Online Appendix: For Online Publication Only

Choosing Your Pond: Location Choices and Relative Income

Nicolas L. Bottan & Ricardo Perez-Truglia

A Further Information and Analysis

A.1 Estimation of the Feedback Provided to Subjects

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.[1] 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.[2] 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,

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

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

and the correlation is 0.91. Although both sources are similar in levels, there is plenty of exogenous variation between them when comparing pairwise differences of chosen locations. We show this variation in Figure A.1.a, where the R-squared of regressing the pairwise differences for the ACS on the pairwise differences for the CPS is 0.430.

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 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 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 A.1.b, where the R-squared for regressing the pairwise differences of RPP on COLI is only 0.436.

A.2 Descriptive Statistics and Randomization Balance

Table A.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 male, the average age was 27 years, 35.4% of respondents were single, 23.9% were married, and 40.7% 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.[3] To verify that the randomization was successful, Table A.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. Last, Figure A.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.

A.3 Further Information about the Subject Pool

We recruited 27 of the 135 accredited medical schools in the U.S. to participate in our study (see geographical distribution of participating medical schools in Figure A.3). 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

[3]For more details, see Appendix A.4.

ranks for universities and hospitals). We present descriptive statistics for the universe of medical schools, non-participating and participating schools in Table A.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 A.3, 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 A.4, 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 A.4 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).

A.4 Comparing Experimental Subjects to U.S. Population of Earners

In Table A.5 we present average characteristics of medical students in our sample to the U.S. population of earners obtained from the 2015 American Community Survey PUMS. On average, medical students are significantly younger (27 vs 41 years old) and slightly more likely to be a female. Additionally, medical students are half as likely to be married than the average earner. During the residency, medical students on average expect to earn a wage that is close to that of the average earner in the U.S.

A.5 Variation in Nominal Income, Earnings Rank and Cost of Living

Residency programs do not compensate for differences in cost of living or earnings rank through wages. Figure A.5.a 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.

In a similar spirit, Figure A.5.b presents a scatterplot of the the costs of living versus the (log) nominal residency income. The strength of the association between these two vari-

ables 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.

We use one possible definition of earnings rank: the position in the distribution of individual income among all income earners. There are alternative ways in which we could define this rank that the subjects could find more or less relevant. However, in practice this may matter little to the extent that all of the different definitions may be highly correlated. For example, Figure A.6 shows the correlation between the measure we use in this study versus an alternative definition: the position in the distribution of individual income among all working-age individuals (between 18 and 64 years of age) regardless of whether they earn some income or not. This figure shows that these two measures of relative income are highly correlated. The $R^2 = 0.890$ suggests that the results would have been probably identical had we used the alternative definition of earnings rank.

A.6 Accuracy of Prior Beliefs, Pairwise Differences

Respondents may have a poor idea of the levels of earnings rank and cost of living, 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 A.7.a for earnings rank and Figure A.7.b for cost of living. 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.

We must note, however, that the elicitation of prior beliefs was not incentivized. When designing the survey, we decided not to introduce incentives for the following reasons. On the one hand, the benefit from incentivizing beliefs is that it may lead to more truthful elicitations. The main worry is usually that subjects may lie if they are not provided with

incentives to tell the truth. In other contexts, this can be crucial. For example, if individuals are asked about whether they donated to charity, voted or helped a friend, there is evidence that a non-negligible fraction lie unless provided with incentives. Presumably, the reason for these lies is social desirability bias: a subject who did not vote may want to lie and say that he or she voted to appear like a good citizen in the eyes of the interviewer. While we know this social desirability bias can be quite perverse in other contexts, we thought it would be a minor nuisance in our context. Unlike giving to charity or turning out to vote, it is doubtful the surveyor could judge responses on cost of living and relative income. In the case of giving to charity or turning out to vote, the surveyor does not know the truth (e.g., whether the subject voted or gave to charity), while in our context the surveyor clearly knows the truth (i.e., the actual cost of living or income distribution) so there is no point in lying.

Additionally, social desirability bias might be a concern if we were only interested in identifying subject's beliefs about the levels of earnings ranking for each city. However, our identification relies on differences in these perceptions between cities. Therefore, this would only be a concern if such bias were stronger for one city, but not for the other – which seems even less likely. Last but not least, our survey is a confidential survey conducted online, which has been shown to be less subject to the social desirability concerns (as opposed to, for example, phone surveys or face-to-face surveys in which the "presence" of the surveyor is more tangible).

On the other hand, there would have been a downside from incentivizing the question: we would need to ask subjects to guess what the statistics say about cost of living, instead of asking them what they expect their own cost of living to actually be. It is possible that a subjects think that the statistics say the cost of living in city A is X, but they expect that their own cost of living in city A would be different. This is not only because the surveyor may distrust that the cost of living statistics reflect the true cost of living, but most importantly because the cost of living statistics are designed to reflect population averages and the respondent may not think of himself or herself as average. For example, you may

think that in city A the average cost of living is higher but it would be even higher for you because you are married, or have kids, or have some special consumption needs. In our choice model, we need a measure of the cost of living that the subject truly expects, instead of the subject's perception about the statistics. For that reason, the thought we would be better off by not incentivizing this question. On top of that, if we had incentivized the questions then would not been able to have a first stage for the regression because subjects would have had a strong incentive to just regurgitate back to us exactly the feedback that they received.

A.7 Estimating Learning Rates

We begin by demonstrating that the degree of learning can be estimated from the relationship between the initial perception gaps (i.e., the signal received minus the prior belief) and the belief revisions (i.e., the posterior belief minus the prior belief).[4] 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 ($b_k^{signal} -$

[4]For a discussion about the estimation of learning models with survey experiments, see Armantier et al. (2016) and Cavallo et al. (2017). Also, in relation to the identification of preferences from information-provision experiments, see Wiswall and Zafar (2014) and Cullen and Perez-Truglia (2018).

b_k^{prior}).

One concern with our experimental design is that individuals may have updated their beliefs diferentially 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 A.8 we present the same figures as in section 5.2 by information source. Learning rates for earnings ranking and cost of living are almost identical between sources, showing that respondent's reactions to information did not depend on the source.

A.8 Marginal Effects

Given that it is not possible to directly interpret coefficients from Probit regressions, in Table A.6 we present estimates from columns (1) to (3) from Table 1 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 (3) for the baseline sample implies that an increase of 1 percentage point in earnings rank in location 1 would increase the probability of choosing that location by 0.186 percent (or, in other words, an implied behavioral elasticity of 0.186).

A.9 Robustness 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 earnings rank and cost of living. In the first row of Table A.7 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 C.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 A.7 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.

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 relative income and cost of living across choices. Our survey data indicates that no more than 6% of individuals tried to select the same metro area.[5] For those subjects, the comparison was between two of their top programs but not necessarily the top two. In Table A.8 we show that the results are robust if we exclude those individuals from the sample. Note that point estimates are very similar, both qualitatively (i.e., sign and statistical significance) and quantitatively. This should not be surprising, as the group that we exclude constitutes such a small share of the sample.

Another potential concern is whether the results are driven by a few specific metro areas. We address those concerns in Table A.9. The first row corresponds to the baseline specification. The second row excludes the three largest metro areas in the country: New York, Los Angeles and Chicago. A subject is dropped if either of the top-2 residency programs is located in one of those 3 metro areas, which results in a substantial drop in the number of observations: from 1,080 observations in the baseline specification to 814 in the second row. Despite the significant loss in statistical power, the coefficients on relative income are, if anything, larger in magnitude and more statistically significant. In the third row, we keep the largest metro areas but drop the smallest metro areas instead. Dropping the smallest 3

[5]These are identified based on the number of clicks recorded on that question relative to the identical previous question. If the cities matched, they had to re-answer the questions which would require at least 5 additional clicks. We identify around 6% of individuals who met this criteria.

metro areas makes little sense as the data would barely change. For that reason, we include a more aggressive approach, we drop the smallest 25 metro areas (which have less than 160,000 people living in them). Again, a subject is dropped if either of the top-2 residency programs is located in one of those 25 smallest metro areas. This results in a still significant, but more reasonable, drop in the number of observations: from 1,080 observations in the baseline specification to 1,023 in the third row. The coefficients on relative income are again qualitatively and quantitatively consistent with those from the first row.

A.10 Robustness to Additional Controls: Hospitals and Residency Programs

In this section, we provide some additional details about the quality of residency programs as control variables.

We start by simply showing the raw correlation between the variables of interest (earnings rank and cost of living) and a simple measure of the quality of the residency program: the residency program's percentile ranking as measured by Doximity. The results are presented in Figure A.9. Figure A.9.a corresponds to the relationship between the earnings rank and the residency program quality. The $R^2 = 0.038$ implies that there is plenty of orthogonal variation between these two characteristics. In turn, Figure A.9.a corresponds to the association between the residency reputation and cost of living. This correlation is quite low too: the $R^2 = 0.001$ implies that these two characteristics are almost perfectly orthogonal.

Given the low correlations presented above, it should come at no surprise that adding residency program quality as control variable does not affect the coefficients on earnings rank and cost of living. One remaining concern, however, is that maybe we are not focusing on the "correct" variable. To address this concern, Table A.10 shows the results with different sets of control variables. The first row presents our baseline results. The second row controls for the supply or availability of residency programs in each city by including the relative number of programs (for the selected specialty) in the metro areas of choice. The third row also controls for the overall quality of residency programs in the metro areas by including

the relative average percentile ranking and controlling for the share of programs ranked in the top 10th percentile. In addition to these variables, the fourth row includes variables controlling for objective and subjective measures of quality for the *chosen* programs (as in Table 2). Notably, estimates are quite stable across each row.

