it is unlikely that individuals would need to rely on information about earnings rank to learn about other features of the locations and programs, given that they can learn

[45] Studies show that driving a more expensive car makes other drivers more patient (Doob and Gross 1968), and wearing an expensive shirt makes someone more persuasive (Fennis 2008) and more likely to be recommended for a job (Nelissen and Meijers 2011). Also, there is suggestive evidence that individuals purposefully overspend in highly visible goods to appear richer in the eyes of their peers (Charles et al. 2009; Heffetz 2011; Bursztyn et al. 2017).

from those other aspects directly. After all, these subjects devote time and attention to studying various aspects of the residency programs and their locations, which includes visiting those locations, talking to others and doing research online. Second, we know from the distribution of prior beliefs and from the official statistics that earnings ranks are correlated to the cost of living. However, the evidence presented in section 5 indicates that the information about earnings ranks did not affect posterior beliefs about cost of living. Thus, if the information about earnings ranks did not spill over to a closely-related belief such as cost of living, it is unlikely to have large effects on other beliefs.

7 Conclusions

We presented results from a survey with 1,100 medical students participating in the NRMP. These data provide unique revealed-preference evidence that, when choosing where to live, individuals care about their relative consumption in addition to their absolute consumption. Furthermore, we found that individuals can differ dramatically in their preferences for relative consumption: while non-single individuals want to live in poorer ponds, single individuals prefer to live in richer ponds.

Regarding the external validity of our results, it is possible that doctors have stronger relative concerns than the rest of the population due to the competitive nature of their profession. Also, given that most doctors make well above the subsistence level, they may care about positional externalities to an extent that poor individuals would not. A first avenue for future research is to find other contexts in which this revealed-preference method could be used to estimate preferences for relative consumption. For instance, even though the settings may not be as clear-cut as for the medical residency, there are multiple job markets in which job seekers must choose between job offers in different cities. Using a broader subject pool will help generalize the findings from this study, and will also provide more room to study heterogeneity in preferences.

Future research should also work towards understanding the precise mechanisms why individuals care about relative consumption, such as instrumental and non-instrumental motives.[46] In particular, there is little evidence of whether individuals care mostly about their own perceptions of relative consumption (i.e., self image) or their beliefs about the perceptions of peers (i.e., social image).[47] These additional hypotheses can be explored by using

[46] For instance, Cullen and Pakzad-Hurson (2017) show suggestive evidence that, in the context of an online work platform, concerns for relative wages operate though the instrumental channel.

[47] One exception is Bursztyn et al. (2017) which shows suggestive evidence that, in the context of demand for premium credit cards, at least some of the conspicuous consumption operates through the self-image channel.

the same empirical framework proposed in this paper, but with additional treatment arms designed to test specific mechanisms.

References

[1] Agarwal, N. (2015), "An Empirical Model of the Medical Match," American Economic Review, Vol. 105(7), pp. 1939-1978.

[2] Albouy, D. (2008), "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas," NBER Working Paper No. 14472.

[3] Albouy, D. (2016), "What are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities," Review of Economics and Statistics,98(3) July 2016, pp. 477–487.

[4] Armantier, O.; Nelson, S.; Topa, G.; van der Klaauw, W. and Zafar, B. (2016), "The Price Is Right: Updating of Inflation Expectations in a Randomized Price Information Experiment," Review of Economics and Statistics, Vol. 98 (3), pp. 503-523.

[5] Bénabou, R.J.M. and Tirole, J. (2006), "Incentives and Prosocial Behavior," American Economic Review, Vol. 96 (5), pp. 1652-1678.

[6] Benjamin, D.J.; Heffetz, O.; Kimball, M.S. and Rees-Jones, A. (2012), "What Do You Think Would Make You Happier? What Do You Think You Would Choose?" American Economic Review, Vol. 102 (5), pp. 2083–2110.

[7] Benjamin, D.J.; Heffetz, O.; Kimball, M.S. and Rees-Jones, A. (2014), "Can Marginal Rates of Substitution Be Inferred From Happiness Data? Evidence from Residency Choices." American Economic Review, Vol. 104 (11), pp. 3498-3528.

[8] Benjamini, Y. and Yekutieli, D. (2001), "The control of the false discovery rate in multiple testing under dependency," Annals of Statistics, Vol. 29, pp. 1165-1188.

[9] Boskin, M. and Sheshinski, E. (1978), "Optimal Redistributive Taxation When Individual Welfare Depends on Relative Income," Quarterly Journal of Economics, Vol. 92 (4), pp. 589-601.

[10] Bertrand, M.; Kamenica, E.; and Pan, J. (2015), "Gender Identity and Relative Income Within Households," Quarterly Journal of Economics, Vol. 130(2), pp. 571-614.

[11] Bursztyn, L.; Fujiwara, T. and Pallais, A. (2017), "Acting Wife: Marriage Market Incentives and Labor Market Investments," American Economic Review, forthcoming.

[12] Bursztyn, L.; Ferman, B.; Fiorin, S.; Kanz, M. and Rao, G. (2017), "Status Goods: Experimental Evidence from Platinum Credit Cards," Mimeo.

[13] Bursztyn, L. and Jensen, R. (2015), “How Does Peer Pressure Affect Educational Investments?” Quarterly Journal of Economics, Vol. 130 (3), pp. 1329–1367.

[14] Charles, K.K.; Hurst, E. and Roussanov, N.L. (2009), “Conspicuous consumption and race,” The Quarterly Journal of Economics, Vol. 124 (2), pp. 425-467.

[15] Clark, A.E. and Oswald, A.J. (1996), “Satisfaction and Comparison Income,” Journal of Public Economics, Vol. 61, pp. 359–381.

[16] Clark, A.E.; Frijters, P. and Shields, M.A. (2008), “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles,” Journal of Economic Literature, Vol. 46 (1), pp. 95-144.

[17] Clark, A.E.; Senik, C. and Yamada, K. (2017), “When experienced and decision utility concur: The case of income comparisons,” Journal of Economic Behavior and Organization, Vol. 70, pp. 1-9.

[18] Clark, A.E.; Kristensen, N. and Westergård-Nielsen, N. (2009), “Economic Satisfaction and Income Rank in Small Neighbourhoods,” Journal of the European Economic Association, Vol. 7, pp. 519-527.

[19] Cole, H.L.; Mailath, G.J.; and Postlewaite, A. (1992), “Social norms, savings behavior, and growth,” Journal of Political economy, Vol. 100 (6), pp. 1092–1125.

[20] Couture, V. and Handbury, J. (2015), “Urban Revival in America, 2000 to 2010,” Mimeo.

[21] Cullen, Z.B., and Pakzad-Hurson, B. (2017), “Equilibrium Effects of Pay Transparency,” HBS Working Paper No. 52648.

[22] Cruces, G.; Perez-Truglia, R. and Tetaz, M. (2013), “Biased Perceptions of Income Distribution and Preferences for Redistribution: Evidence from a Survey Experiment,” Journal of Public Economics, Vol. 98, pp. 100-112.

[23] Daniel, C. and O’Brien, M. (2008), “Why Study Medicine?” The Student Doctor Network, April 24, 2008.

[24] Dellavigna, S.; List, J.A.; Malmendier, U.; Rao, G. (2017), “Voting to Tell Others,” Review of Economic Studies, Vol. 84 (1), pp. 143–181.

[25] Doob, A. and Gross, A. (1968), “Status of Frustrator as an Inhibitor of Horn- Honking Responses,” Journal of Social Psychology, Vol. 76, pp. 213-218.

[26] Dubins, L. and Freedman, D. (1981), “Machiavelli and the Gale-Shapley Algorithm,” American Mathematical Monthly, Vol. 88 (7), pp. 485-494.

[27] Duesenberry, J. (1949), "Income, Saving and the Theory of Consumption Behavior," Cambridge, Mass.: Harvard University Press, 1949.

[28] Easterlin, R.A. (1974), "Does economic growth improve the human lot? Some empirical evidence," Nations and households in economic growth, Vol. 89, pp. 89-125.

[29] Fennis, B.M. (2008), "Branded into submission: Brand attributes and hierarchization behavior in same-sex and mixed-sex dyads," Journal of Applied Social Psychology, Vol. 38, pp. 1993-2009.

[30] Ferrer-i-Carbonell, A. (2005), "Income and Well-Being: An Empirical Analysis of the Comparison Income Effect," Journal of Public Economics, Vol. 89, pp. 997–1019.

[31] Fisman, R.J.; Iyengar, S.S.; Kamenica, E.; and Simonson, I. (2006), "Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment," Quarterly Journal of Economics, Vol. ?, pp. 673–697.

[32] Frank, R.H. (1985a), "The Demand for Unobservable and Other Positional Goods," The American Economic Review, Vol. 75, pp. 101-116.

[33] Frank, R.H. (1985b), "Choosing the Right Pond: Human behavior and the quest for status." Oxford University Press, 1985.

[34] Heffetz, O. (2011), "Conspicuous Consumption and Expenditure Visibility: Measurement and Application," Review of Economics and Statistics, Vol. 93 (4), pp. 1101–1117.

[35] Hitsch, G.; Hortacsu, A. and Ariely, D. (2010), "What Makes You Click: An Empirical Analysis of Online Dating," Quantitative Marketing and Economics, Vol. 8 (4), pp. 393-427.

[36] Holmes, T. J. and Sieg, H. (2015), "Structural Estimation in Urban Economics," Handbook of Regional and Urban Economics, Vol. 5, Ch. 2, pp. 69-114.

[37] Johansson-Stenman, O.; Carlsson, F. and Daruvala, D. (2002), "Measuring future grandparents' preferences for equality and relative standing," Economic Journal, Vol. 112, pp. 362–383.

[38] Kahneman, D. and Tversky, A. (1972), "Subjective probability: A judgment of representativeness," Cognitive Psychology, Vol. 3 (3), pp. 430–454.

[39] Karadja, M.; Mollerstrom, J. and Seim, D. (2017), "Richer (and Holier) than Thou? The Effect of Relative Income Improvements on Demand for Redistribution," Review of Economics and Statistics, forthcoming.

[40] Kuziemko, I.; Buell, R.W.; Reich, T. and Norton, M. (2014), "Last-place Aversion: Evidence and Redistributive Implications," Quarterly Journal of Economics, Vol. 129 (1), pp. 105–149.

[41] Luttmer, E.F.P. (2005), “Neighbors as Negatives: Relative Earnings and Well-Being,” Quarterly Journal of Economics, Vol. 120 (3), pp. 963-1002.

[42] Nelissen, R. and Meijers, M. (2011), “Social benefits of luxury brands as costly signals of wealth and status,” Evolution and Human Behavior, Vol. 32(5), pp. 343-355.

[43] Perez-Truglia, R. (2015), “The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment,” Mimeo.

[44] Perez-Truglia, R. (2017), “Political Conformity: Event-Study Evidence from the United States,” Review of Economics and Statistics, forthcoming.

[45] Perez-Truglia, R. and Cruces, G. (2017), “Partisan Interactions: Evidence from a Field Experiment in the United States,” Journal of Political Economy, Vol. 125 (4), pp. 1208–1243.

[46] Rees-Jones, A. (2017), “Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match,” Games and Economic Behavior, forthcoming.

[47] Rees-Jones, A. (2017), “Mistaken Play in the Deferred Acceptance Algorithm: Implications for Positive Assortative Matching,” American Economic Review: Papers and Proceedings, Vol. 107 (5), pp. 225-229.

[48] Rees-Jones, A. and Skowronek, S. (2017), “Why Do We Lie in Incentive-Compatible Mechanisms? Evidence from the Residency Match,” Mimeo.

[49] Richins, M.L. and Dawson, S. (1992), “A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation,” Journal of Consumer Research, Vol. 19, pp. 303–16.

[50] Roth, A.E. and Peranson, E., (1999), “The redesign of the matching market for American physicians: Some engineering aspects of economic design,” American Economic Review, Vol. 89 (4), p.748-780.

[51] Senik, C. (2004), “When Information Dominates Comparison. Learning from Russian Subjective Panel Data,” Journal of Public Economics, Vol. 88, pp. 2099-2133.

[52] Shigeoka, H. and Yamada, K. (2016), “Income-comparison Attitudes in the U.S. and the UK: Evidence from Discrete-choice Experiments,” NBER Working Paper No. 21998.

[53] Smither, R.D. and Houston, J.M., (1992), “The nature of competitiveness: The development and validation of the competitiveness index,” Educational and Psychological Measurement, Vol. 52 (2), pp. 407-418.

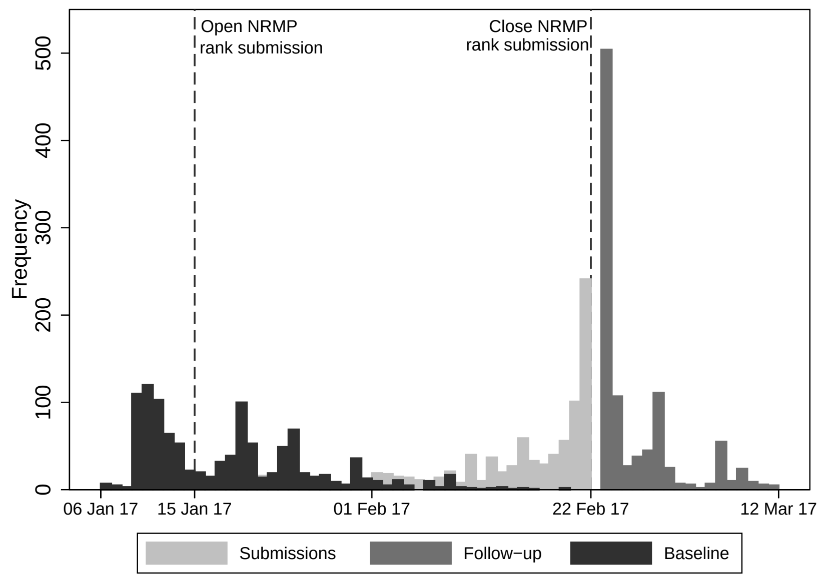
[54] Solnick, S. and Hemenway, D. (1998), “Is more always better? A survey on positional concerns,” Journal of Economic Behavior and Organization, Vol. 37, pp. 373–383.

[55] van de Stadt, H.; Kapteyn, A. and van de Geer, S. (1985), “The Relativity of Utility: Evidence from Panel Data,” Review of Economics and Statistics, Vol. 67, pp. 179–187.