At the same time, medical students may not care so much about residency program characteristics themselves, but rather *hospital* characteristics. Data on hospital characteristics were obtained from the CMS Provider of Services File and CMS Hospital Compare data for 2016. In the fifth row we estimate the baseline model including the relative number of hospitals in the metro areas selected. The sixth row controls for various measures of hospital quality in the metro area (relative metro area averages of mortality rate, readmission rate and overall patient satisfaction score; and shares of hospitals in the metro area that are top 10th percentile nationally for the respective measure). Again, the seventh row adds controls for objective and subjective measures of residency quality (as in row 4). The final row includes all variables described above. Again, point estimates and significance remains quite similar (if anything, magnitudes are slightly larger). This suggests that the coefficients on earnings rank and cost of living are not capturing preferences for residency program or hospital quality.

A.11 Preferences over Subjective Program Characteristics

To better understand the magnitude of our results, we compare the estimates for preferences over earnings rank and cost of living 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 A.11. 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 cost of living.[6]

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 decrease in earnings rank. According to column (2) of Table A.11, a one standard deviation increase in earnings rank would correspond to a Probit coefficient of 0.112 (i.e., the non-standardized coefficient, 1.147, multiplied by the standard deviation of earnings rank, 0.098). This standardized coefficient can be compared to the coefficient of 0.441 corresponding to a one standard deviation increase in the sense of purpose. This comparison implies that the sense of purpose of a program is 3.9 times as important as the earnings rank. By the same metric, the career prospects (column (3)) and sense of prestige (column (4)) are 3.4 and 2.2 times as important as earnings rank. In sum, the characteristics of a program are systematically more important for the choice of residency than relative income during the residency.

A.12 Other Forms of Heterogeneity in Preferences

In this section we explore additional heterogeneity over preferences for earnings rank and cost of living. We first decompose the results of heterogeneity by relationship status in two ways. In columns (1) and (2) of Table A.12, we show that within non-single respondents, preferences over relative income are similar for married or long-term relationship respondents. However, it seems that preferences for cost of living are mostly driven by married respondents (though the difference is borderline insignificant, p-value=0.109). In columns (3) to (6) of Table A.12, we estimate preferences by gender, within relationship status. Preferences over earnings rank 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

[6]The results are similar if we do the comparison with respect to the preferences for earnings rank instead.

across different dimensions in Table A.13. In columns (1) to (4) we explore heterogeneity according to differences in hypothetical choices of changes in earnings rank and cost of living. Interestingly, we find that those who believe they would be better off if cost of living were to decrease care significantly more about earnings rank 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 earnings rank.

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 C.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 earnings rank, while those who are less "materialistic" care more about cost of living. 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.

A.13 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, such as Luttmer (2005).[7] Let absolute consumption be the nominal earnings divided by the cost of living index, and let relative consumption be the individual's rank in the distribution of absolute consumption in the same city. If the

[7]Luttmer (2005) is particularly important because it uses data for the United States and is then the most comparable sample.

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 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. As a result, the ratio $-\frac{\beta^{ER}}{\beta^{COL}}$ would correspond to the marginal rate of substitution between relative consumption and absolute consumption.

Luttmer (2005) and 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 A.14 shows the estimates of a and b reported in other studies, and the resulting estimate of $\frac{a-b}{b}$.[8]

The key specification from Luttmer (2005), which is estimated on the sample of non-single individuals, implies that most of the utility from consumption derives from relative consumption rather than absolute consumption: non-single individuals are willing to give up 1% of absolute consumption to decrease the median consumption of neighbors by 0.22%. Accord-

[8]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).

ing to our baseline estimates from column (1) of Table 3, non-single individuals are willing to give up 1% of their absolute consumption to decrease the median consumption of their peers by 0.91% (90% confidence interval: [−0.18%, 2.05%]).[9] Compared to Luttmer (2005), our baseline estimates suggest a weaker role for relative concerns; however, this difference is not statistically significant. Compared to Luttmer (2005), the results from our auxiliary experiment also suggest a weaker role of relative concerns, but this time with a statistically significant difference: our auxiliary estimates suggest that subjects are willing to give up 1% of absolute consumption to decrease the median consumption of their peers by 2.79% (90% confidence interval: [1.13%, 4.52%]).[10] Assuming that the estimates from Luttmer (2005) reflect the real degree of neighbor externalities, our findings suggest that individuals anticipate, at least partially, the negative externalities from more affluent neighbors.

We can also provide comparisons with respect to other studies. According to our baseline estimates for non-singles (column (1) of Table 3), the average individual is willing to give up 1 percent of her absolute consumption to decrease the median consumption of her peers by 0.91%.[11] The other studies that use happiness data suggest a corresponding trade-off of 0.89% (Clark et al., 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, 2016). All of these estimates are in the ballpark of our own estimate of 0.91%, implying that, relative to these other studies, our estimates suggest a similar role for relative concerns.

Last, we must note that some studies find the opposite effect. For instance, Senik (2004)

[9]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 0.519 (= 1/1.928) percentage points in the earnings rank.

[10]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. Furthermore, we would need the standard errors from Luttmer (2005) to directly compare with our estimates.

[11]This result arises because, for the average individual in the sample, we would need to decrease the median earnings in the area by 0.91% to allow the individual to climb up 0.518 (= 1/1.928) percentage points in the earnings rank.

and Clark et al. (2009) find that life satisfaction is increasing in the mean income of the reference group. And Yamada and Sato (2016) show estimates from a hypothetical choice experiment with mixed results: while the U.K. respondents prefer poorer peers, the opposite is true for their U.S. respondents.

A.14 Estimating Happiness 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 A.15 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^{ER} is 0.995 (s.e. 0.539) for choice and 0.957 (s.e. 0.516) for happiness; while β^{COL} is -1.073 (s.e. 0.485) for choice and -0.403 (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.

A.15 Beliefs are Persistent in Follow-up Survey

Since posterior beliefs on earnings ranking and cost of living 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 earnings rank and cost of living in Figure A.10. Indeed, the overall persistence of beliefs suggests that these

individuals were still incorporating a lot of relevant information during this period of time; for example, a 1 percentage point increase in the (posterior) belief about the relative earnings in the baseline survey is associated with 0.46 percentage points higher belief in the follow-up survey. The corresponding association for cost of living is 0.84.

A.16 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 C.2).

In one scenario, we elicited the subjects' preferences for an increase in earnings rank, while holding the cost of living constant. 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 income, and thus these responses probably lead to an underestimation of concerns for relative income. Figure A.11.a shows the distribution of responses. Consistent with individuals having direct preferences over relative income, 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.

In the alternative scenario, we elicited the subjects' preferences for a reduction in cost of living while holding the earnings rank constant. More precisely, we asked the respondents whether they would be better off, the same, or worse off if their own cost of living and the cost of living of all other individuals in the city went down by 10%. Figure A.11.b 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.

A.17 Robustness to Alternative Regression Models

All the baseline results are based on a simple Binary Probit model. In practice, the results are similar when based on alternative models. This robustness check is presented in Table A.16. Each row corresponds to the same regression but based on different models: Binary Probit, Binary Logit, Ordered Probit and Linear Probability Model.[12] In addition to the raw coefficients, the Binary Probit and Binary Logit specifications report the marginal effects estimated at the means of the independent variables, to make a more quantitative comparison easier. This table shows that the results are very similar regardless of the specific model used.

A.18 Details about the Instrumental Variable Regression

We break down the Instrumental Variables regression into the first-stage and reduced-form regressions. Table A.17.a presents the same experimental estimates as those found in the second row of Table 3. In the next panel we focus on the first stages. As discussed in section 5.2, 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 169 to 42. In the final panel of Table 3 we show that the reduced form estimates are similar to those obtained by IV.

A.19 Robustness Check: Using Relative Income to Learn about Cost of Living

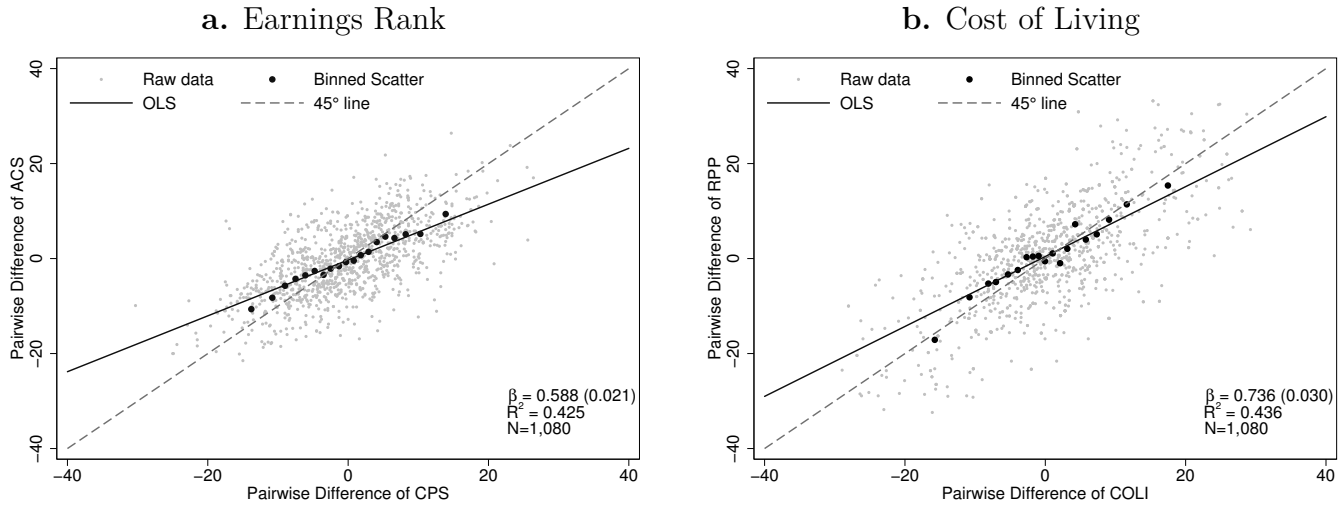
In Figure A.12 we use similar learning regressions from Section 5.2, only that we measure the effect of relative income feedback on beliefs about cost of living (and vice versa). Figure

[12] 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. In contrast, the Ordered Probit model exploit the full variation given by the likelihood scale.

A.12.a shows the effect of earnings rank feedback on posterior beliefs about the cost of living from the baseline survey (i.e., the short-term effect). The slope is close to zero (-0.003), precisely estimated (s.e. 0.006), and statistically insignificant. This coefficient suggests that increasing the observed earnings rank by 1 percentage point reduces posterior beliefs about cost of living by 0.003 percentage points. To put this magnitude in context, Figure 2.d suggests that the effect of earnings rank on posterior beliefs about earnings rank is 0.873 (s.e. 0.011). The difference between this 0.873 and the -0.003 effect is economically large and statistically highly significant. Figure A.12.b shows the effect of earnings rank feedback on beliefs about the cost of living from the follow-up survey (i.e., the long-term effect). Again, the effect is close to zero (-0.011), precisely estimated (s.e. 0.011), and statistically insignificant.[13]

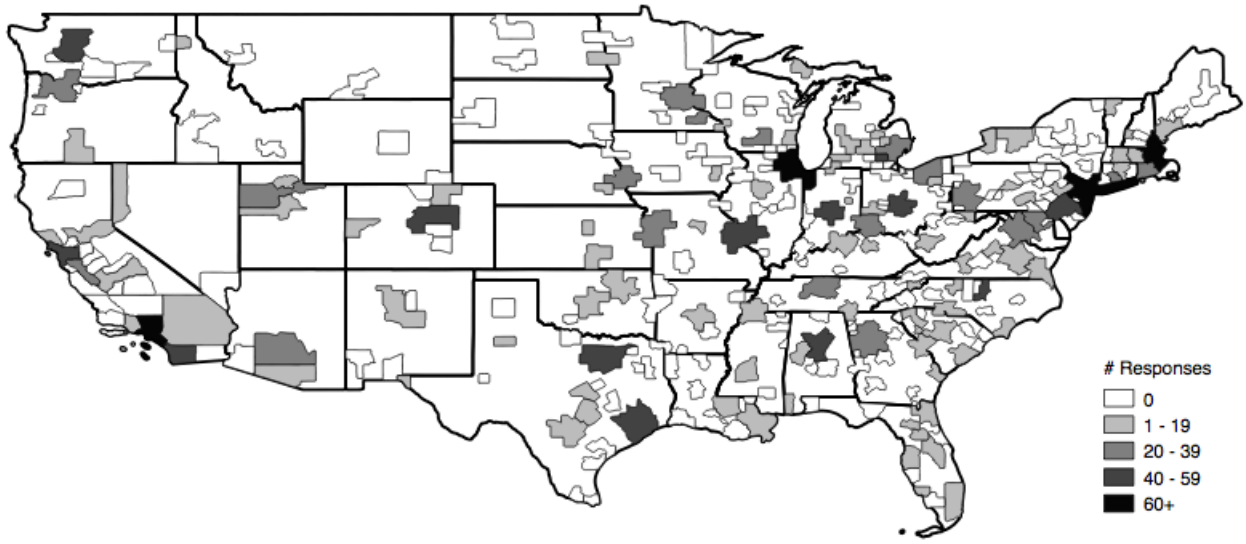
[13]As additional evidence that subjects see relative income and cost of living as two distinct features of the city, Figure A.12.c and A.12.d show that the converse also is true: feedback about cost of living does not affect short-term or long-term beliefs about relative income.

Figure A.1: Comparison of Feedback Across the Two Data Sources



Notes: Pairwise differences of statistics from different sources of earnings rank and cost of living 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.

Figure A.2: Geographic Distribution of Choice Set

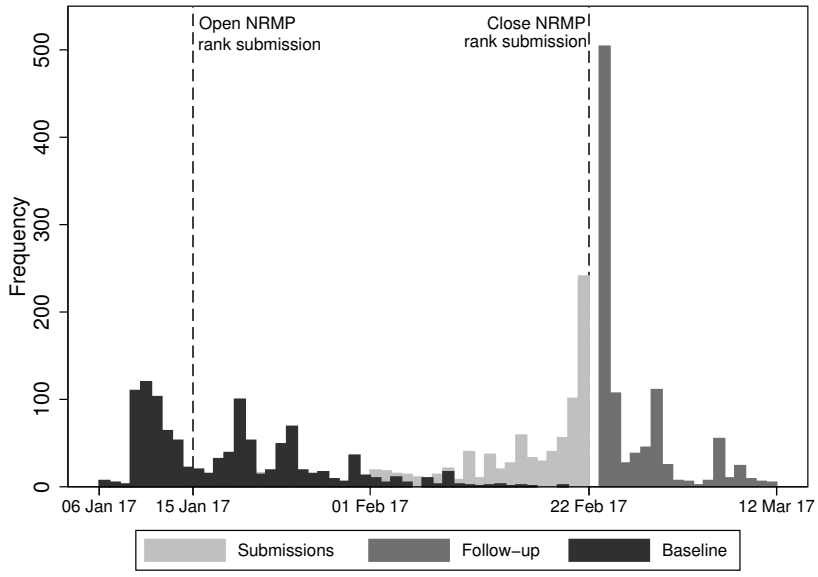


Notes: Geographical distribution of metropolitan areas where top-2 residency programs of respondents are located, for the continental United States. No responses were located in Hawaii, while Alaska only has 2 responses. Only metropolitan areas with a residency program participating in the 2017 NRMP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

Figure A.3: 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.

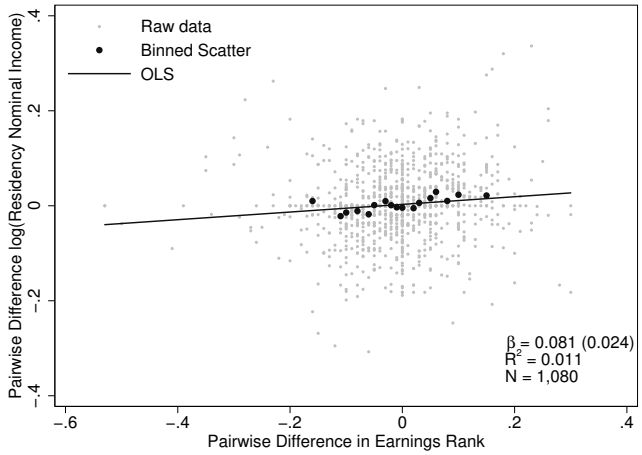
Figure A.4: Distribution Over Time of Survey Responses and NRMP Rank Submissions



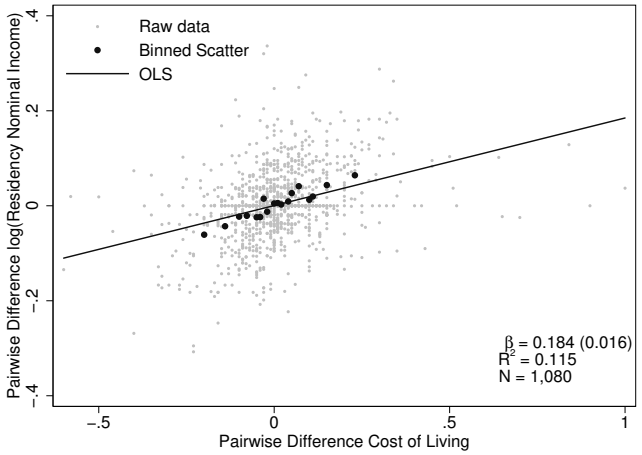
Notes: Distribution of timing of responses to Baseline and Follow-up Surveys, and NRMP rank submission dates (as reported by respondents in the follow-up survey).

Figure A.5: Variation in Nominal Income, Earnings Rank and Cost of Living

a. Earnings Rank vs. Nominal Income

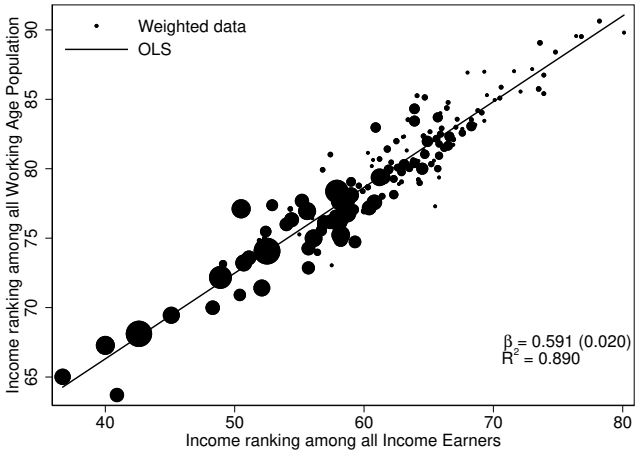


b. Cost of Living vs. Nominal Income



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).