[56] van Gelder, M.M.H.J.; Bretveld, R.W.; and Roeleveld, N. (2010), “Web-based Questionnaires: The Future in Epidemiology?” American Journal of Epidemiology, Vol. 172 (11), pp. 1292-1298.

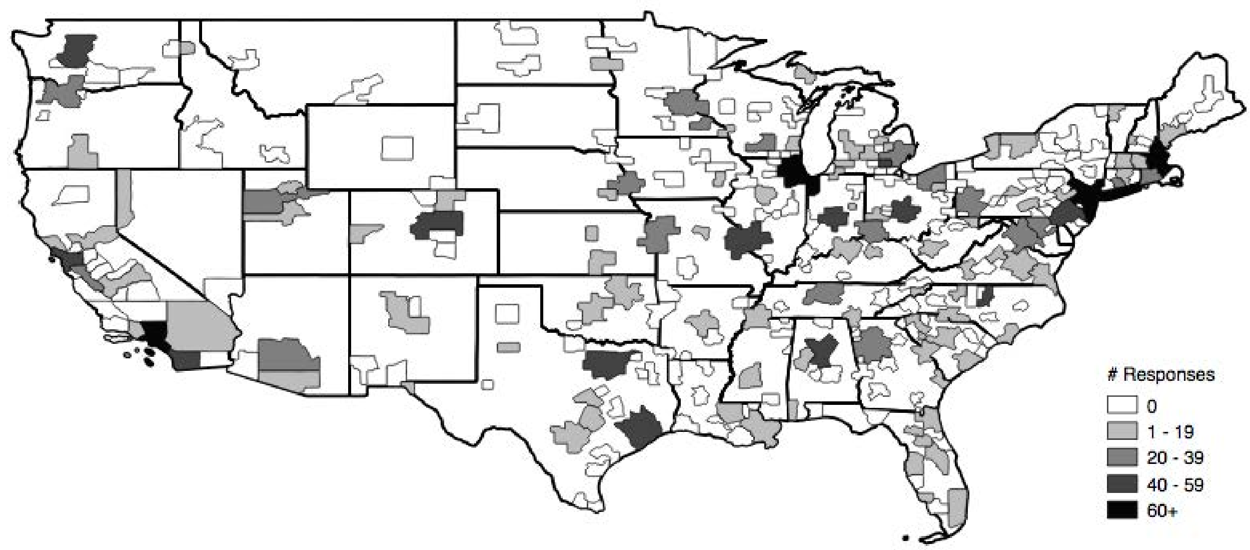
[57] Yamada, K. and Sato, M. (2013), “Another avenue for anatomy of income comparisons: Evidence from hypothetical choice experiments,” Journal of Economic Behavior & Organization, Vol. 89, pp. 35-57.

Figure 1: Distribution Over Time of Survey Responses and NRMP Rank Submissions



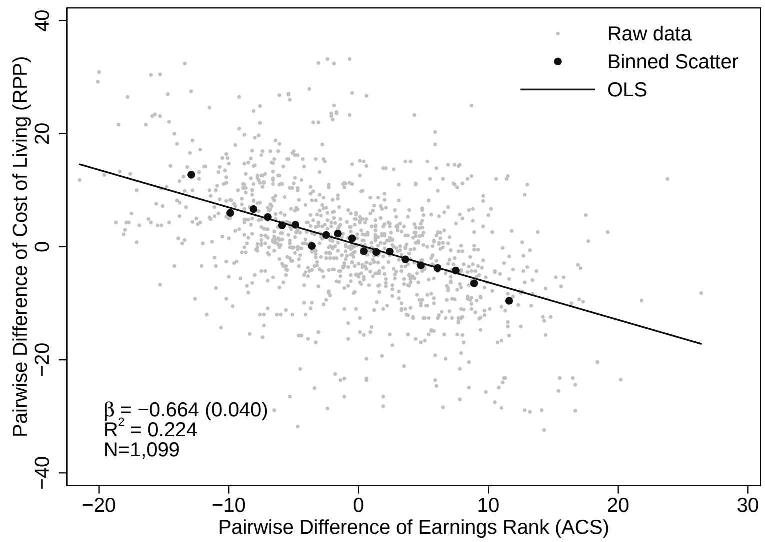
Notes: Distribution of timing of responses to Baseline and Follow-up Surveys, and NRMP rank submission dates (as reported by respondents in the follow-up survey).

Figure 2: Geographic Distribution of Choice Set



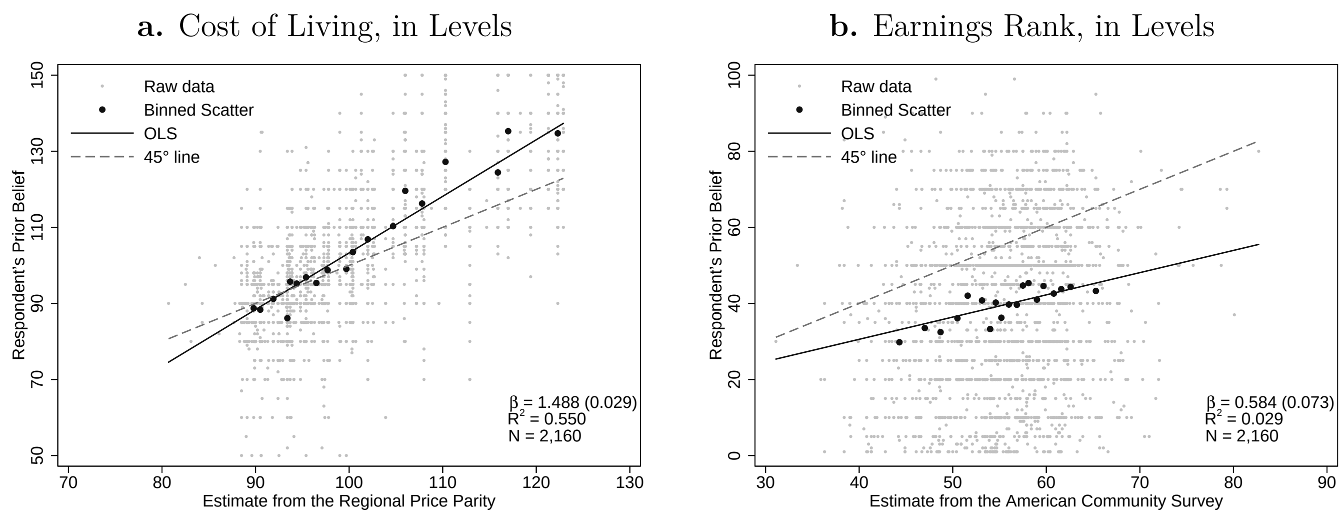
Notes: Geographical distribution of metropolitan areas where top-2 residency programs of respondents are located, for the continental United States. No responses were located in Hawaii, while Alaska only has 2 responses. Only metropolitan areas with a residency program participating in the 2017 NRMP are displayed (279 in total).

Figure 3: Variation in Cost of Living and Earnings Rank



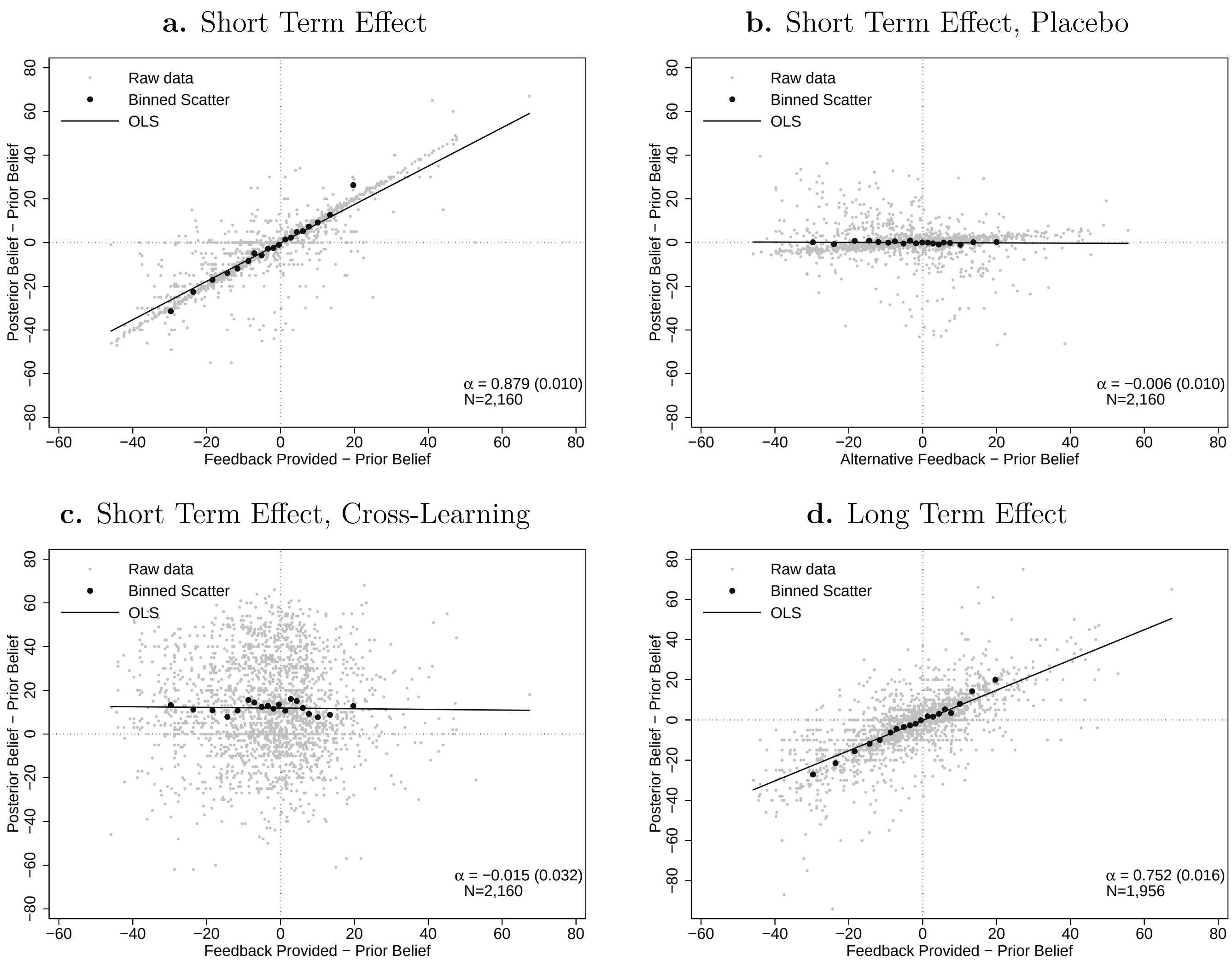
Notes: The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Slopes (β, with robust standard errors in parentheses) and R^2 are based on a linear regression. All variables for x-axis and y-axis correspond to pairwise differences across the two cities that the subject is considering submitting to the algorithm. Data from survey responses, the Regional Price Parity Index (for cost of living) and the American Community Survey (for earnings rank).

Figure 4: Comparison Between Prior Beliefs and Statistics



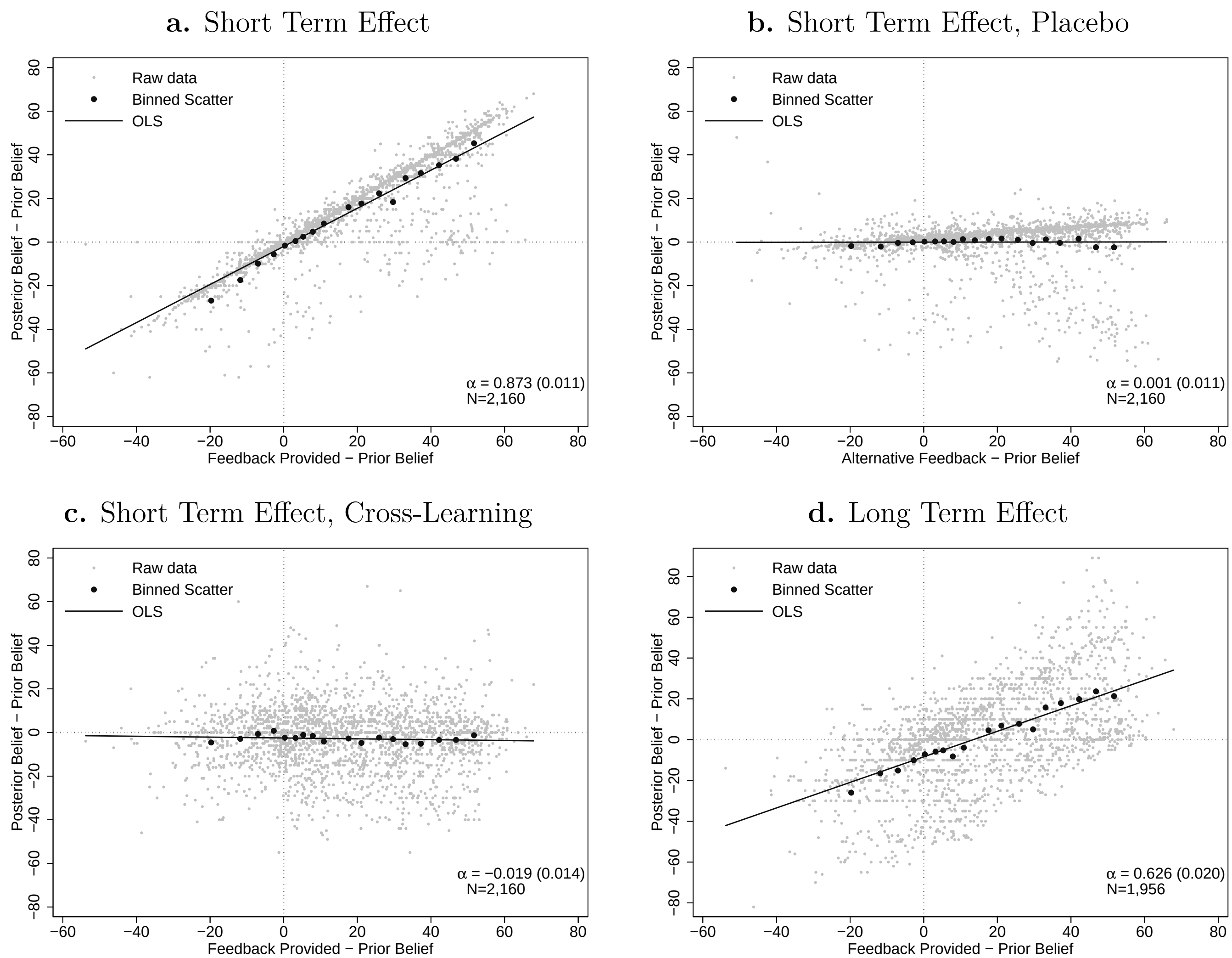
Notes: Comparison between respondent's perceptions before the information provision (i.e., prior beliefs) and statistics. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panels a and b present data in levels (i.e., two observations per individual, one for each of their options). The slope (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

Figure 5: Learning from the Experimental Feedback: Cost of Living



Notes: Comparison between the difference in statistics and respondent's perceptions before the information provision (i.e., prior beliefs), and difference in respondent's perceptions after the information provision (i.e., posterior beliefs) and prior beliefs. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panel b shows a placebo test where we compare the difference between the alternative feedback and prior belief to the difference between the posterior and prior beliefs, adjusting for the shown statistic. Panel c shows the extent to which respondents adjust their perceptions on earnings rank as a result in their perception gap in cost of living (adjusting for the perception gap in earnings rank). Panel d uses respondent's perceptions measured in the follow-up survey as posterior belief. The slope (α, with robust standard errors in parentheses) is based on a linear regression.

Figure 6: Learning from the Experimental Feedback: Earnings Rank



Notes: Comparison between the difference in statistics and respondent's perceptions before the information provision (i.e., prior beliefs), and difference in respondent's perceptions after the information provision (i.e., posterior beliefs) and prior beliefs. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panel b shows a placebo test where we compare the difference between the alternative feedback and prior belief to the difference between the posterior and prior beliefs, adjusting for the shown statistic. Panel c shows the extent to which respondents adjust their perceptions on cost of living as a result in their perception gap in earnings rank (adjusting for the perception gap in cost of living). Panel d uses respondent's perceptions measured in the follow-up survey as posterior belief. The slope (α, with robust standard errors in parentheses) is based on a linear regression.on a linear regression.

Table 1: Descriptive Statistics and Randomization Balance

	All	RPP; ACS	RPP; CPS	COLI; ACS	COLI; CPS	F-test P-value
Male (=1)	0.481	0.452	0.491	0.481	0.502	0.688
	(0.015)	(0.030)	(0.031)	(0.031)	(0.030)	
Age	27.091	27.092	27.104	26.985	27.181	0.863
	(0.083)	(0.164)	(0.165)	(0.145)	(0.187)	
Nr Kids	0.132	0.125	0.164	0.104	0.137	0.553
	(0.014)	(0.027)	(0.033)	(0.026)	(0.029)	
Single (=1)	0.354	0.401	0.312	0.343	0.358	0.189
	(0.015)	(0.030)	(0.028)	(0.029)	(0.029)	
Dual Match (=1)	0.074	0.077	0.059	0.104	0.055	0.157
	(0.008)	(0.016)	(0.014)	(0.019)	(0.014)	
US News Rank	58.81	58.849	59.104	58.604	58.683	0.996
	(0.787)	(1.612)	(1.560)	(1.568)	(1.565)	
Prior: $COL^i_{2,1}$	-0.409	-0.445	0.238	0.567	-1.982	0.506
	(0.640)	(1.364)	(1.134)	(1.308)	(1.303)	
Prior: $ER^i_{1,2}$	0.394	0.162	0.71	-0.526	1.221	0.595
	(0.467)	(0.903)	(0.925)	(0.906)	(0.999)	
Observations	1,080	272	269	268	271	

Notes: Individual characteristics obtained from baseline survey. Column (1) corresponds to all respondents, and columns (2) through (4) correspond to each of the four treatment groups given by all the possible combinations from the source-randomization experiment. RPP and COLI are the two sources used to compute the cost of living feedback (corresponding to the Regional Price Parity Index and the Cost of Living Index, respectively). ACS and CPS are the two sources used to compute the earnings ranking feedback (corresponding to the American Community Survey and the Current Population Survey, respectively). The final column presents p-value for test of the null hypothesis that the mean characteristic is equal across all four treatment groups. All variables constructed from the survey data, except for the U.S. News Rank which was taken from the U.S. News rank of medical schools for 2016.

Table 2: Preference for Relative Consumption: Baseline Estimates

		By Relationship Status		By Gender		By Specialty Salary	
	All	Non-Single	Single	Female	Male	$\leq$ \$229,000	$>$ \$229,000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
β^{rel}	0.989*	2.195***	-1.527*	1.034	0.898	1.411*	0.780
	(0.540)	(0.670)	(0.875)	(0.757)	(0.784)	(0.730)	(0.797)
β^{abs}	1.080**	1.095*	1.042	0.961	1.440*	0.691	1.232*
	(0.484)	(0.658)	(0.750)	(0.671)	(0.752)	(0.701)	(0.684)
Diff. P-value *[q-value]:*							
Relative		0.001 *[0.038]*		0.900 *[0.974]*		0.559 *[0.896]*	
Absolute		0.957 *[0.994]*		0.635 *[0.896]*		0.580 *[0.896]*	
Pseudo R^2	0.025	0.046	0.026	0.042	0.032	0.028	0.032
Observations	1,080	698	382	560	520	549	531

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit coefficients. Each column corresponds to a different Probit regression of expected rank order submission on posterior beliefs about cost of living and earnings rank, from the baseline survey, including the baseline controls listed in section 3. Columns (2) through (7) show estimates when restricting sample to different subgroups: columns 2 and 3 by Non-Single (i.e., married or in a long-term relationship) or Single, columns 4 and 5 by gender, columns 5 and 7 by expected salary of specialty after residency (over and below the median value of \$229,000). P-values corresponds to the test of the null hypothesis that the coefficients are equal between the two sub-groups, multiple-testing q-values based on Benjamini and Yekutieli (2001) presented in brackets.

Table 3: Preference for Relative Consumption: Robustness to Alternative Control Variables

	Panel A: β^{rel}			Panel B: β^{abs}			Pseudo R^2		
	All (1)	Non-Single (2)	Single (3)	All (4)	Non-Single (5)	Single (6)	All (7)	Non-Single (8)	Single (9)
No Controls	0.873*	1.961***	-1.480*	0.894**	0.812	1.131*	0.015	0.032	0.017
	(0.531)	(0.663)	(0.841)	(0.382)	(0.523)	(0.589)			
Baseline	1.022*	2.195***	-1.485*	1.106**	1.095*	1.140	0.026	0.046	0.034
	(0.542)	(0.669)	(0.895)	(0.484)	(0.658)	(0.746)			
Demographic	1.172**	2.313***	-0.987	1.359***	1.231**	1.677**	0.031	0.054	0.045
	(0.588)	(0.722)	(1.003)	(0.465)	(0.625)	(0.701)			
Amenities	0.958*	2.056***	-1.381	0.898*	0.718	1.265	0.018	0.037	0.022
	(0.538)	(0.669)	(0.853)	(0.481)	(0.630)	(0.816)			
Geography	1.001*	2.064***	-1.551	1.572***	1.626**	1.783***	0.039	0.059	0.054
	(0.593)	(0.733)	(1.004)	(0.461)	(0.652)	(0.646)			
Economic	0.946*	1.914***	-1.191	0.868*	0.467	1.647**	0.019	0.036	0.047
	(0.566)	(0.684)	(0.941)	(0.498)	(0.670)	(0.812)			
State Dummies	1.084*	2.901***	-1.943**	0.968*	1.090	1.219	0.049	0.105	0.149
	(0.555)	(0.703)	(0.907)	(0.502)	(0.671)	(0.939)			
Obj. Program Chars.	0.964*	1.985***	-1.366	0.942**	0.866	1.289**	0.019	0.037	0.028
	(0.541)	(0.686)	(0.857)	(0.385)	(0.527)	(0.589)			
Subj. Program Chars.	1.199**	2.222***	-1.320	1.277***	1.210**	1.678***	0.123	0.142	0.137
	(0.605)	(0.730)	(1.040)	(0.425)	(0.587)	(0.619)			

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit coefficients. Each row corresponds to a separate specification of a Probit regression of expected rank order submission on posterior beliefs about cost of living and earnings rank, from the baseline survey. All regressions include as controls the log difference in nominal income and a constant. The first row does not include any additional controls. The second row includes the baseline controls listed in section 3. The third to last rows use different sets of additional controls, listed in section 6.4. Results are based on 1,080 individual responses (698 from non-singles and 382 from singles), except for the last row, which is restricted to the follow-up sample (978 responses, 595 from non-singles and 311 from singles).

Table 4: Preference for Relative Consumption: Experimental Estimates

	Panel A: β^{rel}			Panel B: β^{abs}		
	All (1)	Non-Single (2)	Single (3)	All (4)	Non-Single (5)	Single (6)
Baseline	1.130*	2.337***	-1.666*	1.271**	1.230*	1.401*
	(0.578)	(0.703)	(0.995)	(0.529)	(0.739)	(0.780)
Experimental	0.858	2.955**	-4.984**	-0.653	-0.336	-1.660
	(1.150)	(1.332)	(1.950)	(0.880)	(1.163)	(1.288)
Experimental, Long Term	-0.055	1.946*	-5.279***	-1.068	-1.723*	-0.250
	(1.072)	(1.183)	(1.986)	(0.822)	(1.001)	(1.357)
Experimental, Falsification	-0.032	-0.039	-0.021	0.004	-0.008	0.039
	(0.650)	(0.854)	(1.119)	(0.836)	(0.997)	(1.728)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit (or IV-Probit) coefficients. All regressions include the baseline controls listed in section 3. The independent variables are the posterior beliefs about cost of living and earnings rank, from the baseline specification. The first row corresponds to the baseline Probit specification. The second through third row correspond to IV-Probit regressions, using the variation in perceptions generated by the source-randomization experiment as instrumental variables. The first and second rows use the expected rank order submission (from the baseline survey) as dependent variable. The third row uses the final rank order submission (from the follow-up survey) as dependent variable. The fourth row provides corresponds to a falsification test that uses the same IV-Probit specification from the second row, but using the list order (i.e., the order in which programs are listed at the beginning of the survey) as dependent variable instead of the rank order. To estimate this IV-Probit model, we randomly assign programs to be program 1 and program 2, and then we use as dependent variable a dummy that takes the value 1 if program 1 was listed first at the beginning of the survey. We repeat this procedure 1,000 times and report the average and standard error from the distribution of coefficients. All results based on the sample of individuals who completed the follow-up survey (978 responses, 647 from non-singles and 311 from singles).

Online Appendix: For Online Publication Only

A Snapshots of Invitations and Surveys

Here we include snapshots with a sample of the baseline survey (A.1) and follow-up survey (A.2). Additionally, this Appendix also includes a snapshot of an invitation sent out to the deans (A.3), the invitation sent out to the medical students to participate in the baseline survey (A.4), the invitation sent out to students inviting them to the follow-up survey (A.5), and a snapshot of the project's website.

A.1 Sample Questionnaire: Baseline Survey

This survey has the objective of understanding how participants of the 2017 NRMP make their ranking decisions. Even though it may not benefit you directly, the results from this survey may benefit the medical students participating in future years.

We anticipate that this survey will take between 8 to 10 minutes to complete. Eligible participants completing the entire survey will be paid $10 in the form of an Amazon Gift Card (note: you must have a .edu email address).

Your participation is voluntary, and is greatly appreciated. You may withdraw from the study at any time. Your responses will be used solely for research purposes and will be kept strictly confidential, used only by the Principal Investigators. For more details about this survey, including contact information, please visit the project's website.

To be eligible to participate in this survey, **you must be a medical student participating in the 2017 Main Residency Match and not yet submitted your rankings.**

- YES, I am participating in the 2017 Main Residency Match and would like to complete the survey

>>

NOTE: Please answer questions carefully, it is not possible to go back and change an answer.

Where are you attending Medical School?

State: Illinois

Medical School: University of Illinois

Which **match** will you be participating in?
(Note: this is referring to the match, not necessarily your specialty)

Main Residency (Opens Jan 15)

Will you register with the NRMP for a dual match?

- ○ Yes
- ● No

Did you already submit your ranking to the NRMP?

- ○ Yes
- ● No

>>

In the next couple of weeks you will be submitting your rankings to the Main Residency Match.

Please tell us (in no particular order) the top two Residency Programs you are thinking about ranking in the Main Residency Match.

Enter information for first program.

State	California
Metropolitan Area	Los Angeles-Long Beach-Anaheim, CA
Program	Los Angeles County-Harbor-UCLA Medical Center

Specialty:

Internal Medicine (IM)

What is the annual salary you are being offered here? (pre-tax, in dollars)

62000

>>

Enter information for second program.

State: Illinois

Metropolitan Area: Champaign-Urbana, IL

Program: Carle Foundation Hospital

Specialty:

Internal Medicine (IM)

What is the annual salary you are being offered here? (pre-tax, in dollars)

60000

>>

Now, we want to ask you a couple of questions about the two cities you are considering living in.

Let's start with the expected cost of living. You probably noticed that the average prices of goods and services are different across different cities. As a result, with the same income, you would be able to buy more things in some cities and less in other cities.

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- ○ cheaper
- ● more expensive

How much more expensive is the Los Angeles-Long Beach-Anaheim, CA metro area than the U.S. average?

10% ▴▾

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- (selected) cheaper
- more expensive

How much cheaper is the Champaign-Urbana, IL metro area than the U.S. average?

7%

>>

Now we want to ask you about your expected earnings ranking. This ranking is defined as the share of the working individuals of a city who earn less than you. You probably noticed that the distribution of earnings is different across different cities. As a result, with the same earnings, you may be relatively rich in some cities but relatively poor in other cities.

Imagine that you chose to work in **Los Angeles-Long Beach-Anaheim, CA**. With your individual annual earnings of **$ 62000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

Richer than 37% of individual earners

>>

Imagine that you chose to work in **Champaign-Urbana, IL**. With your individual annual earnings of **$ 60000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s individual earners?

Richer than 52% of individual earners

>>

Now, we want to share some information with you, related to the characteristics of the two cities that you are considering living in. Please take a moment to review the information carefully.

***Note**: this information is only shown once and you will not be able to come back to it.*

First, find below some estimates of the cost of living:

The **Los Angeles-Long Beach-Anaheim, CA** metro area is **17.0% more expensive** than the U.S. average.

The **Champaign-Urbana, IL** metro area is **6.6% cheaper** than the U.S. average.

Source: based on most recent data from the Bureau of Economic Analysis.

>>

Second, find below some estimates of the earnings ranking:

With your individual annual earnings of **$ 62000**, you would be richer than **64.6%** of **Los Angeles-Long Beach-Anaheim, CA**'s population.