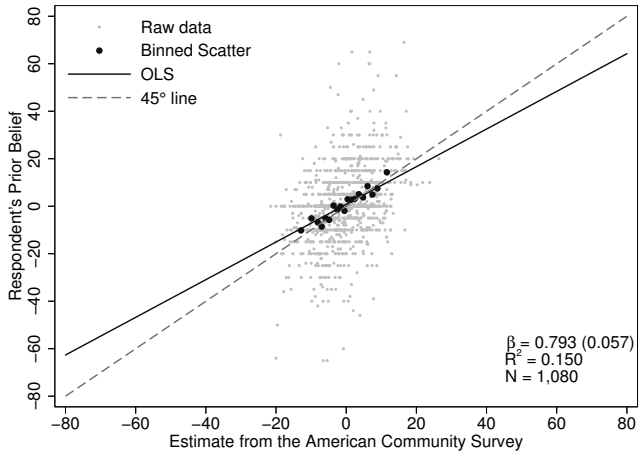
Figure A.6: Alternative Definition of Earnings Rank



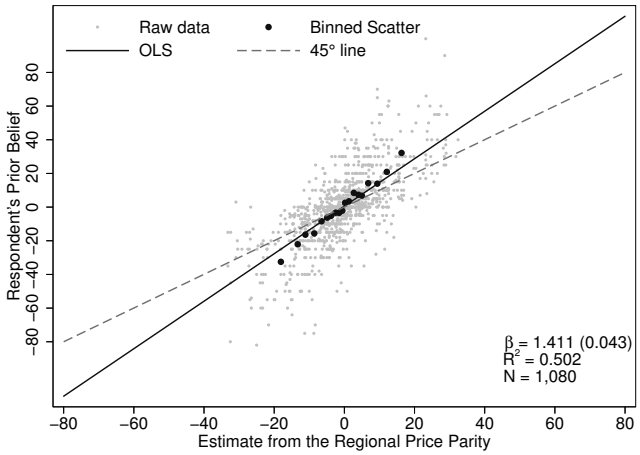
Notes: Dots are weighted responses for each metro area chosen by subjects. The x-axis shows our main measure for income rank (based on $54,000) for all income earners obtained from the American Community Survey (ACS) in levels. The y-axis corresponds to levels of income rank (based on $54,000) among all working age population (also from the ACS). Slopes (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

Figure A.7: Comparison Between Prior Beliefs and Statistics

a. Earnings Rank, Pairwise Differences

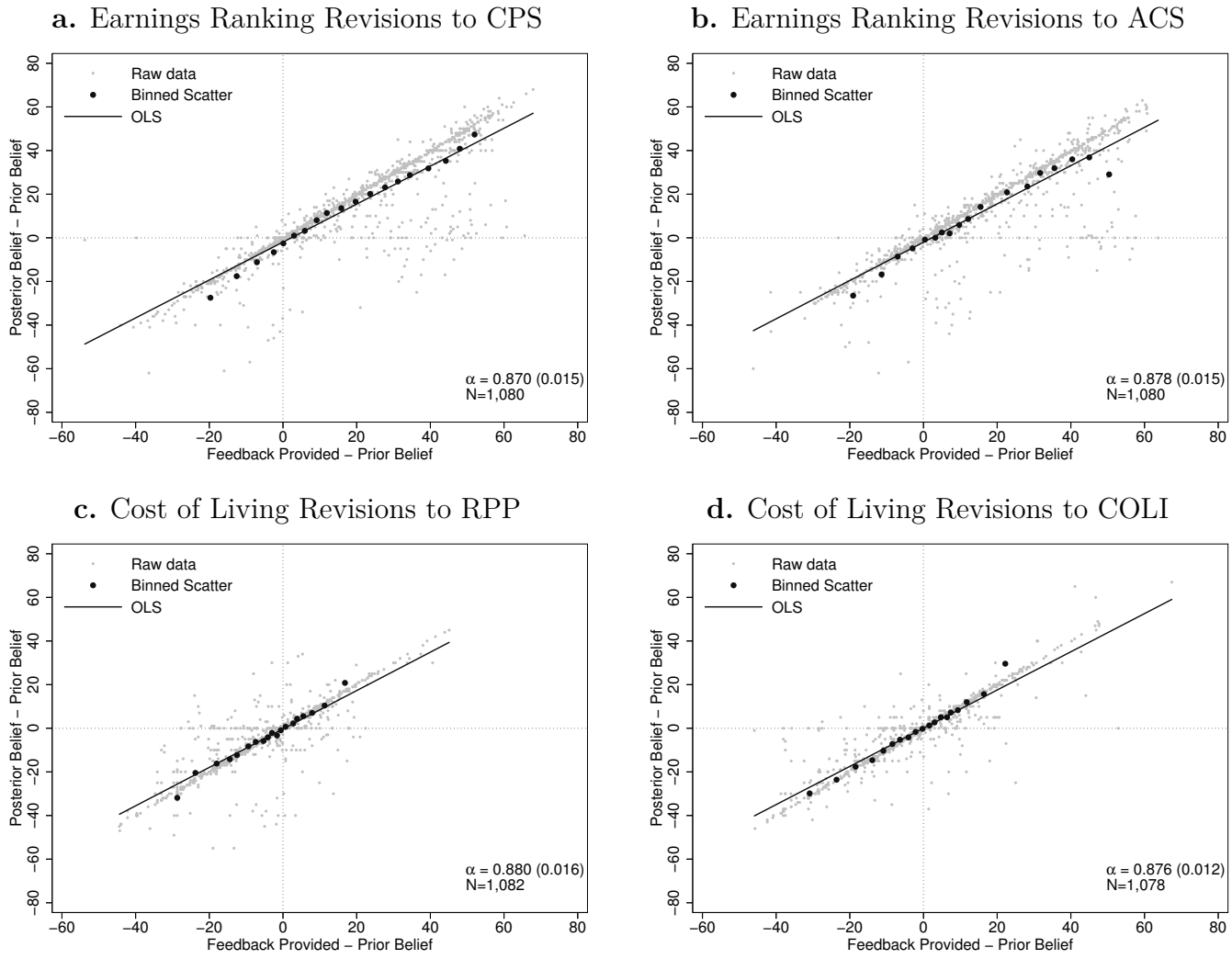


b. Cost of Living, Pairwise Differences



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.

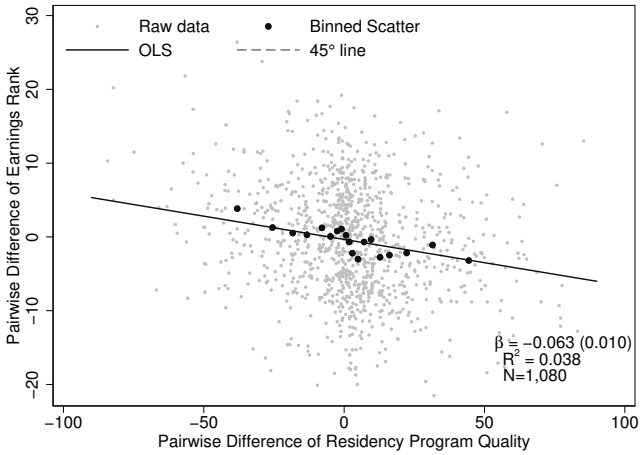
Figure A.8: Reduced-Form Evidence of Learning in the Information-Provision Experiment by Feedback Source



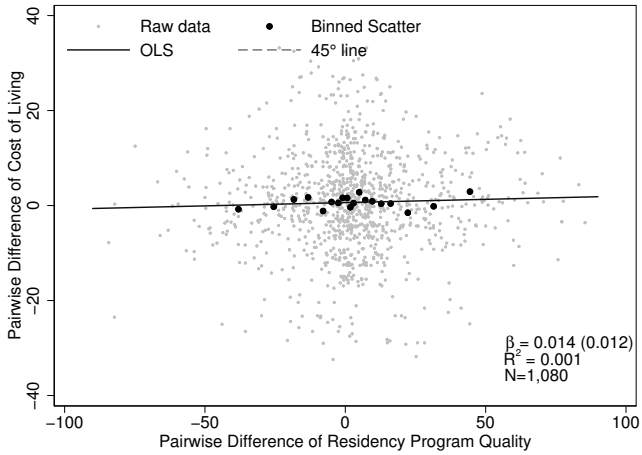
Notes: Comparison between the difference in statistics and respondent's perceptions before the information provision (i.e., prior beliefs), and difference in respondent's perceptions after the information provision (i.e., posterior beliefs) and prior beliefs. The gray dots correspond to the raw scatterplot,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

Figure A.9: Correlation between Earnings Rank, Cost of Living and Program Quality

a. Earnings Rank, Pairwise Differences



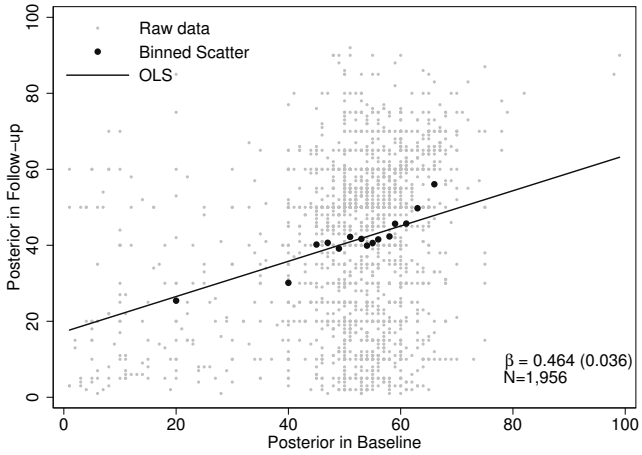
b. Cost of Living, Pairwise Differences



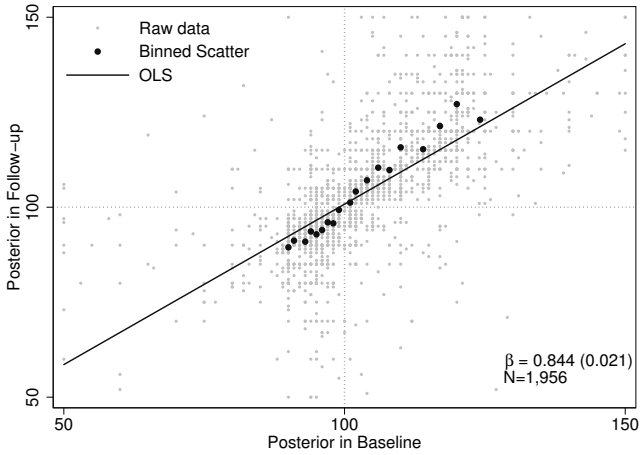
Notes: Pairwise differences of statistics for earnings rank or cost of living compared to residency program quality, based on cities, wages and programs indicated by respondents in the survey. The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. The sources in Panel (a) correspond to ACS (American Community Survey). The sources from Panel (b) correspond to RPP (Regional Price Parity Index). Residency program quality is the percentile rank of the residency program obtained from Doximity. The slope (β, with robust standard errors in parentheses) and R^2 are based on a linear regression.

Figure A.10: Correlation between (Posterior) Beliefs in Baseline and Follow-Up Surveys

a. Earnings Rank



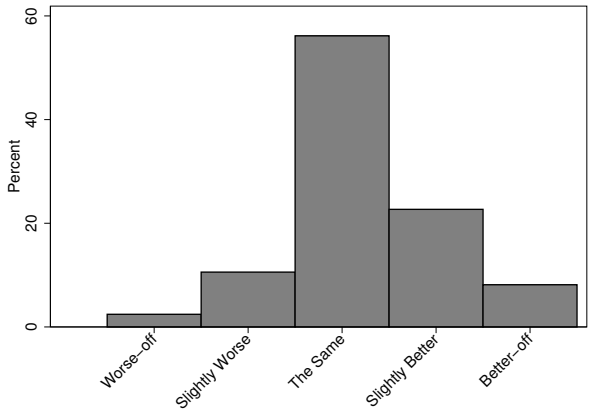
b. Cost of Living



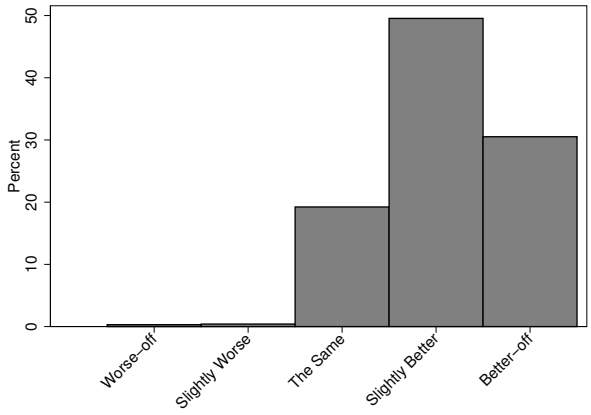
Notes: The gray dots correspond to the raw scatterplot, and the darker dots correspond to the binned-scatterplot based on 20 bins. Panels a and b present data in levels (i.e., two observations per individual, one for each of their options). The slope (β, with robust standard errors in parentheses) and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

Figure A.11: Survey Responses, Preferences over Hypothetical Changes to Earnings Rank and Cost of Living

a. Hypothetical Increase in Earnings Rank

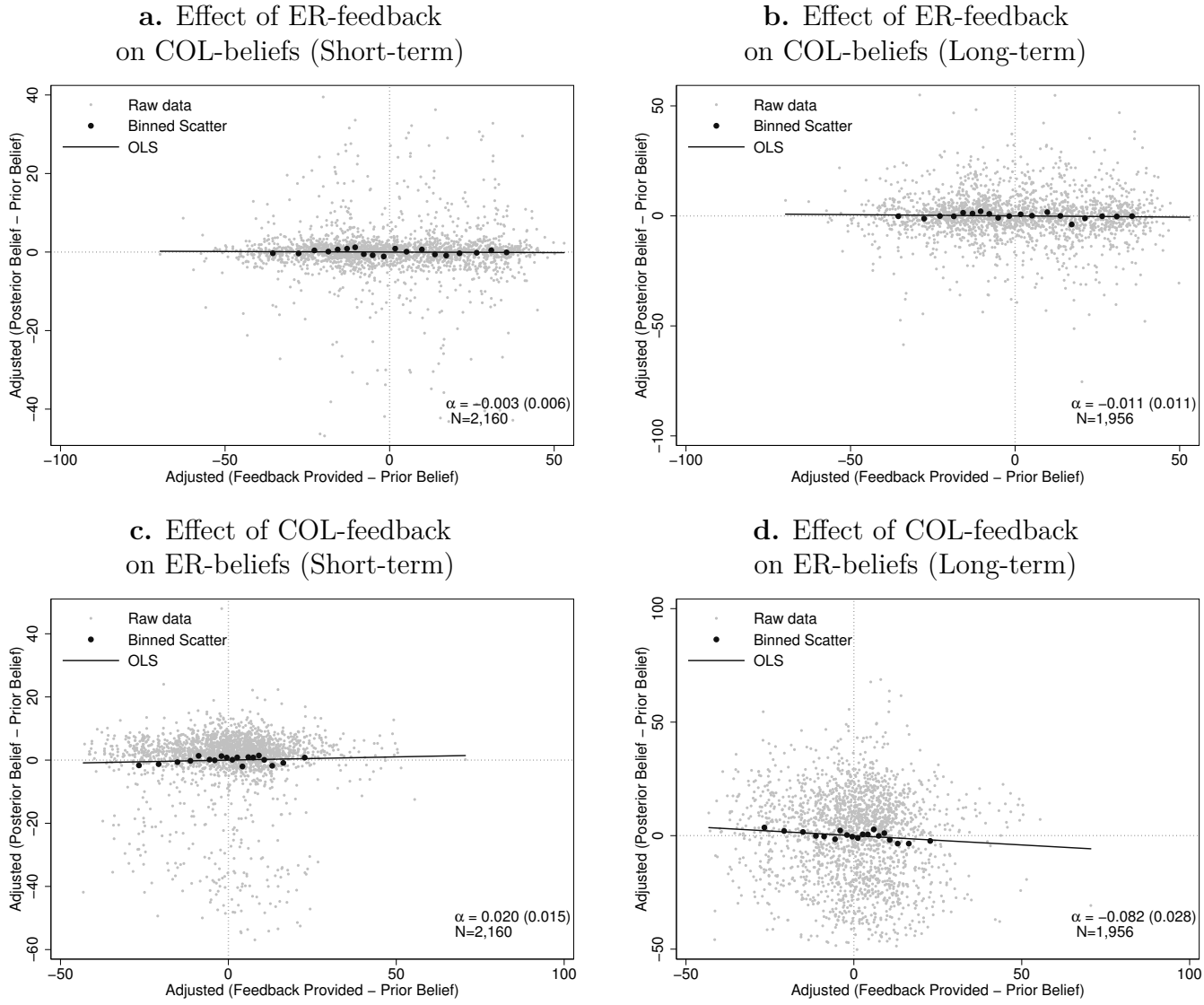


b. Hypothetical Decrease in Cost of Living



Notes: 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 C.2.

Figure A.12: Effect of Earning Rank Feedback on Posterior Belief on Cost of Living (and vice-versa)



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 shows the extent to which respondents adjust their perceptions on earnings rank (cost of living) as a result in their perception gap in cost of living (earnings rank) adjusted for the perceptions gap in ER (COL). The slope (α, with robust standard errors in parentheses) is based on a linear regression.

Table A.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: $COL^i_{1,2}$	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 A.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 A.3: 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 A.4: 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 $COL_{1,2}$	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 $COL_{1,2}$	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 A.5: Comparison of Characteristics between Experimental Subjects and the General U.S. Population

	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.