With your individual annual earnings of **$ 60000**, you would be richer than **61.2%** of **Champaign-Urbana, IL**'s population.

Source: based on most recent data from the Current Population Survey.

>>

That was all the information that we wanted to share with you. Now that you have reviewed this information, we would like to ask you again about your expected cost of living and earning rankings.

Let's start with the cost of living:

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- ○ cheaper
- ● more expensive

How much more expensive is the Los Angeles-Long Beach-Anaheim, CA metro area than the U.S. average?

17%

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- (•) cheaper
- ( ) more expensive

How much cheaper is the Champaign-Urbana, IL metro area than the U.S. average?

7%

>>

Imagine that you chose to work in **Los Angeles-Long Beach-Anaheim, CA**. With your individual annual earnings of **$ 62000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

Richer than 64% of individual earners

>>

Imagine that you chose to work in **Champaign-Urbana, IL**. With your individual annual earnings of **$ 60000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s individual earners?

Richer than 62% of individual earners

>>

We understand this is a lot of information to process, so we will help you make the comparison simpler. According to your final answers about incomes, cost of living and relative earnings:

- If you chose to live in Los Angeles-Long Beach-Anaheim, CA, you would be able to afford 17.9% less than if you chose to live in Champaign-Urbana, IL.

- If you chose to live in Los Angeles-Long Beach-Anaheim, CA, your earnings ranking would be 3.2% higher than if you chose to live in Champaign-Urbana, IL.

>>

As of this moment: of the two programs discussed so far, which one do you expect to **rank higher** for the NRMP?

- Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- Likely Los Angeles County-Harbor-UCLA Medical Center
- Leaning Los Angeles County-Harbor-UCLA Medical Center
- Leaning Carle Foundation Hospital
- Likely Carle Foundation Hospital
- Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

>>

If assigned to it, in which of the two programs would you expect to **live a happier life**?

- ○ Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely Los Angeles County-Harbor-UCLA Medical Center
- ○ Leaning Los Angeles County-Harbor-UCLA Medical Center
- ○ Leaning Carle Foundation Hospital
- ○ Likely Carle Foundation Hospital
- ○ Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

>>

To get a general picture of the people answering this survey, we would like to ask you a few things about yourself. Please remember that your answers are confidential and that your name is not collected as part of this study. Please indicate your gender:

- ( ) Female
- (●) Male

How old are you? [30 ▾]

What is your relationship status?

- ( ) Single
- ( ) In a long-term relationship
- (●) Married

How many children do you have?

[None ▾]

Recent research on decision making shows that choices are affected by the context in which they are made. Differences in how people feel, in their previous knowledge and experience, and in their environment can influence the choices they make. To help us understand how people make decisions, we are interested in information about you, specifically whether you actually take the time to read the instructions; if you don't, some results may fail to tell us very much about decision making in the real world. To help us confirm that you have read these instructions, please ignore the question below about how you are feeling and instead check only the "none of the above" option. Thank you very much.

☐ Interested	☐ Enthusiastic	☐ Inspired
☐ Distressed	☐ Proud	☐ Determined
☐ Excited	☐ Irritable	☐ Attentive
☐ Scared	☐ Alert	☑ None of the above

>>

Thank you so much for completing the survey! As a token of our appreciation, we want to send you a $10 Amazon Gift Card. Please note that you may only participate once.

We need your official university email address (.edu) to be able to: (i) email you the Amazon gift card; and (ii) verify that you are a medical student participating in the 2017 NRMP.

I certify that I am a medical student participating in the 2017 NRMP match.

Please sign with your university (.edu) email address:

email@university.edu

As a reminder, your email address and survey responses will be kept strictly confidential.

>>

A.2 Sample Questionnaire: Follow-Up Survey

Thank you for volunteering to participate in our follow-up survey! Remember that your responses will be used solely for research purposes and will be kept strictly confidential. You may withdraw from the survey at any time.

We estimate that it will take you around 5 minutes to complete the survey. As a token of our appreciation, we will send you a $5 Amazon gift card for completing this survey.

For more details about the survey, including contact information, please visit the project's website.

>>

On what date did you submit your preference ranking to the 2017 Main Residency Match?

Feb 18

>>

In the initial survey you listed two of your favorite programs.

When you submitted your preference ranking to the 2017 Main Residency Match on Feb 18, which of these two programs did you **rank higher**?

- ○ Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Carle Foundation Hospital (Champaign-Urbana, IL)

>>

If assigned to it, in which of the two programs would you expect to **live a happier life**?

- ○ Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely Los Angeles County-Harbor-UCLA Medical Center
- ○ Leaning Los Angeles County-Harbor-UCLA Medical Center
- ○ Leaning Carle Foundation Hospital
- ○ Likely Carle Foundation Hospital
- ○ Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

>>

Now, we want to ask you a couple of questions about the two cities where you may live. When you took the survey a month ago, we asked these same questions. We are asking them again to see if your perceptions have changed.

Let's start with the expected cost of living. You probably noticed that the average prices of goods and services are different across different cities. As a result, with the same income, you would be able to buy more things in some cities and less in other cities.

Imagine that you chose to work in the Los Angeles-Long Beach-Anaheim, CA metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- ○ cheaper
- ● more expensive

How much more expensive is the Los Angeles-Long Beach-Anaheim, CA metro area than the U.S. average?

17%

>>

Imagine that you chose to work in the Champaign-Urbana, IL metro area. Would you expect your cost of living in this city to be cheaper or more expensive than the U.S. average?

- [x] cheaper
- [ ] more expensive

How much cheaper is the Champaign-Urbana, IL metro area than the U.S. average?

7%

>>

Now we want to ask you about your expected earnings ranking. This ranking is defined as the share of the working individuals of a city who earn less than you. You probably noticed that the distribution of earnings is different across different cities. As a result, with the same earnings, you may be relatively rich in some cities but relatively poor in other cities.

Imagine that you chose to work in Los Angeles-Long Beach-Anaheim, CA. With your individual annual earnings of **$** 62000, you would be richer than what percentage of Los Angeles-Long Beach-Anaheim, CA's individual earners?

Richer than 64% of individual earners

>>

Imagine that you chose to work in Champaign-Urbana, IL. With your individual annual earnings of **$** 60000, you would be richer than what percentage of Champaign-Urbana, IL's individual earners?

Richer than 62% of individual earners

>>

Now we want to ask you to compare other aspects of these two programs.

In which program do you expect to have a greater sense of purpose in life?

- Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- Likely Los Angeles County-Harbor-UCLA Medical Center
- Leaning Los Angeles County-Harbor-UCLA Medical Center
- Leaning Carle Foundation Hospital
- Likely Carle Foundation Hospital
- Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

Which program do you think will give you higher prestige and status?

- Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- Likely Los Angeles County-Harbor-UCLA Medical Center
- Leaning Los Angeles County-Harbor-UCLA Medical Center
- Leaning Carle Foundation Hospital
- Likely Carle Foundation Hospital
- Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

Which program do you think will give you better future career prospects?

- Very likely Los Angeles County-Harbor-UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- Likely Los Angeles County-Harbor-UCLA Medical Center
- Leaning Los Angeles County-Harbor-UCLA Medical Center
- Leaning Carle Foundation Hospital
- Likely Carle Foundation Hospital
- Very likely Carle Foundation Hospital (Champaign-Urbana, IL)

>>

From 1 (most important) to 5 (least important): How would you rank the following aspects of life? *(no ties)*

	1	2	3	4	5
Happiness	○	○	○	○	○
Health	○	○	○	○	○
Sense of purpose	○	○	○	○	○
Spirituality	○	○	○	○	○
Control over your life	○	○	○	○	○

>>

When you submitted your preference ranking to the 2017 Main Residency Match on Feb 18, how many programs did you rank in total?

17

>>

Now we want to ask you a few more questions about your background, your beliefs and your values.

Did you grow up in the United States?

- ○ Yes
- ○ No

>>

More precisely, in which of the following did you spend the most time while growing up?

State

Metro area

>>

Imagine that you face the following situation. You earn $50,000 per year and have an earnings ranking of 50% (that is, you earn more than 50% of the individuals living in your same city). Now consider the following two events:

EVENT A: The cost of living in this city decreases by 10%, so you and all other individuals in the city would be able to afford 10% more consumption. After this event, you think you would be:

- Better off
- Slightly better off
- The same
- Slightly worse off
- Worse off

EVENT B: Your own income and your own cost of living do not change, so your own consumption stays the same. However, all other individuals in the city face an income reduction. As a result, your earnings ranking increases from 50% to 60%. After this event, you think you would be:

- Better off
- Slightly better off
- The same
- Slightly worse off
- Worse off

>>

Please indicate the degree to which you agree or disagree with the following statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I like competition	O	O	O	O	O
I am a competitive individual	O	O	O	O	O
I enjoy competing against an opponent	O	O	O	O	O
I don't like competing against other people	O	O	O	O	O
I get satisfaction from competing with others	O	O	O	O	O
I find competitive situations unpleasant	O	O	O	O	O

>>

We are almost done, this is the last question of the survey. Please indicate the degree to which you agree or disagree with the following statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I admire people who own expensive homes, cars, and clothes	O	O	O	O	O
The things I own say a lot about how I'm doing in life	O	O	O	O	O
Buying things gives me a lot of pleasure	O	O	O	O	O
I like a lot of luxury in my life	O	O	O	O	O
My life would be better if I owned certain things I don't have	O	O	O	O	O
I'd be happier if I could afford to buy more things	O	O	O	O	O

>>

A.3 Sample Invitation Email to Deans

Dear Dean X,

I am a Graduate Student in Economics at the University of Illinois. Along with Ricardo Perez-Truglia (Economics Professor at University of California Los Angeles, Anderson School of Management), we are working on a research project about how people make important life decisions. I am writing you in your capacity as Dean in the hope that you would allow us to survey the students at University X about their choices in the National Residency Matching Program next January, before the ranking submission window opens.

Completing the online survey would take the students less than 10 minutes, and as a token of appreciation, we will send each respondent a $10 Amazon gift card. I have attached a draft of the survey for your reference. The questions are non-controversial, responses will remain strictly confidential, and we are open to incorporating your feedback into the survey.

The NRMP provides a perfect context to study important life decisions. We hope that the results from our study could provide useful information and insights to future generations of medical students applying to residency programs, and provide new insights to residency programs.

If you have any questions about the survey, we would be happy to answer them over e-mail or schedule a time for a brief phone conversation. We will be surveying students from medical schools around the country, and would love to add University X to our list of participating medical schools. Can we please count with your collaboration?

Best regards,

A.4 Sample Email to Students with Invitation to Baseline Survey

Dear graduating medical student,

We would like to invite you to participate in a brief, confidential survey about the Main Residency Match. **It takes less than 10 minutes to complete the survey and, as a token of our appreciation, respondents will be sent a $10 Amazon gift card by email.**

To participate in the survey, you must be registered in the 2017 Main Residency Match. If you want to participate, you must fill out the survey before you submit your rankings to the NRMP.

The survey can be accessed here: [LINK]

The results of this study will provide better information on how medical students select residency programs, and can assist in the advising and preparation of future generations of students.

We thank you and deeply appreciate your time and participation,

Ricardo Perez-Truglia, University of California, Los Angeles
Nicolas Bottan, University of Illinois at Urbana-Champaign
[Project's URL]

A.5 Sample Email to Students with Invitation to Follow-Up Survey

Dear graduating medical student,

Thank you for participating in our study! We wanted to invite you to participate in a very short follow-up survey. Your participation is voluntary and all responses will be kept strictly confidential. **It takes less than 5 minutes to complete the survey, and, as a token of our appreciation, we will send you a $5 Amazon gift card by email.**

Follow this link to the Survey: [LINK]

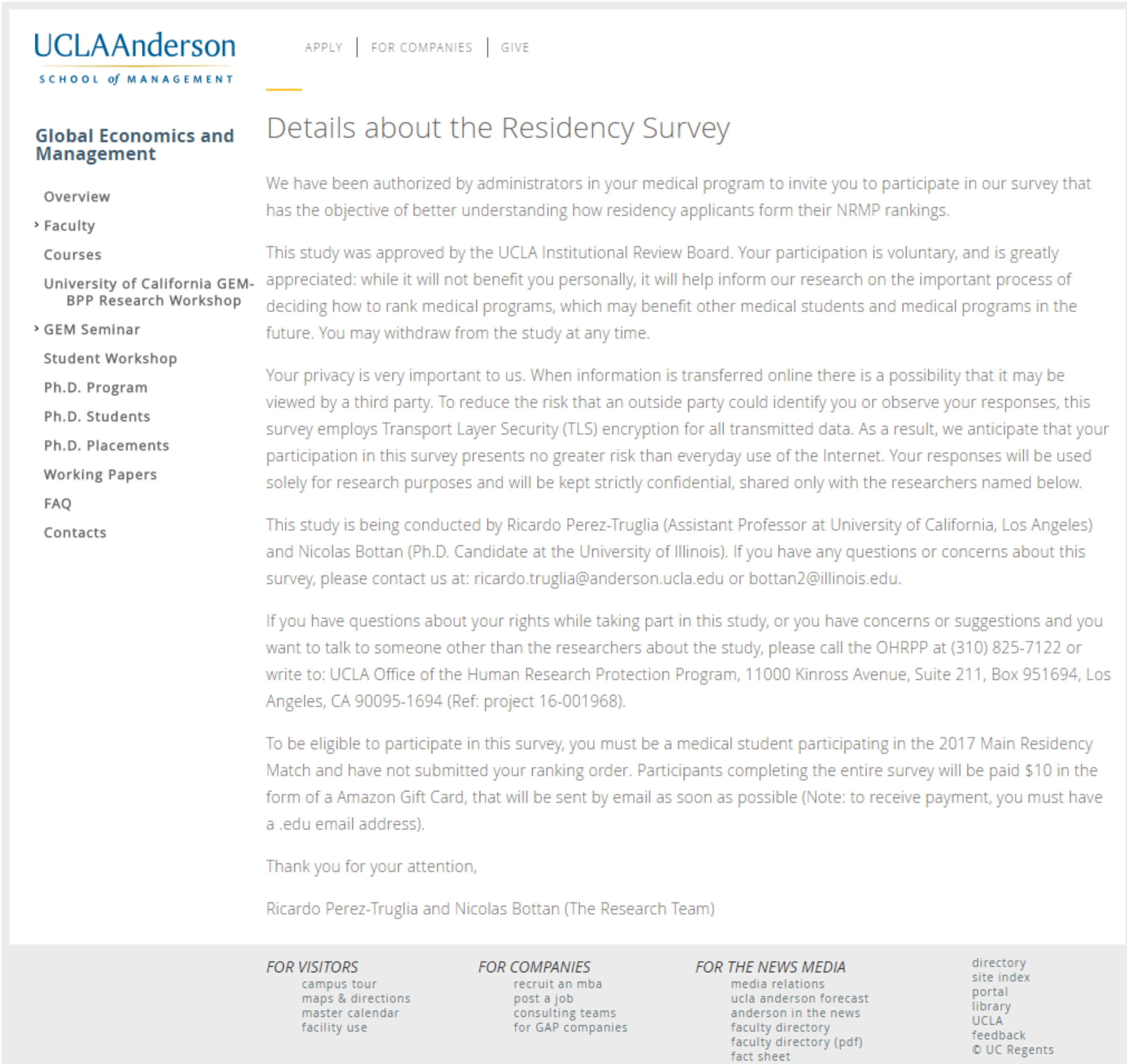
Or copy and paste the URL below into your internet browser: [URL]

After you complete this follow-up survey, your contact information will erased and we will not contact you again.

We thank you again and deeply appreciate your time and participation,

Ricardo Perez-Truglia, University of California, Los Angeles
Nicolas Bottan, University of Illinois at Urbana-Champaign
[Project's URL]
[Unsubscribe LINK]

A.6 Project's Website



UCLA Anderson
SCHOOL of MANAGEMENT

APPLY | FOR COMPANIES | GIVE

Global Economics and Management

- Overview
- › Faculty
- Courses
- University of California GEM-BPP Research Workshop
- › GEM Seminar
- Student Workshop
- Ph.D. Program
- Ph.D. Students
- Ph.D. Placements
- Working Papers
- FAQ
- Contacts

Details about the Residency Survey

We have been authorized by administrators in your medical program to invite you to participate in our survey that has the objective of better understanding how residency applicants form their NRMP rankings.

This study was approved by the UCLA Institutional Review Board. Your participation is voluntary, and is greatly appreciated: while it will not benefit you personally, it will help inform our research on the important process of deciding how to rank medical programs, which may benefit other medical students and medical programs in the future. You may withdraw from the study at any time.

Your privacy is very important to us. When information is transferred online there is a possibility that it may be viewed by a third party. To reduce the risk that an outside party could identify you or observe your responses, this survey employs Transport Layer Security (TLS) encryption for all transmitted data. As a result, we anticipate that your participation in this survey presents no greater risk than everyday use of the Internet. Your responses will be used solely for research purposes and will be kept strictly confidential, shared only with the researchers named below.

This study is being conducted by Ricardo Perez-Truglia (Assistant Professor at University of California, Los Angeles) and Nicolas Bottan (Ph.D. Candidate at the University of Illinois). If you have any questions or concerns about this survey, please contact us at: ricardo.truglia@anderson.ucla.edu or bottan2@illinois.edu.

If you have questions about your rights while taking part in this study, or you have concerns or suggestions and you want to talk to someone other than the researchers about the study, please call the OHRPP at (310) 825-7122 or write to: UCLA Office of the Human Research Protection Program, 11000 Kinross Avenue, Suite 211, Box 951694, Los Angeles, CA 90095-1694 (Ref: project 16-001968).

To be eligible to participate in this survey, you must be a medical student participating in the 2017 Main Residency Match and have not submitted your ranking order. Participants completing the entire survey will be paid $10 in the form of a Amazon Gift Card, that will be sent by email as soon as possible (Note: to receive payment, you must have a .edu email address).

Thank you for your attention,

Ricardo Perez-Truglia and Nicolas Bottan (The Research Team)

FOR VISITORS
campus tour
maps & directions
master calendar
facility use

FOR COMPANIES
recruit an mba
post a job
consulting teams
for GAP companies

FOR THE NEWS MEDIA
media relations
ucla anderson forecast
anderson in the news
faculty directory
faculty directory (pdf)
fact sheet

directory
site index
portal
library
UCLA
feedback
© UC Regents

B Information about the Subject Pool

We recruited 27 of the 135 accredited medical schools in the U.S. to participate in our study. In order to compare school characteristics from our sample with those not participating in our study, we obtained data from U.S. News (that is best known for compiling data and publishing ranks for universities and hospitals). We present descriptive statistics for the universe of medical schools, non-participating and participating schools in Table B.2. Medical schools participating in our study have slightly higher enrollment, lower average MCAT score, and are a little lower ranked on average than non-participating schools. However, none of these differences are statistically significant at conventional levels. The only statistically significant difference we do find is that the faculty to student ratio in participating schools is lower than in non-participating schools. Overall, it seems that participating medical schools are fairly representative of the overall universe of schools and not substantially different from non-participating schools.

Next, in Table B.1, we present the list of participating medical schools, along with the estimated size of the senior cohort, number of finished surveys and response rates. Around half of the schools reported the exact number of senior students who were participating in the Main Residency Match. For the remaining schools, we imputed the values of these variable using the average for the reporting schools (22% of the total enrollment). The overall response rate was almost 30%. Note that in the table we are excluding 20 observations that were deemed invalid either because answers to key questions were missing or feedback did not display correctly. These issues were due to technical difficulties most likely due to using a outdated internet browser without the proper Javascript support required to display and interact correctly with the survey. We have significant variation in response rates across medical schools. The response rate at Penn State is particularly low due to the fact that instead of forwarding the invitation by email, fliers were posted in the student lounge.

The day after the rank order submission deadline to the NRMP, we sent email invitations to the follow-up survey directly to respondents who had participated in the baseline survey. In Table B.3, we present descriptive statistics for our entire sample, and by whether respondents participated in the follow-up or not. The overall response rate to the follow-up was 90.6%. We do not find any statistically significant differences between the follow-up and non-follow-up respondents for all variables with the exception for single, where it appears that single students were less likely to participate in the follow-up survey. Additionally, participants to the follow-up survey reported slightly higher prior beliefs in cost of living than non-follow-up respondents. However, they were similarly "accurate" in their prior belief of cost of living.

Figure B.1: Distribution of Medical Schools in the U.S.

Notes: Each dot represents one of the 135 accredited medical schools contacted to participate in the study (excluding one in Hawaii). Dots do not denote exact location since they were moved to avoid overlap. Dark dots denote medical schools that agreed to participate in our study.

Table B.1: Survey Participation

State	University	Est. Senior Cohort	Nr Finished Surveys	Est. Response Rate (%)
Alabama	University of Alabama	174	47	27.0
Alabama	University of South Alabama	73	21	28.8
Arizona	University of Arizona	72	18	25.0
California	UC San Diego	124	39	31.5
Connecticut	Yale University	121	24	19.8
Florida	University of Florida	135	52	38.5
Illinois	Loyola University	145	66	45.5
Illinois	University of Illinois	20	8	40.0
Indiana	Indiana University	345	89	25.8
Massachusetts	Tufts University	194	42	21.6
Michigan	Michigan State University	183	76	41.5
Missouri	Saint Louis University	165	70	42.4
Missouri	University of Missouri (Kansas City)	101	34	33.7
Nebraska	University of Nebraska	125	46	36.8
New Mexico	University of New Mexico	97	27	27.8
New York	Stony Brook University	126	16	12.7
New York	University of Rochester	103	37	35.9
Ohio	Ohio State University	172	61	35.5
Oklahoma	University of Oklahoma	147	47	32.0
Pennsylvania	Pennsylvania State University	139	4	2.9
Rhode Island	Brown University	126	34	27.0
South Carolina	University of South Carolina	90	21	23.3
Texas	Baylor	180	44	24.4
Texas	Paul L. Foster School of Medicine (TTU)	89	30	33.7
Vermont	University of Vermont	105	39	37.1
Virginia	Virginia Commonwealth University	215	65	30.2
West Virginia	West Virginia University	110	23	20.9
	Total	3,676	1,080	29.38

Notes: 20 responses were excluded because they were deemed invalid (e.g., they did not received feedback due to a technical issue with their Internet Browser). Estimated senior cohort based on actual cohort size for schools that reported, and estimated as 22% of total enrollment for those that did not report cohort size (where 22% is the average proportion of seniors to total enrollment for schools that reported senior cohort size).

Table B.2: Comparison of Characteristics between Participating and Non-Participating Medical Schools

	All schools	Non-Participants	Participants	P-value
Enrollment	630.98	619.338	671.727	0.398
	(23.117)	(24.891)	(57.213)	
NR	0.267	0.287	0.185	0.245
	(0.038)	(0.044)	(0.076)	
Avg. MCAT	32.222	32.364	31.727	0.253
	(0.252)	(0.293)	(0.475)	
NR	0.267	0.287	0.185	0.245
	(0.038)	(0.044)	(0.076)	
Undergrad GPA	3.735	3.734	3.737	0.902
	(0.009)	(0.010)	(0.019)	
NR	0.267	0.287	0.185	0.245
	(0.038)	(0.044)	(0.076)	
Acceptance rate	0.066	0.067	0.062	0.458
	(0.003)	(0.004)	(0.005)	
NR	0.274	0.296	0.185	0.206
	(0.039)	(0.044)	(0.076)	
US News Ranking	45.451	43.478	51.636	0.166
	(2.784)	(3.309)	(4.872)	
NR	0.326	0.361	0.185	0.049
	(0.040)	(0.046)	(0.076)	
Tuition	51,404.98	51,333.526	51,651.818	0.913
	(1,097.842)	(1,193.139)	(2,689.180)	
NR	0.274	0.296	0.185	0.206
	(0.039)	(0.044)	(0.076)	
Faculty per student	2.363	2.518	1.827	0.039
	(0.221)	(0.279)	(0.177)	
NR	0.274	0.296	0.185	0.206
	(0.039)	(0.044)	(0.076)	
Peer Assessment score	3.14	3.139	3.145	0.961
	(0.076)	(0.093)	(0.106)	
NR	0.222	0.231	0.185	0.59
	(0.036)	(0.041)	(0.076)	
Observations	135	108	27	

Notes: Data for 135 accredited medical schools contacted by authors to participate in study. Data obtained from U.S. News for 2016. NR indicates the proportion of observations for which the statistic was either not published or missing. P-value in final column for the difference in means between participating and non-participating medical schools. Standard deviations reported in parenthesis.

Table B.3: Comparison of Characteristics between Respondents to Baseline and Follow-Up Surveys

	All	No Follow-up	Follow-up	P-value
Male (=1)	0.481	0.505	0.479	0.621
	(0.015)	(0.050)	(0.016)	
Age	27.091	26.921	27.108	0.482
	(0.083)	(0.253)	(0.088)	
Nr Kids	0.132	0.079	0.138	0.160
	(0.014)	(0.039)	(0.015)	
Single (=1)	0.354	0.505	0.338	0.001
	(0.015)	(0.050)	(0.015)	
Dual match (=1)	0.074	0.079	0.074	0.841
	(0.008)	(0.027)	(0.008)	
RPP treatment (=1)	0.499	0.525	0.496	0.588
	(0.015)	(0.050)	(0.016)	
ACS treatment (=1)	0.500	0.475	0.503	0.601
	(0.015)	(0.050)	(0.016)	
Average Residency Salary ($1000s)	0.013	0.019	0.012	0.871
	(0.013)	(0.042)	(0.014)	
Relative residency percentile	0.025	0.026	0.024	0.944
	(0.007)	(0.025)	(0.007)	
Pass Attention Check (=1)	0.964	0.950	0.965	0.509
	(0.006)	(0.022)	(0.006)	
Prior $ER_{1,2}$	0.004	0.008	0.004	0.775
	(0.005)	(0.013)	(0.005)	
Prior $COL_{2,1}$	-0.004	0.007	-0.005	0.479
	(0.006)	(0.016)	(0.007)	
Posterior $ER_{1,2}$	-0.009	-0.012	-0.008	0.639
	(0.003)	(0.008)	(0.003)	
Posterior $COL_{2,1}$	-0.010	-0.008	-0.010	0.856
	(0.004)	(0.014)	(0.004)	
Observations	1,080	101	979	

Notes: Standard deviations reported in parenthesis. P-values correspond to the test of the null hypothesis of equal means between follow-up and non-follow-up samples. Relative residency percentile based on residency quality ranks computed by Doximity. All variables constructed with data from the baseline survey.

Table B.4: Comparison of Characteristics between Experimental Subjects and U.S. Population of Earners

	Survey	ACS 2015
	Med. Students	Adult Earners
Age	27.091	41.258
	(2.725)	(12.330)
% Male	0.481	0.515
	(0.500)	(0.500)
% Married	0.240	0.531
	(0.427)	(0.499)
Wage	54,203.4	50,877.0
	(3,447.0)	(56,438.8)
US Born	0.950	0.809
	(0.218)	(0.393)
% More than College	1	0.125
	(0.000)	(0.331)

Notes: Data from 2015 American Community Survey PUMS for the subsample of adults in between 21 and 65 years of age and who receive positive wage income.

C Estimation of the Feedback Provided to Subjects

C.1 Earnings Rank

To provide feedback on the earnings rank of each metropolitan and wage offered at the location, we used data for the American Community Survey (ACS) at the metro area level for 2015 and the latest data from the Current Population Survey (CPS), as stated in the debriefing message.[48] From the data we estimated the parameters (μ and σ) for fitting a log-normal distribution. In the ACS we based this on the proportion of total full-time year round workers with earnings in each earnings bin, over which we estimated the parameters of fitting a log-normal distribution using maximum-likelihood for each metropolitan area. For the CPS, we combined weekly earnings with overtime earnings in order to obtain as close a measure as possible to that in the ACS. We obtained the parameters for fitting a log-normal distribution by estimating, for each metro area, a right-censored Tobit of annualized log earnings on the intercept. In the ACS, only 2% of metro areas were missing, while 20% of metro areas were missing for the CPS. Most of the metro areas with missing values from the ACS were imputed using the corresponding values obtained from the 2011-2015 5-year ACS.[49] The missing values in the CPS were imputed using the state-averages.

From these parameters, given a wage in dollars, it is easy to calculate the percentile using a simple formula: $\Phi(log(wage_j) - \mu_j)/\sigma_j$. This way it was possible to provide personalized feedback according to different wages reported in the survey. For our sample of metro areas, the average percentile rank for earnings of $55,000 is 59.2% and 68.9% for the ACS and CPS, and the correlation is 0.91.

Although both sources are very similar in levels, there is plenty of exogenous variation between them when comparing pairwise differences of chosen locations. We show this variation in Figure C.1.a, where the R-squared of regressing the pairwise differences for the ACS on the pairwise differences for the CPS is 0.430.