Table A.6: Location Preferences: Probit Marginal Effects

	Panel A: β^{ER}			Panel B: β^{COL}		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)
Baseline Sample						
Raw Probit	2.236***	-1.538*	0.995*	-1.087	-1.058	-1.073**
	(0.669)	(0.880)	(0.539)	(0.663)	(0.749)	(0.485)
Marginal Effect	0.418***	-0.267*	0.186*	-0.203	-0.183	-0.201**
	(0.125)	(0.155)	(0.100)	(0.124)	(0.130)	(0.090)
Follow-up Sample						
Raw Probit	2.380***	-1.656*	1.141**	-1.234*	-1.379*	-1.262**
	(0.702)	(0.991)	(0.577)	(0.743)	(0.772)	(0.531)
Marginal Effect	0.425***	-0.253*	0.202**	-0.221*	-0.211*	-0.224**
	(0.125)	(0.154)	(0.102)	(0.132)	(0.118)	(0.094)

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 earnings rank and cost of living 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 A.7: Location Preferences: Sensitivity to Sample Definition

	Panel A: β^{ER}			Panel B: β^{COL}		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)
Baseline Sample	2.236***	-1.538*	0.995*	-1.087	-1.058	-1.073**
	(0.669)	(0.880)	(0.539)	(0.663)	(0.749)	(0.485)
Pass Attention Check	2.248***	-1.380	1.077**	-0.928	-1.283	-1.087**
	(0.681)	(0.897)	(0.542)	(0.679)	(0.784)	(0.497)
Drop Dual Matches	2.158***	-1.308	1.005*	-1.080	-1.134	-1.119**
	(0.698)	(0.854)	(0.550)	(0.669)	(0.773)	(0.495)

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 earnings rank and cost of living 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 A.8: Location Preferences: Exclude Subsample

	All			Drop Repeat City		
		By Relationship Status			By Relationship Status	
	All	Non-Single	Single	All	Non-Single	Single
	(1)	(2)	(3)	(4)	(5)	(6)
β^{ER}	0.995*	2.236***	-1.538*	1.057*	2.377***	-1.726*
	(0.539)	(0.669)	(0.880)	(0.552)	(0.689)	(0.918)
β^{COL}	-1.073**	-1.087	-1.058	-1.039**	-1.056	-0.966
	(0.485)	(0.663)	(0.749)	(0.489)	(0.671)	(0.754)
Pseudo R^2	0.025	0.047	0.026	0.025	0.049	0.032
Observations	1,080	698	382	1,018	657	361

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 earnings rank and cost of living estimated by sample (i.e., coefficients in the same panel are are from a single regression based on the same sample). There are two panels: the left panel row shows estimates for the baseline sample (1,080 responses, 698 from non-singles and 382 from singles). The right panel excludes respondents who attemped to list a second program from the same city.

Table A.9: Location Preferences: Sensitivity to Largest and Smallest Metro Areas

	Panel A: β^{ER}			Panel B: β^{COL}			N
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)	(7)
Baseline	2.236***	-1.538*	0.995*	-1.087	-1.058	-1.073**	1,080
	(0.669)	(0.880)	(0.539)	(0.663)	(0.749)	(0.485)	
Exclude:							
3 Largest Metro Areas	2.981***	-1.007	1.828***	-0.232	-0.134	-0.268	814
	(0.796)	(1.073)	(0.639)	(0.815)	(0.813)	(0.591)	
Small Metro Areas	2.281***	-1.840**	0.904	-0.964	-0.787	-0.914*	1,023
	(0.685)	(0.926)	(0.551)	(0.689)	(0.806)	(0.505)	

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Raw Probit coefficients. The coefficients from the same row are estimated from the same regression. The first row shows estimates for baseline sample (1,080 responses, 698 from non-singles and 382 from singles). The second row excludes respondents who reported at least one residency program from one of the three largest metro areas: New York, Los Angeles and Chicago. The third row excludes respondents who reported at least one residency program from one the 25 smallest metro areas.

Table A.10: Robustness Check: Additional Controls on Hospital and Residency Programs

	Panel A: β^{ER}			Panel B: β^{COL}			Pseudo R^2		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)	Non-Single (7)	Single (8)	All (9)
Baseline	2.236***	-1.538*	0.995*	-1.087	-1.058	-1.073**	0.047	0.026	0.025
	(0.669)	(0.880)	(0.539)	(0.663)	(0.749)	(0.485)			
+Metro Nr Programs	2.334***	-1.612*	1.015*	-1.052	-0.894	-1.020**	0.049	0.034	0.026
	(0.677)	(0.906)	(0.544)	(0.659)	(0.766)	(0.486)			
+Metro Program Quality	2.272***	-1.600*	1.004*	-1.178*	-0.914	-1.130**	0.060	0.035	0.032
	(0.666)	(0.911)	(0.533)	(0.666)	(0.746)	(0.492)			
+Chosen Program Quality	2.321***	-1.678	1.214**	-1.512**	-1.511**	-1.481***	0.183	0.172	0.154
	(0.722)	(1.158)	(0.618)	(0.751)	(0.761)	(0.536)			
Baseline									
+Metro Nr Hospitals	2.343***	-1.594*	1.064**	-1.150*	-1.047	-1.103**	0.052	0.029	0.027
	(0.674)	(0.891)	(0.540)	(0.671)	(0.753)	(0.487)			
+Metro Hospital Quality	2.858***	-1.540*	1.235**	-1.141*	-1.044	-1.068**	0.084	0.084	0.038
	(0.703)	(0.922)	(0.552)	(0.649)	(0.800)	(0.485)			
+Chosen Program Quality	2.857***	-1.816	1.452**	-1.234*	-1.720**	-1.252**	0.192	0.226	0.151
	(0.772)	(1.210)	(0.644)	(0.731)	(0.832)	(0.531)			
All	3.144***	-1.810	1.565**	-1.311*	-1.288*	-1.268**	0.215	0.255	0.172
	(0.794)	(1.298)	(0.662)	(0.744)	(0.780)	(0.539)			

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 earnings rank and cost of living 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 corresponds to the baseline specification. Each row progressively add control variables. Row 2 controls for the relative number of programs within selected specialty. Row 3 adds controls for overall quality of residency programs in the metro areas: the relative average percentile ranking and the share of programs ranked in the top 10th percentile. Row 4 adds objective and subjective characteristics for the chosen programs as in Table 2. Row 5 presents baseline results including the relative number of *hospitals* in each metro area. Row 6 adds measures of hospital quality in the metro area (relative averages of mortality rate, readmission rate and overall patient satisfaction score; and shares of hospitals in the city that are in the top 10th percentile nationally for the measure). The final row includes all referenced controls.

Table A.11: Location Preferences: Subjective Program Characteristics

	(1)	(2)	(3)	(4)
β^{ER}	1.141**	1.147*	1.172*	1.172**
	(0.577)	(0.609)	(0.602)	(0.584)
β^{COL}	-1.262**	-1.211**	-1.470***	-1.412***
	(0.531)	(0.529)	(0.525)	(0.515)
$\beta^{purpose}$		0.441***		
		(0.064)		
$\beta^{prospects}$			0.379***	
			(0.070)	
$\beta^{prestige}$				0.249***
				(0.061)
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 earnings rank and cost of living 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 A.12: Location Preferences: Additional Results on Heterogeneity with Respect to Marital Status

	Non-Single		Non-Single		Single	
	Married	LT Relationship	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
β^{ER}	2.002*	2.345***	2.754***	1.733*	-2.472*	-1.023
	(1.187)	(0.841)	(0.964)	(0.966)	(1.282)	(1.318)
β^{COL}	-2.403**	-0.311	-1.172	-1.366	-0.630	-1.634
	(0.999)	(0.844)	(1.023)	(0.952)	(0.805)	(1.294)
Diff. P-value *[q-value]:*						
ER	0.813 *[0.883]*		0.430 *[0.746]*		0.454 *[0.784]*	
COL	0.109 *[0.640]*		0.509 *[0.919]*		0.890 *[0.746]*	
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 earnings rank and cost of living (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 A.13: Location Preferences: Heterogeneity with Respect to Other Individual Characteristics

	Hypothetical decrease cost of living		Hypothetical increase earnings rank		By Materialism		By Competitiveness		By Life Dimension	
	Better off	Same/Worse off	Better off	Same/Worse off	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
β^{ER}	1.713**	-0.683	1.812*	0.893	1.698**	0.828	1.229*	0.838	1.656*	0.667
	(0.669)	(1.186)	(1.023)	(0.734)	(0.708)	(0.952)	(0.664)	(1.197)	(0.908)	(0.779)
β^{COL}	-1.189**	-1.483	-1.721**	-1.117	-0.638	-2.283***	-1.659***	0.076	-1.028	-1.945**
	(0.585)	(1.159)	(0.799)	(0.710)	(0.757)	(0.751)	(0.608)	(0.958)	(0.748)	(0.869)
Diff. P-value *[q-value]:*										
ER	0.078 *[0.467]*		0.465 *[0.751]*		0.463 *[0.751]*		0.775 *[0.917]*		0.408 *[0.714]*	
COL	0.820 *[0.936]*		0.572 *[0.811]*		0.123 *[0.536]*		0.126 *[0.536]*		0.424 *[0.729]*	
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 earnings rank and cost of living (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 A.14: 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 A.15: Location Preferences: Happiness Trade-Offs

	Panel A: β^{ER}			Panel B: β^{COL}		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)
Baseline	1.506**	-0.004	0.957*	-0.793	0.414	-0.403
	(0.627)	(0.947)	(0.516)	(0.619)	(0.756)	(0.478)
Experimental	2.958***	-1.617	1.751*	0.028	1.319	0.471
	(1.082)	(2.022)	(0.965)	(1.045)	(1.235)	(0.791)
Experimental + Long Term	2.845***	-2.192	1.323	0.664	-1.176	-0.072
	(1.065)	(2.189)	(0.975)	(0.945)	(1.215)	(0.759)

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 earnings rank and cost of living (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).