C.2 Cost of Living

To provide feedback on cost of living in the metropolitan areas we use the Regional Price Parity Index (RPP) compiled by the Bureau of Economic Analysis and the Cost of Living Index (COLI). The Cost of Living Index has been published since 1968 (formerly known as ACCRA) and has been used extensively in academic research. For the Regional Price Parity Index we used their final index for 2014 (the latest available at the time we conducted the

[48] At the time, the latest two months available were September and October of 2016.

[49] Only 3 metro areas were still missing for which we imputed with the average values for the country.

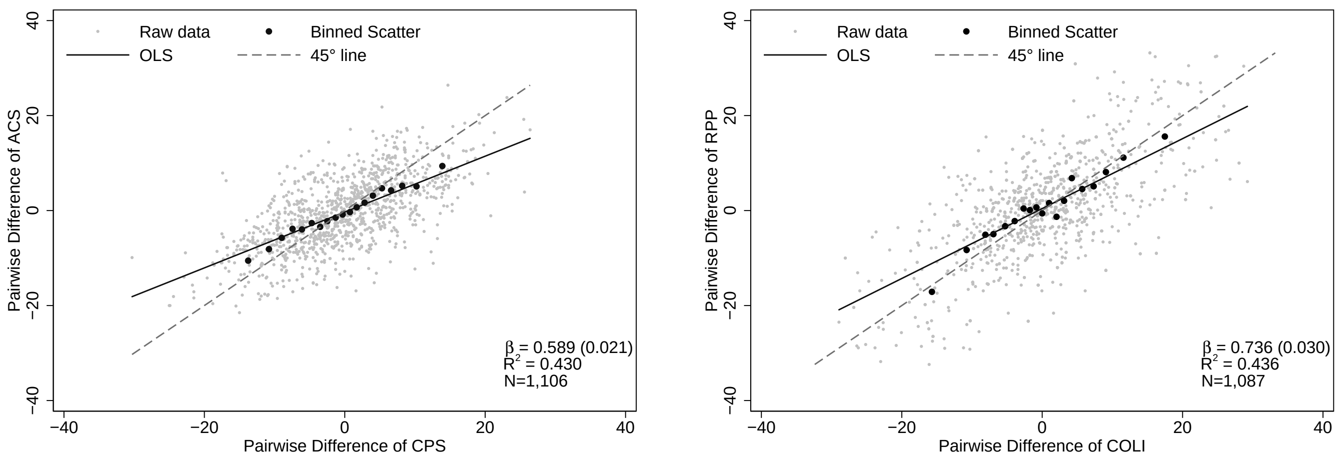
survey), and for the Cost of Living Index we used their raw data for the first quarter of 2016, and calculated our own index by computing a weighted average over the expenditure categories grocery, utilities, transport, health, and miscellaneous (excluding housing).

Both sources are quite similar: for the sample of potential metro areas that respondents can choose from based on the residency programs available, the mean cost of living for the RPP and COLI are 102.4 and 110.2, while the correlation between levels is 0.95. Note that the original indices have an average of 100. However, our sample of metro areas only include those for which there was a potential residency program to apply to. Therefore only 286 metro areas are included in our sample. These are mostly large and more expensive metro areas, which explains why the average is larger than 100. Additionally, 37 and 117 metro areas were imputed for RPP and COLI, respectively. We imputed values using predictions based on OLS regressions that included metro area census characteristics such as population, average household size, income, population density, racial and educational composition, housing characteristics and state dummies. The R^2 for those regressions was 93% and 86%. In our survey, less than 1% of metro options receiving RPP feedback were imputed, while only 11% of COLI feedback metro options were imputed.

Even though both sources are very similar when comparing them in levels, there is substantial exogenous variation when comparing the pairwise differences between chosen locations. The variation is presented in Figure C.1.b, where the R-squared for regressing the pairwise differences of RPP on COLI is only 0.436.

Figure C.1: Comparison of Statistics between the Two Different Data Sources

a. Earnings Rank **b.** Cost of Living



Notes: Pairwise differences of statistics from different sources of cost of living and earnings rank based on cities and wages indicated by respondents in the survey. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. The sources in Panel a correspond to ACS (American Community Survey) and CPS (Current Population Survey). The sources from Panel b correspond to RPP (Regional Price Parity Index) and COLI (Cost of Living Index). The slope (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

D Additional Results

D.1 Variation in Nominal Income, Cost of Living and Earnings Rank

Residency programs do not compensate for differences in cost of living or earnings rank through wages. Figure 3.a presents a scatterplot of the the costs of living versus the (log) nominal residency income. The strength of the association between these two variables represent the degree to which residencies compensate for local cost of living. The low value of the R^2 (0.115) shows that residencies compensate only partially, leaving substantial orthogonal variation between cost of living differences and differences in nominal income.

In a similar spirit, Figure 3.b explores the extent to which programs compensate for differences in the distribution of income through their nominal wages. This figure shows a scatterplot of the earnings rank at the residency income versus the nominal income. The strength of the association indicates how much of the differences in earnings rank are generated by differences in nominal income. Again, the low value of the R^2 (0.011) indicates that the vast majority of the variation in earnings rank is orthogonal to the nominal income.

D.2 Accuracy of Prior Beliefs, Pairwise Differences

Respondents may have a poor idea of the levels of cost of living and earnings rank, but they may have a better understanding of relative differences—ultimately the relevant statistic in decision making. We repeat our previous analysis, examining the pairwise differences instead of levels. in Figure D.2.a for cost of living and Figure D.2.b for earnings rank. Although the results for cost of living are almost identical, prior beliefs about earnings ranks are somewhat more accurate with pairwise differences. For example, the slope coefficient increases to 0.793, while the R^2 increases to 0.15 (5 times larger than in levels). In any case, even under this alternative specification, the accuracy of prior beliefs about earnings rank remain far less accurate than those for cost of living.

D.3 Learning Rates by Information Source

One concern with our experimental design is that individuals may have updated their beliefs differentially depending on the source used. For example, if respondents believe one source to be less trustworthy than another they may disregard that feedback. We explore the extent to which this could have happened by separately examining learning by the information source used. In Figures D.3 we present the same figures as in section 5.3 by information source.

Learning rates for cost of living and earnings ranking are almost identical between sources, showing that respondent's reactions to information did not depend on the source.

D.4 Persistence of Beliefs

Since posterior beliefs on cost of living and earnings ranking were elicited directly after providing respondents feedback, we are interested in examining how persistent these beliefs are a month later. We show that posterior beliefs are persistent for both cost of living and earnings rank in Figure D.4. The persistence in cost of living is twice as large as that of earnings rank (correlation of 0.844 versus 0.464), most likely due to respondents reverting to their prior beliefs over time and the fact that their priors were significantly more accurate for cost of living than earnings rank.

D.5 Complementary Evidence: Hypothetical Questions

To provide some additional suggestive evidence that individuals care about their consumption rank, we included a couple of hypothetical questions at the end of the follow-up survey (for the exact wording of this question, see Questionnaire Appendix A.2).

In the first scenario, we elicited the subjects' preferences for a reduction in absolute consumption while holding constant the relative consumption. More precisely, we asked the respondents whether they would be better off, the same, or worse off if their own consumption and the consumption of all other individuals in the city went up by 10%. Figure D.5.a shows the distribution of responses. Consistent with preferences for absolute consumption, 80% of respondents answered that they would be better off with this change, with 19% reporting that they would be the same and less than 1% responding that they would be worse off.

In the second scenario, we elicited the subjects' preferences for an increase in relative consumption, while holding constant the absolute consumption. To do so, we asked the respondents whether they would be better off, the same, or worse off if their own income and cost of living stayed the same but all other individuals in the city faced an income reduction of 10%. Because of the social desirability bias, individuals may not want to "confess" so directly that they care about relative consumption, and thus these responses probably lead to an underestimation of concerns for relative consumption. Figure D.5.b shows the distribution of responses. Consistent with individuals having direct preferences over relative consumption, 44% of individuals responded that they would be either better or worse off, with significant heterogeneity. While 31% of individuals reported that they would be better off with the poorer neighbors, 13% of individuals reported that they would be worse off.

D.6 Marginal Effects

Given that it is not possible to directly interpret coefficients from Probit regressions, in Table D.1 we present estimates from columns (1) to (3) from Table 2 along with their corresponding marginal effects at the average in the first two rows. The third and fourth row restrict the sample to respondents from the follow-up survey. For example, the coefficient in column (1) for the baseline sample implies that an increase of 1 percentage point in relative consumption in location 1 would increase the probability of choosing that location by 0.185 percent (or, in other words, an implied behavioral elasticity of 0.185).

D.7 Preferences over Subjective Program Characteristics

To better understand the magnitude of our results, we compare the estimates for preferences over relative and absolute consumption with those of subjective perceptions of residency program characteristics (prestige, career prospects and sense of purpose). These perceptions were elicited by the end of the follow-up survey and are standardized to have mean zero and standard deviation of one. We estimate the baseline model presented in section 3.1, introducing the three perceived program characteristics one by one. Since we only observe these perceptions in the follow-up survey, we restrict the sample to those respondents.

The results are presented in Table D.2. The coefficients on the three subjective perceptions of the program are positive, as expected, and highly statistically significant (all p-values<0.001). This means that individuals prefer programs associated with higher purpose, career prospects and prestige. Furthermore, we can compare the strength of these preferences to the strength of preferences for absolute consumption.[50]

We cannot compare the raw Probit coefficients directly, because the independent variables are measured in different units. For a meaningful comparison, we can calculate the standardized coefficient corresponding to a one standard deviation increase in absolute consumption. According to column (2) of Table D.2, a one standard deviation increase in absolute consumption would correspond to a Probit coefficient of 0.196 (i.e., the non-standardized coefficient, 1.422, multiplied by the standard deviation of absolute consumption, 0.138). This standardized coefficient can be compared to the coefficient of 0.437 corresponding to a one standard deviation increase in the sense of purpose. This comparison implies that the sense of purpose of a program is 2.22 times as important as the absolute consumption. By the same metric, the career prospects (column (3)) and sense of prestige (column (4)) are 1.93 and 1.27 times as important as absolute consumption. In sum, the characteristics of a program are system-

[50]The results are similar if we do the comparison with respect to the preferences for relative consumption instead.

atically more important for the choice of residency than the absolute consumption during the residency.

D.8 Preference Heterogeneity

In this section we explore additional heterogeneity over preferences for relative and absolute consumption. We first decompose the results of heterogeneity by relationship status in two ways. In columns (1) and (2) of Table D.3, we show that within non-single respondents, preferences over relative consumption are similar for married or long-term relationship respondents. However, it seems that preferences for absolute consumption are mostly driven by married respondents (though the difference is borderline insignificant, p-value=0.121). In columns (3) to (6) of Table D.3, we estimate preferences by gender, within relationship status. Preferences over relative consumption seem to be stronger for females in general, though the difference is not statistically significant for non-singles or singles.

In addition to the dimensions explored in the paper, we present results for heterogeneity across different dimensions in Table D.4. In columns (1) to (4) we explore heterogeneity according to differences in hypothetical choices of changes in absolute and relative consumption. Interestingly, we find that those who believe they would be better off if absolute consumption were to increase care significantly more about relative consumption than respondents who claimed they would be the same or worse off. However, we do not find any significant differences for the hypothetical question of a change in relative consumption.

Next, we explore whether there is preference heterogeneity across different individual traits, such as degree of materialism, competitiveness or life dimensions valued the most. The materialism index is based on questions that typically reflect status from consumption (see follow-up survey questionnaire in Appendix A.2, based on Richins and Dawson, 1992). Even though we do not find statistically significant difference in the effects in columns (5) and (6), the point estimates are different and reflect that those who are classified as more "materialistic" (or in other words, those most concerned by the signaling value of material goods) care more about relative consumption, while those who are less "materialistic" care more about absolute consumption. In columns (7) and (8) we explore heterogeneity by the degree of competitiveness using commonly used indices in psychology (Smither and Houston, 1992). We do not find any significant differences across these traits. Finally, in columns (9) and (10) we explore heterogeneity according to a principal component score of the rank of different life dimensions by importance (happiness, health, sense of purpose, spirituality, control over life). We do not find any statistically significant differences in these dimensions.

D.9 Sensitivity to Dropping Specific Subgroups

In this section we explore the sensitivity of our baseline results to dropping specific subgroups that may potentially attenuate our estimates for preferences over relative and absolute consumption. In the first row of Table D.5 we report the baseline estimates. In the second row, we re-estimate the model dropping respondents that did not successfully answer a question at the end of the baseline survey designed to test whether they were paying attention and reading the questions carefully. In this question we describe how emotions can play a role in influencing responses and respondents have a menu of emotions to choose from. However, at the end of the paragraph we instruct respondents to only select the option "none of the above" (see Appendix A.1 for the full question). Only 3.6% of respondents failed to answer this question correctly. Estimates do not change much when dropping these respondents – if anything, the coefficients are slightly larger in magnitude.

One additional concern is that respondents may not choose according to their own preferences but define it jointly with their spouse when they are both participating as a dual match. In the third row of Table D.5 we drop respondents who are participating in a dual match (7.4% of the sample). Again, the results are similar when we drop these respondents.

D.10 Binary Probit Vs. Ordered Probit

In the baseline survey we asked respondents about their intention to rank using a likelihood scale, that we later converted in to a binary variable in order to directly compare it to their final rank submission in the follow-up survey. However, we could also exploit the full variation of using the likelihood scale by means of estimating an ordered Probit model. The results are presented in Table D.6. Overall, the results are quite similar regardless of using the binary or likelihood variables.

D.11 Instrumental Variable Regression

We break down the Instrumental Variables regression into the first-stage and reduced-form regressions. Table D.7.a presents the same experimental estimates as those found in the second row of Table 4. In the next panel we focus on the first stages. As discussed in section 5.3, respondents learn from our information provision experiment, where learning rates are close to 1 for both earnings rank and cost of living. It does not seem that weak instruments are a problem overall. However, the instruments are substantially weaker for the sample of singles compared to the non-singles, where the Cragg-Donald F-statistic drops from 172 to 43. In the final panel of Table 4 we show that the reduced form estimates are very similar to those obtained by IV.