Table A.16: Location Preferences: Alternative Regression Models

	Panel A: β^{ER}			Panel B: β^{COL}		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)
Binary Probit:						
Raw Coefficient	2.236***	-1.538*	0.995*	-1.087	-1.058	-1.073**
	(0.669)	(0.880)	(0.539)	(0.663)	(0.749)	(0.485)
Marginal Effect	0.418***	-0.267*	0.186*	-0.203	-0.183	-0.201**
	(0.125)	(0.155)	(0.100)	(0.124)	(0.130)	(0.090)
Binary Logit:						
Raw Coefficient	4.069***	-2.635	1.925*	-2.017	-1.930	-2.008**
	(1.178)	(1.631)	(0.986)	(1.232)	(1.378)	(0.882)
Marginal Effect	0.391***	-0.233	0.186**	-0.194*	-0.171	-0.194**
	(0.111)	(0.146)	(0.094)	(0.118)	(0.122)	(0.084)
Ordered Probit:						
Raw Coefficient	1.342***	-0.301	0.738**	-0.893**	-0.098	-0.599*
	(0.476)	(0.596)	(0.373)	(0.394)	(0.493)	(0.309)
Linear Probability Model:						
Raw Coefficient	0.400***	-0.228	0.202*	-0.182	-0.186	-0.225**
	(0.147)	(0.146)	(0.112)	(0.125)	(0.150)	(0.107)

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Each row corresponds to the same regression but based on different models: Binary Probit, Binary Logit, Ordered Probit and Linear Probability Model. In addition to the raw coefficients, the Binary Probit and Binary Logit specifications report the marginal effects estimated at the means of the independent variables. All regressions include the same set of baseline controls listed in section 3. All results are based on the same sample of individuals who completed the baseline survey (1,080 responses, 698 from non-singles and 382 from singles).

Table A.17: Location Preferences: IV, First Stage, and Reduced Form Estimates

	Non-Single	Single	All
	(1)	(2)	(3)
Panel A: IV-Probit Estimates			
β^{ER}	2.977**	-4.964**	0.867
	(1.331)	(1.974)	(1.151)
β^{COL}	0.353	1.663	0.662
	(1.160)	(1.286)	(0.881)
Panel B: First Stage			
Dep. Var.: $ER^i_{1,2}$			
$\Delta ER^i_{1,2}$	0.855***	0.687***	0.797***
	(0.055)	(0.082)	(0.045)
$\Delta COL^i_{1,2}$	0.019	0.007	0.012
	(0.049)	(0.064)	(0.039)
Dep. Var.: ${COL_{1,2}}^i$			
$\Delta ER^i_{1,2}$	-0.101***	0.035	-0.058
	(0.036)	(0.089)	(0.037)
$\Delta COL^i_{1,2}$	0.893***	0.985***	0.928***
	(0.064)	(0.070)	(0.048)
Wald test of exog. p-val.	0.324	0.004	0.061
Cragg-Donald F-stat.	172.03	42.85	207.29
Panel C: Reduced Form			
$DeltaER^i_{1,2}$	2.494**	-3.493**	0.661
	(1.151)	(1.602)	(0.920)
$\Delta COL^i_{1,2}$	0.501	1.777	0.722
	(1.071)	(1.382)	(0.847)
Observations	647	331	978

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 earnings rank and cost of living, 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.

B Auxiliary Experiment

B.1 Survey Design

In this section, we present complementary evidence from an auxiliary experiment. We designed a variation of the survey instrument, attached in Appendix C.7, that can be used in other contexts besides the medical residency match. At the beginning of this survey, we ask respondents to list two cities that they know well to which they would consider moving to. The following questions are identical to the baseline survey instrument from our main experiment with medical students: we elicit prior and posterior beliefs about earnings rank and cost of living, we conduct the information-provision experiment, and we elicit preferences for the two cities under consideration.

We conducted this auxiliary experiment using a sample of 1,245 U.S. respondents recruited through Amazon Mechanical Turk.[14] Compared to the main residency match experiment, the primary difference is that the subjects in the auxiliary experiment are not moving anytime soon, so we cannot followup with them to measure the effects of the information provision experiment on their actual location choices. Instead, we measure the effects on their expected location choices. This is a limitation of the auxiliary experiment.[15] On the other hand, this auxiliary sample has some advantages. Due to a wide availability of subjects, it is possible to run additional experiments on demand to address alternative mechanisms and to disentangle the mechanisms at play. Also, this secondary sample is more diverse than the main sample of medical students in many observable characteristics.[16] As a result, this secondary sample can help to assess the external validity of the results – for example, due to their high incomes or the competitive nature of their profession, it is possible that doctors have stronger relative concerns compared to the general population.

Also, we extended the survey instrument used in this auxiliary sample to test a specific explanation for the preference for relative income: individuals may use information about relative income to learn about other city characteristics such as school quality and crime rates. Towards the end of the auxiliary survey, after individuals received feedback about relative income and the cost of living, we

[14] Details about the recruitment are presented in Appendix B.2.

[15] The results from the previous section suggests that using hypothetical choices may not be as problematic as generally thought, to the extent that preferences inferred from expected choices are consistent with preferences inferred from actual choices.

[16] For example, on average, participants in the auxiliary sample are older and less educated. Descriptive statistics comparing sample characteristics between the main and auxiliary experiment are presented in Appendix B.3.

included a set of additional questions eliciting beliefs about other attributes of the two cities under consideration. We picked eight attributes that individuals could arguably perceive to be correlated to the affluence of a city: quality of schools, crime rates, quality of health services, quality of public spaces, quality of the environment, quality of entertainment, share of college graduates, and share of supporters of Donald Trump. If individuals learn about these attributes from the earnings rank, then controlling for those attributes should mute the association between perceived earnings rank and choice.

B.2 Recruitment of Auxiliary Experiment

We conducted an auxiliary experiment using a sample of respondents from Amazon Mechanical Turk (mTurk), an online job market for crowdsourcing small tasks. During September of 2017, we recruited the auxiliary sample through work postings (or HIT - "Human Intelligence Task") on mTurk. Participants were invited to participate in a 8 minute survey "about city perceptions". When accepting the task, participants were re-directed to the survey. After successful completion of the survey, participants were given a code to redeem their payment of $0.60 for completing the task. We restricted the survey to participants located in the United States.

B.3 Comparing Subjects in Main and Auxiliary Experiments

In Table B.1 we compare respondent characteristics between subjects in the main experiment and the auxiliary experiment. Notably, there are various significant differences between these samples. On average, the medical student sample is 10 years younger than the auxiliary experiment, and more likely to be male. Additionally, respondents in the main experiment are half as likely to be married than respondents in the auxiliary experiment (24% compared to 46%). Additionally, over half of respondents in the auxiliary experiment have at least one child, while only 8.9% of respondents in the main experiment have at least one child.

B.4 Experimental Variation, Learning and Perceptions in Auxiliary Experiment

After replicating the information experiment and obtaining posterior beliefs about the earnings ranking and cost of living, subjects were asked about perceptions regarding various other city characteristics. The order of these other perceptions were randomized for each respondent. In Table B.2, we

present the results of the effect of the experimental variation on each perception, presenting the corresponding coefficient, p-value and multiple-testing q-value based on Benjamini and Yekutieli (2001). The main results of the experiment replicate and we do not find significant evidence of cross-learning between earnings rank and cost of living. Overall, respondents do not seem to make significant inferences about other city characteristics. However, there does seem to be some learning about other city characteristics, though in many cases in the opposite direction than expected. For example, increasing a respondent's earnings ranking (meaning, locating in a less affluent pond) is associated to believing that the quality of schools or public places or the percentage of college graduates would be higher – when the opposite would be expected.

B.5 Results: Replication of the Main Experiment

Table B.3 replicates the preference estimation from Table 3, but using data from the auxiliary experiment instead of the main experiment. The comparison between Tables B.3 and 3 can shed light on the robustness and external validity of the results from the main experiment.

The first row of Table B.3 corresponds to the baseline estimates, which uses the experimental and non-experimental variations in beliefs. The coefficients β^{ER} and β^{COL} are similar between the main experiment and the auxiliary experiment. Focusing on the entire sample, the estimated β^{ER} and β^{COL} are 1.141 and -1.262 in the main experiment (p-value=0.048 and p-value=0.017), and 1.293 and -1.962 in the auxiliary experiment (p-value<0.001 for both). That is, the coefficients have the same signs and similar magnitudes. The coefficients are more precisely estimated in the auxiliary sample, in part due to the larger sample size.

To compare the magnitude of relative income concerns, we can compare the marginal rate of substitution between relative income and cost of living: i.e., $\frac{\beta^{ER}}{-\beta^{COL}}$. In the main experiment, $\frac{\beta^{ER}}{-\beta^{COL}}$ is 0.90 (s.e. 0.64; from the first row, columns (3) and (6) of Table 3). In the auxiliary experiment, we find a corresponding ratio of 0.66 (s.e. 0.20; from the first row, columns (3) and (6) of Table B.3). That is, the auxiliary experiment suggests slightly weaker preferences for relative income than the main experiment, but that difference is statistically insignificant. This constitutes suggestive evidence that, despite large observable differences in observable characteristics, medical students are not special in terms of their preferences for relative income.

The second row presents the experimental estimates. The results from the auxiliary experiment

are even more robust than the results from the main experiment. In the main experiment, the baseline estimates for β^{ER} are similar to the experimental estimates, and this is true again in the auxiliary experiment. In the main experiment, the experimental estimates for β^{COL} are statistically insignificant and smaller in magnitude than the baseline estimates. In the auxiliary sample, the experimental estimates for β^{COL} are negative, precisely estimated, and statistically significant at the 1% level.

Another important finding from the main experiment is the heterogeneity in β^{ER} by relationship status. In the auxiliary experiment, we find evidence consistent with this heterogeneity, although it is less extreme. Table B.3 shows that, when we break down β^{ER} by relationship status, the coefficient of β^{ER} is smaller among singles than among non-singles. This difference, however, becomes stronger in the experimental estimates: β^{ER} is 2.578 for singles and 0.664 for non-singles. However, there are two notable differences: the difference is statistically insignificant in the auxiliary experiment (p-value=0.244), and less pronounced in magnitude than the corresponding heterogeneity in the main experiment. One potential interpretation for this finding is that the results from the main sample exaggerate the degree of heterogeneity by relationship status. Another potential interpretation is that the difference in results are due to differences in the characteristics across the two samples. For example, compared to singles in the auxiliary sample, singles in the medical student sample are much more likely to be in prime dating age, less likely to have children from previous relationships, and they expect a much higher permanent income.