D.12 Comparison to Studies using Subjective Data

We are interested in comparing our results to those obtain in previous studies based on happiness surveys or hypothetical choices. It is important to note that these other studies measure relative concerns in a slightly different way. They compare the effects of own consumption versus the mean consumption of peers. They present an econometric model along the following lines:

$$U = a \cdot log(y) - b \cdot log(\bar{y})$$

Where y is the individual's own income and $\bar{y}$ is the average income in the individual's reference group. With parameters a and b, we can calculate the trade-off between absolute and relative income. The effect of absolute income is given by $a - b$: i.e., what would happen if increase my income by 1% if I am also increasing everyone else's income by 1%. The effect of relative income is just b: i.e., what happens if you increase everyone else's income by 1% while leaving my own income unchanged. An individual with parameters a and b should be indifferent between a 1% increase in her absolute consumption and a $\frac{a-b}{b}$ decrease in her relative consumption. Table D.8 shows the estimates of a and b reported in other studies, and the resulting estimate of $\frac{a-b}{b}$.[51]

Section 6.3 compares our estimates with respect to the findings from Luttmer (2005). In this section, we provide comparisons with respect to other studies. According to our baseline estimates (column (1) of Table 2), the average individual is willing to give up 1 percent of her absolute consumption to decrease the median consumption of her peers by 4.35%.[52] The other studies that use happiness data suggest a corresponding trade-off of 0.89% (Clark, Senik and Yamada, 2017) and 1.02% (Ferrer-i-Carbonell, 2005); while the studies using hypothetical choices suggest a corresponding trade-off of 1.85% (Johansson-Stenman et al., 2002) and 1.18% (Yamada and Sato, 2013). All of these estimates are below our own estimate of 4.35%, implying that, relative to these other studies, our estimates suggest a weaker role for relative concerns.

Last, we must note that some studies find the opposite effect. For instance, Senik (2004) and Clark, Kristensen and Westergård-Nielsen (2009) find that life satisfaction is increasing in the mean income of the reference group. And Shigeoka and Yamada (2016) show estimates from a hypothetical choice experiment with mixed results: while the U.K. respondents prefer

[51] The table does not include standard errors or confidence intervals because we do not have sufficient information to compute those ($\frac{a-b}{b}$ is a non-linear function, and thus it does not suffice with the standard errors of a and b).

[52] This result arises because, for the average individual in the sample, we would need to decrease the median earnings in the area by 4.35% to allow the individual to climb up 1.240 (= 1/0.806) percentage points in the earnings rank.

poorer peers, the opposite is true for their U.S. respondents.

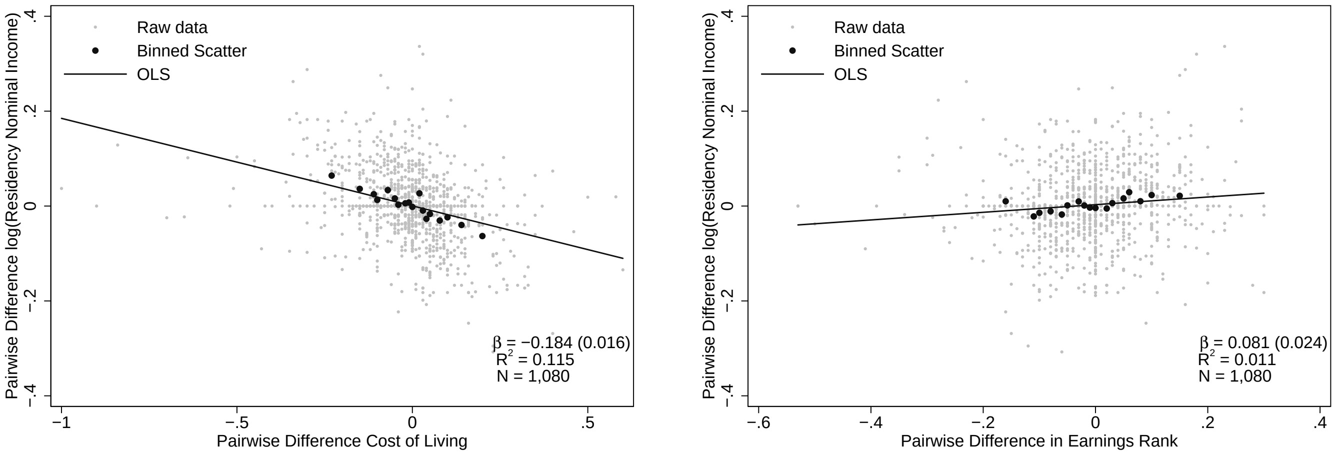
D.13 Comparing Happiness and Choice Trade-Offs

We can also exploit a different outcome variable, the happiness rank between the options, to compare the preferences inferred from choice versus happiness. Consistent with Benjamin et al. (2014), we observe a significant correlation (0.456) between the choice ranks and happiness ranks of these individuals. However, this association is far from perfect, which suggest that individuals are not choosing to maximize their happiness only. As a result, it is not obvious that preferences inferred from choice will be similar to preferences inferred from happiness.

Table D.9 presents results using happiness as outcome variables. These coefficients are of course not directly comparable to those of choice, because they are based on different dependent variables with different distributions. The baseline preferences are roughly consistent. For instance, for the full sample, β^{rel} is 0.989 (s.e. 0.540) for choice and 0.957 (s.e. 0.517) for happiness; while β^{abs} is 1.080 (s.e. 0.484) for choice and 0.401 (s.e. 0.478) for happiness. We cannot reject the null hypotheses that these two pairs of coefficients are equal. This evidence suggests that the happiness and choice trade-offs may be similar – however, given the precision of the estimates, we cannot reject the possibility of substantial discrepancies.

Figure D.1: Variation in Nominal Income, Cost of Living and Earnings Rank

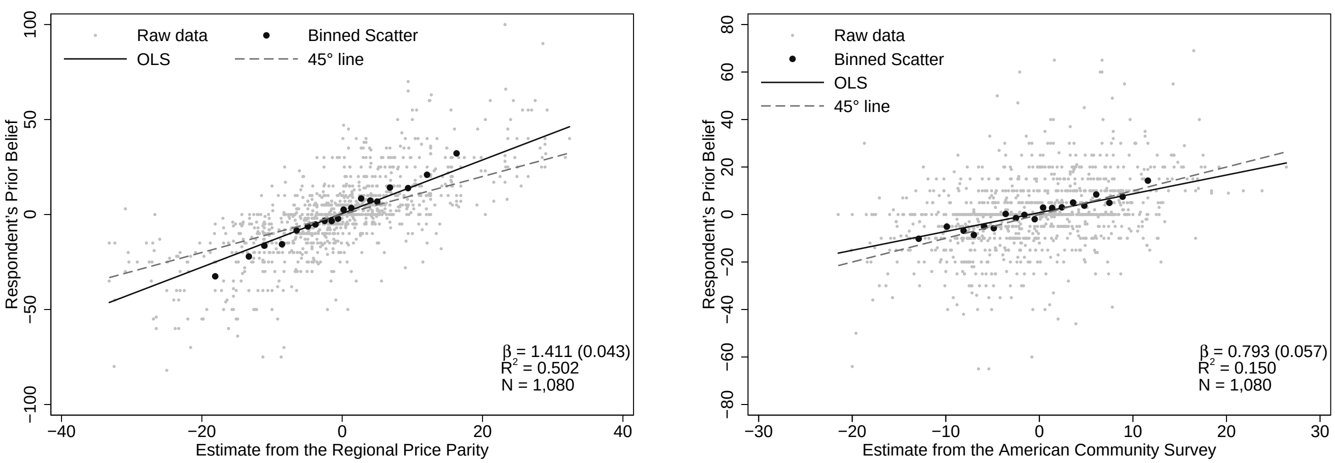
a. Cost of Living vs. Nominal Income **b.** Earnings Rank vs. Nominal Income



Notes: The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Slopes (β, with robust standard errors in parentheses) and R^2 are based on a linear regression. All variables for x-axis and y-axis correspond to pairwise differences across the two cities that the subject is considering submitting to the algorithm. Data from survey responses, the Regional Price Parity Index (for cost of living) and the American Community Survey (for earnings rank).

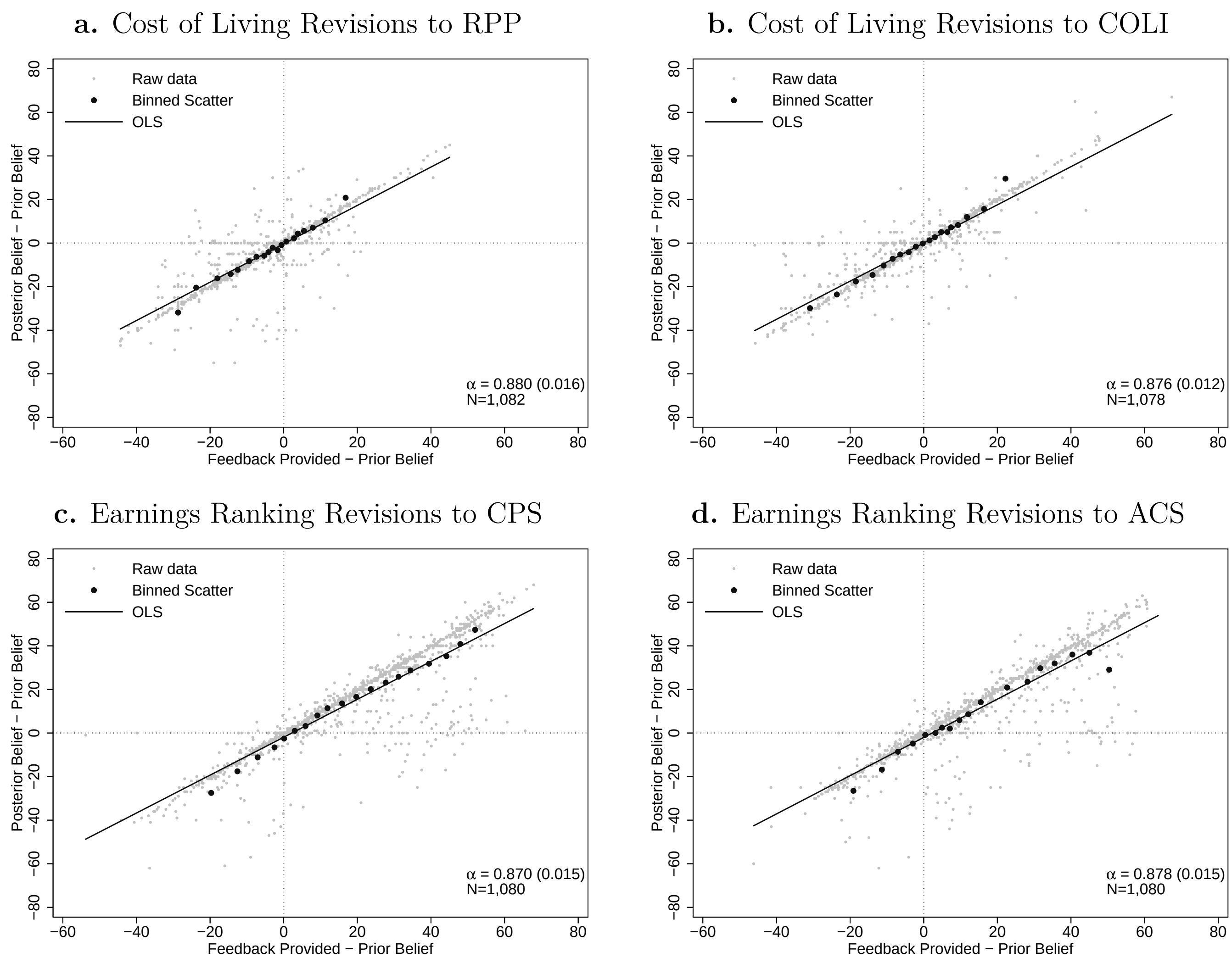
Figure D.2: Comparison Between Prior Beliefs and Statistics

a. Cost of Living, Pairwise Differences **b.** Earnings Rank, Pairwise Differences



Notes: Comparison between respondent's perceptions before the information provision (i.e., prior beliefs) and statistics. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panels a and b present pairwise differences between an individual's options (i.e., value for first option minus that of the second option). The slope (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

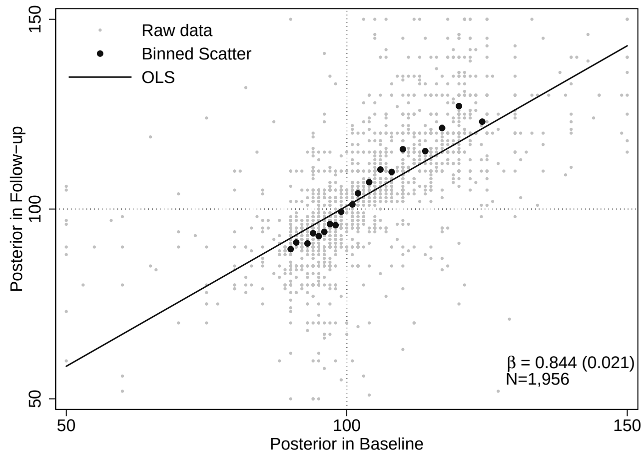
Figure D.3: Reduced-Form Evidence of Learning in the Information-Provision Experiment by Feedback Source



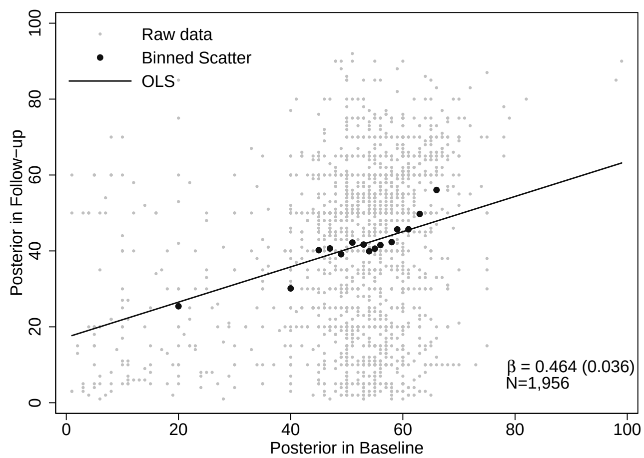
Notes: Comparison between the difference in statistics and respondent's perceptions before the information provision (i.e., prior beliefs), and difference in respondent's perceptions after the information provision (i.e., posterior beliefs) and prior beliefs. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panels a and b show cost of living revisions to statistics from RPP (Regional Price Parity Index) and COLI (Cost of Living Index). Panels c and d show earnings rank revisions to statistics from CPS (Current Population Survey) and ACS (American Community Survey). The slope (α, with robust standard errors in parentheses) is based on a linear regression.

Figure D.4: Correlation between (Posterior) Beliefs in Baseline and Follow-Up Surveys

a. Cost of Living



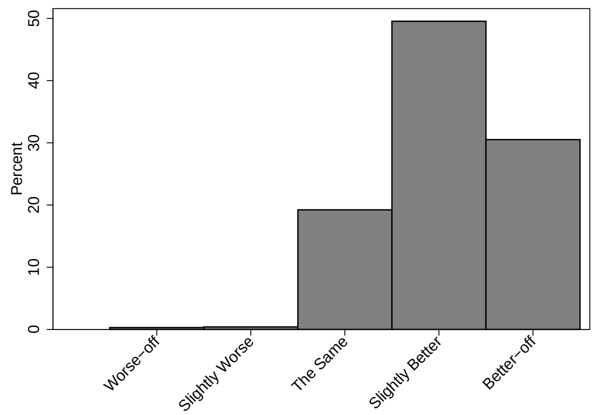
b. Earnings Rank



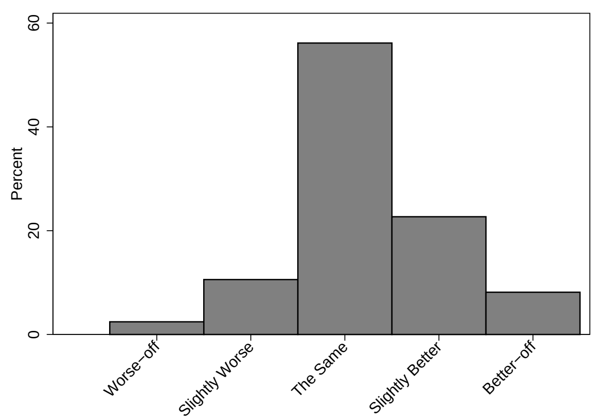
Notes: The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panels a and b present data in levels (i.e., two observations per individual, one for each of their options). The slope (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

Figure D.5: Survey Responses, Preferences over Hypothetical Changes to Absolute and Relative Consumption

a. Hypothetical Increase in Absolute Consumption



b. Hypothetical Increase in Relative Consumption



<u>Notes</u>: Distribution of responses to hypothetical choice questions included in follow-up survey. Panel b corresponds to the question labeled "Event A", while panel b corresponds to the question labeled "Event B" in the questionnaire to the follow-up survey in Appendix A.2.

Table D.1: Probit Marginal Effects

	Panel A: β^{rel}			Panel B: β^{abs}		
	All (1)	Non-Single (2)	Single (3)	All (4)	Non-Single (5)	Single (6)
Baseline Sample						
Raw Probit	0.989*	2.195***	-1.527*	1.080**	1.095*	1.042
	(0.540)	(0.670)	(0.875)	(0.484)	(0.658)	(0.750)
Marginal Effect	0.185*	0.412***	-0.265*	0.202**	0.206*	0.181
	(0.100)	(0.125)	(0.154)	(0.090)	(0.123)	(0.130)
Follow-up Sample						
Raw Probit	1.130*	2.337***	-1.666*	1.271**	1.230*	1.401*
	(0.578)	(0.703)	(0.995)	(0.529)	(0.739)	(0.780)
Marginal Effect	0.201**	0.419***	-0.255*	0.226**	0.221*	0.214*
	(0.102)	(0.126)	(0.154)	(0.093)	(0.132)	(0.119)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Raw Probit coefficients and corresponding marginal effects at the mean. Probit regressions of expected rank order submission on posterior beliefs about cost of living and earnings rank estimated by sample (i.e., coefficients of a same row and sample are from a single regression). All specifications include the baseline controls listed in section 3. Results for Baseline Sample are based on the sample of individuals who completed the baseline survey (1,080 responses, 698 from non-singles and 382 from singles). Results for Follow-up Sample are based on the sample of individuals who completed the follow-up survey (978 responses, 647 from non-singles and 311 from singles).

Table D.2: Preferences for Subjective Program Characteristics

	(1)	(2)	(3)	(4)
β^{rel}	1.130*	1.159**	1.158*	1.137*
	(0.578)	(0.585)	(0.604)	(0.612)
β^{abs}	1.271**	1.422***	1.482***	1.228**
	(0.529)	(0.513)	(0.523)	(0.525)
$\beta^{purpose}$		0.437***		
		(0.064)		
$\beta^{prospects}$			0.378***	
			(0.070)	
$\beta^{prestige}$				0.248***
				(0.061)
Pseudo R^2	0.035	0.058	0.093	0.118
Observations	978	978	978	978

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit coefficients. Probit regressions of expected rank order submission on posterior beliefs about cost of living and earnings rank estimated by sample (i.e., coefficients of a same row and sample are from a single regression). All estimates, include the baseline controls listed in section 3. Mean (standard deviation) for $ER_{1,2}^{i,posterior}$ is -0.008 (0.098) and for $COL_{1,2}^{i,posterior}$ is 0.010 (0.138). Measures for subjective program characteristics (prestige, prospects, purpose) are standardized to have mean zero and standard deviation of one.

Table D.3: Preference Heterogeneity with Respect to Marital Status: Additional Results

	Non-Single		Non-Single		Single	
	Married (1)	LT Relationship (2)	Female (3)	Male (4)	Female (5)	Male (6)
β^{rel}	1.999*	2.271***	2.693***	1.729*	-2.372*	-1.022
	(1.180)	(0.843)	(0.956)	(0.975)	(1.287)	(1.315)
β^{abs}	2.401**	0.396	1.111	1.369	0.670	1.631
	(0.999)	(0.821)	(0.993)	(0.949)	(0.804)	(1.295)
Diff. P-value *[q-value]:*						
Relative	0.851 *[0.955]*		0.480 *[0.841]*		0.463 *[0.841]*	
Absolute	0.121 *[0.684]*		0.851 *[0.841]*		0.528 *[0.955]*	
Pseudo R^2	0.093	0.045	0.072	0.052	0.054	0.027
Observations	259	439	360	338	200	182

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Each column corresponds to a separate Probit regression. Coefficients for Probit regressions of expected rank submission (at baseline) on relative and absolute consumption (measured by posterior beliefs in baseline survey), and controls (e.g. relative wage, etc.) as defined in section 3. P-values corresponds to the test of the null hypothesis that the coefficients are equal between the two sub-groups. Multiple-testing q-values based on Benjamini and Yekutieli (2001) presented in brackets.

Table D.4: Preference Heterogeneity with Respect to Other Individual Characteristics

	Hypothetical increase absolute consumption		Hypothetical increase relative consumption		By Materialism		By Competitiveness		By Life Dimension	
	Better off (1)	Same/Worse off (2)	Better off (3)	Same/Worse off (4)	High (5)	Low (6)	High (7)	Low (8)	High (9)	Low (10)
β^{rel}	1.688**	-0.684	1.800*	0.884	1.698**	0.719	1.229*	0.835	1.657*	0.637
	(0.670)	(1.186)	(1.017)	(0.731)	(0.708)	(0.931)	(0.665)	(1.216)	(0.908)	(0.788)
β^{abs}	1.198**	1.486	1.698**	1.113	0.638	2.239***	1.659***	-0.079	1.027	1.928**
	(0.585)	(1.154)	(0.807)	(0.715)	(0.756)	(0.746)	(0.606)	(0.930)	(0.747)	(0.858)
Diff. P-value *[q-value]:*										
Relative	0.081 *[0.471]*		0.464 *[0.745]*		0.402 *[0.741]*		0.776 *[0.918]*		0.396 *[0.741]*	
Absolute	0.824 *[0.943]*		0.587 *[0.865]*		0.132 *[0.559]*		0.117 *[0.558]*		0.428 *[0.745]*	
Pseudo R^2	0.042	0.070	0.131	0.031	0.036	0.061	0.043	0.043	0.046	0.059
Observations	782	194	299	677	516	460	750	226	508	468

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Each panel corresponds to a separate Probit regression. Coefficients for Probit regressions of expected rank submission (at baseline) on relative and absolute consumption (measured by posterior beliefs in baseline survey), and controls (e.g. relative wage, etc.) as defined in section 3. All controls are interacted with indicator variable for heterogeneity variable indicated in panel. In panels c and d, respondents are classified as high/low using the median scores for the competitiveness index (16/30) and materialism index (21/30). Life dimension is based on a principle-component index of rank 5 life dimensions (happiness, health, sense of purpose, spirituality, control over life) that was divided at the median. P-values corresponds to the test of the null hypothesis that the coefficients are equal between the two sub-groups. Multiple-testing q-values based on Benjamini and Yekutieli (2001) presented in brackets.

Table D.5: Robustness to Sample Definition

	Panel A: β^{rel}			Panel B: β^{abs}		
	All (1)	Non-Single (2)	Single (3)	All (4)	Non-Single (5)	Single (6)
Baseline Sample	0.989*	2.195***	-1.527*	1.080**	1.095*	1.042
	(0.540)	(0.670)	(0.875)	(0.484)	(0.658)	(0.750)
Pass Attention Check	1.072**	2.205***	-1.373	1.091**	0.942	1.266
	(0.543)	(0.682)	(0.891)	(0.496)	(0.673)	(0.776)
Drop Dual Matches	1.001*	2.125***	-1.300	1.122**	1.091	1.117
	(0.551)	(0.699)	(0.850)	(0.493)	(0.664)	(0.772)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit coefficients. Probit regressions of expected rank order submission on posterior beliefs about cost of living and earnings rank estimated by sample (i.e., coefficients of a same row and sample are from a single regression). All estimates, include the baseline controls listed in section 3. The first row shows estimates for baseline sample (1,080 responses, 698 from non-singles and 382 from singles). The second row restricts the sample to respondents who pass the attention check question in baseline survey (1,041 responses, 678 from non-singles and 363 from singles), while the third row restricts the sample to respondents who are not participating as dual match (1,000 responses, 641 from non-singles and 359 from singles).

Table D.6: Binary Probit vs. Ordered Probit

	Panel A: β^{rel}			Panel B: β^{abs}		
	All	Non-Single	Single	All	Non-Single	Single
	(1)	(2)	(3)	(4)	(5)	(6)
Probit	0.989*	2.195***	-1.527*	1.080**	1.095*	1.042
	(0.540)	(0.670)	(0.875)	(0.484)	(0.658)	(0.750)
Ordered Probit	0.728*	1.336***	-0.305	0.568*	0.841**	0.088
	(0.373)	(0.475)	(0.596)	(0.310)	(0.401)	(0.490)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit (or Ordered Probit) coefficients. Probit (or Ordered Probit) regressions of expected rank order submission on posterior beliefs about cost of living and earnings rank estimated by sample (i.e., coefficients of a same row and sample are from a single regression). All estimates, include the baseline controls listed in section 3. All results based on the sample of individuals who completed the baseline survey (1,080 responses, 698 from non-singles and 382 from singles).

Table D.7: IV, First Stage, and Reduced Form Estimates

	All	Non-Single	Single
	(1)	(2)	(3)
Panel A: IV-Probit Estimates			
β^{rel}	0.858	2.955**	-4.984**
	(1.150)	(1.332)	(1.950)
β^{abs}	-0.653	-0.336	-1.660
	(0.880)	(1.163)	(1.288)
Panel B: First Stage			
Dep. Var.: $ER^i_{1,2}$			
$\Delta ER^i_{1,2}$	0.796***	0.854***	0.687***
	(0.045)	(0.055)	(0.081)
$\Delta COL^i_{2,1}$	-0.013	-0.021	-0.007
	(0.039)	(0.049)	(0.064)
Dep. Var.: $COL^i_{2,1}$			
$\Delta ER^i_{1,2}$	0.058	0.101***	-0.036
	(0.037)	(0.036)	(0.087)
$\Delta COL^i_{2,1}$	0.928***	0.893***	0.985***
	(0.048)	(0.064)	(0.070)
Wald test of exog. p-val.	0.062	0.334	0.004
Cragg-Donald F-stat.	207.402	172.225	42.998
Panel C: Reduced Form			
$\Delta ER^i_{1,2}$	0.655	2.484**	-3.512**
	(0.918)	(1.153)	(1.577)
$\Delta COL^i_{2,1}$	-0.711	-0.485	-1.768
	(0.845)	(1.067)	(1.385)
Observations	978	647	331

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit (or IV-Probit) coefficients. All regressions include the baseline controls listed in section 3. The independent variables are the posterior beliefs about cost of living and earnings rank, from the baseline specification. Panel A presents raw IV-Probit estimates using model detailed in section 3.2. Panel B shows the first stage for each independent variable. Panel C shows reduced form Probit estimates.

Table D.8: Preference Estimates from Other Studies with Happiness and Hypothetical Data

Reference	Evidence	Country	Parameters	Source	$\frac{a-b}{b}$
Luttmer (2005)	Happiness	U.S.A.	a=0.361, b=0.296	Column (3) of Table 1	0.22
Ferrer-i-Carbonell (2005)	Happiness	Germany	a=0.456, b=0.226	Column (1) of Table 2	1.02
Clark, Senik and Yamada (2016)	Happiness	Japan	a=0.290, b=0.153	Column (1) of Table 3	0.89
Johansson-Stenman, Carlsson and Daruvala (2002)	Hypothetical	Sweden	b/a=0.35	Page 373	1.85
Yamada and Sato (2013)	Hypothetical	Japan	a=0.048, b=0.022	Column (1) of Table 4	1.18

Notes: Authors calculations based on the regression coefficients reported in the papers.

Table D.9: Preferences Inferred from Happiness vs. Choice

	Panel A: β^{rel}			Panel B: β^{abs}		
	All (1)	Non-Single (2)	Single (3)	All (4)	Non-Single (5)	Single (6)
Baseline	0.957*	1.512**	0.012	0.401	0.772	-0.398
	(0.517)	(0.629)	(0.948)	(0.478)	(0.618)	(0.766)
Experimental	1.752*	2.968***	-1.648	-0.474	-0.066	-1.300
	(0.965)	(1.083)	(2.017)	(0.792)	(1.041)	(1.243)
Experimental, Long Term	1.314	2.831***	-2.232	0.039	-0.729	1.169
	(0.976)	(1.067)	(2.171)	(0.761)	(0.946)	(1.217)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Coefficients for Probit regressions of variable indicating that respondent would live happier life at location 1 (at baseline, or at follow-up for "long term") on relative and absolute consumption (measured by posterior beliefs in baseline survey), and controls (e.g. relative wage, etc.) as defined in section 3. All results based on the sample of individuals who completed the follow-up survey (978 responses, 647 from non-singles and 311 from singles).