B.6 Results: Disentangling Mechanisms

One possible interpretation of the coefficient β^{ER} is that individuals use their prospective relative income as a signal for other city attributes. For instance, individuals may use information about relative income to learn about other city characteristics, such as school quality and crime rates. Although these inferences would not be unreasonable, this mechanism would probably work against our main finding: if more affluent ponds tend to have desirable amenities, then individuals should prefer to live in more affluent ponds, which is the opposite of what we find.

We can test this hypothesis using the perceptions about additional characteristics elicited in the auxiliary experiment. One important difference in context is that, whereas subjects in the main experiment made a high-stakes decision for which they obtained substantial information, subjects in

the auxiliary experiment had no immediate incentives to be informed about the attributes of these cities and then may make more sense for them to use information about earnings rank to make inferences about other unobserved city attributes. In other words, this mechanism would arguably play a bigger role in the auxiliary experiment than in the main experiment.

The third row of Table B.3 estimates the experimental model, but adding these eight perceptions as additional control variables. If individuals care about relative income because they learn about the other characteristics, the coefficient on β^{ER} should be muted after controlling for these additional perceptions. On the contrary, the comparison between the second and third rows of Table B.3 suggests that controlling for the additional characteristics, if anything, increases the value of β^{ER}. For instance, among non-singles, the experimental estimate for β^{ER} is 2.578 (p-value=0.011) without these additional controls and 3.048 (p-value=0.004) with these additional controls. The difference between these two coefficients is statistically insignificant. Also, the increase in β^{ER} caused by adding the extra controls is consistent with the previous argument that, if anything, this mechanism leads to an underestimation of preferences for relative income.

References

Armantier, O., S. Nelson, G. Topa, W. van der Klaauw, and B. Zafar (2016). The price is right: Updating inflation expectations in a randomized price information experiment. *Review of Economics and Statistics 98*(3), 503–523.

Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics 29*(4), 1165–1188.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics 9*(3), 1–35.

Clark, A., C. Senik, and K. Yamada (2017). When experienced and decision utility concur: The case of income comparisons. *Journal of Behavioral and Experimental Economics 70*, 1 – 9.

Clark, A. E., N. Westergård-Nielsen, and N. Kristensen (2009). Economic satisfaction and income rank in small neighbourhoods. *Journal of the European Economic Association 7*(2-3), 519–527.

Cullen, Z. and R. Perez-Truglia (2018). How much does your boss make? the effects of salary comparisons. *NBER Working Paper No. 24841*.

Ferrer-i Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics 89*(5), 997–1019.

Johansson-Stenman, O., F. Carlsson, and D. Daruvala (2002). Measuring future grandparents' preferences for equality and relative standing. *The Economic Journal 112*(479), 362–383.

Luttmer, E. F. (2005). Neighbors as negatives: relative earnings and well-being. *The Quarterly Journal of Economics 120*(3), 963–1002.

Senik, C. (2004). When information dominates comparison: Learning from russian subjective panel data. *Journal of Public Economics 88*(9), 2099–2123.

Wiswall, M. and B. Zafar (2014). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

Yamada, K. and M. Sato (2016). *Another avenue for anatomy of income comparisons: Evidence from hypothetical choice experiments*. Behavioral Economics of Preferences, Choices, and Happiness; Springer.

Table B.1: Comparison of Characteristics between Subjects in Main and Auxiliary Samples

	Main Experiment *Med. Students*	Auxiliary Experiment *Online Sample*	Difference
Age	27.091	37.476	-10.385***
	(2.725)	(11.980)	(0.350)
% Male	0.481	0.391	0.090***
	(0.500)	(0.488)	(0.021)
% Married	0.240	0.461	-0.221***
	(0.427)	(0.499)	(0.019)
% Has children	0.089	0.527	-0.438***
	(0.285)	(0.499)	(0.017)
Observations	1,080	1,245	

Notes: Standard deviations and Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. Sample statistics for Main Experiment (Medical Student sample) and Auxiliary Experiment (sample of online respondents on Amazon Mechanical Turk).

Table B.2: Experimental Variation, Learning, and Perceptions in the Auxiliary Sample

			Quality of					Amount of	Percentage of	
Dep. Var.:	$ER_{1,2}^{post}$	$COL_{1,2}^{post}$	Schools	Health	Public Spaces	Environment	Entertainment	Crime	College Graduates	Vote Trump
$\Delta ER_{1,2}$	0.741	-0.095	0.386	-0.066	0.341	0.041	0.077	-0.164	0.574	-0.070
P-value	0.000	0.048	0.000	0.539	0.000	0.689	0.477	0.112	0.000	0.585
Q-value	0.000	0.117	0.001	0.674	0.001	0.783	0.636	0.231	0.000	0.716
$\Delta COL_{1,2}$	0.039	1.017	0.059	-0.136	0.050	0.238	0.092	-0.384	0.001	0.235
P-value	0.513	0.000	0.516	0.139	0.537	0.008	0.349	0.000	0.995	0.033
Q-value	0.674	0.000	0.674	0.273	0.674	0.026	0.504	0.000	0.995	0.091

Notes: Heteroskedasticity-robust standard errors in parenthesis. All regressions include the baseline controls listed in section 3 with the exception of program characteristics. Multiple-testing q-values based on Benjamini and Yekutieli (2001) presented.

Table B.3: Location Preferences: Auxiliary Experiment

	Panel A: β^{ER}			Panel B: β^{COL}		
	Non-Single (1)	Single (2)	All (3)	Non-Single (4)	Single (5)	All (6)
Baseline	1.408***	1.095**	1.293***	-2.203***	-1.618***	-1.962***
	(0.376)	(0.478)	(0.292)	(0.463)	(0.566)	(0.364)
Experimental	2.578**	0.664	1.706**	-2.385***	-2.956***	-2.528***
	(1.019)	(1.272)	(0.816)	(0.666)	(0.917)	(0.531)
Experimental, Additional Controls	3.048***	0.452	1.902**	-2.329***	-3.753***	-2.688***
	(1.064)	(1.430)	(0.872)	(0.691)	(0.906)	(0.563)

Notes: Heteroskedasticity-robust standard errors in parenthesis. All regressions include the baseline controls listed in section 3 with the exception of program characteristics. The independent variables are the posterior beliefs about earnings rank and cost of living, 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 third row includes additional controls for differences in city perceptions: quality of schools, crime rates, quality of health, quality of public spaces, quality of the environment, quality of entertainment, quality of colleges, fraction voting Trump in presidential election. All results based on the sample of respondents in the United States on Amazon Mechanical Turk (1,245 responses, 829 from non-singles and 416 from singles).

C Snapshots of Invitations and Surveys

Here we include snapshots with a sample of the baseline survey (C.1) and follow-up survey (C.2). Additionally, this Appendix also includes a snapshot of an invitation sent out to the deans (C.3), the invitation sent out to the medical students to participate in the baseline survey (C.4), the invitation sent out to students inviting them to the follow-up survey (C.5), and a snapshot of the project's website (C.6). Finally, we present snapshots of the auxiliary experiment (C.7).

C.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

Medical School

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: UCLA Medical Center

Specialty:

Internal Medicine (IM)

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

54000

>>

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)

54000

>>

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?

20% ⇕

>>

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?

10% ⇕

>>

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 **$ 54000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

Richer than 35% of individual earners

>>

Imagine that you chose to work in **Champaign-Urbana, IL**. With your individual annual earnings of **$ 54000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s individual earners?

Richer than 47% of individual earners

>>

8

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 **$ 54000**, you would be richer than **57.9%** of **Los Angeles-Long Beach-Anaheim, CA**'s population.

With your individual annual earnings of **$ 54000**, you would be richer than **60.3%** of **Champaign-Urbana, IL**'s population.

Source: based on most recent data from the American Community 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?

6%

>>

12

Imagine that you chose to work in **Los Angeles-Long Beach-Anaheim, CA**. With your individual annual earnings of **$ 54000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

Richer than 58% of individual earners

>>

Imagine that you chose to work in **Champaign-Urbana, IL**. With your individual annual earnings of **$ 54000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s individual earners?

Richer than 60% 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 19.7% 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 3.3% lower 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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?

Age

What is your relationship status?

- ○ Single
- ○ In a long-term relationship
- ○ Married

How many children do you have?

Recent research on decision making shows that choices are affected by the context in

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:

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

>>

C.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 20

>>

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 20, which of these two programs did you **rank higher**?

- ○ 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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?

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?

[ ]

>>

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 **$** 54000, you would be richer than what percentage of Los Angeles-Long Beach-Anaheim, CA's individual earners?

Richer than 58% of individual earners

>>

Imagine that you chose to work in Champaign-Urbana, IL. With your individual annual earnings of $ 54000, you would be richer than what percentage of Champaign-Urbana, IL's individual earners?

Richer than 60% 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center

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

- ○ Very likely UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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 UCLA Medical Center (Los Angeles-Long Beach-Anaheim, CA)
- ○ Likely UCLA Medical Center
- ○ Leaning 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 20, how many programs did you rank in total?

>>

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	○	○	○	○	○
I am a competitive individual	○	○	○	○	○
I enjoy competing against an opponent	○	○	○	○	○
I don't like competing against other people	○	○	○	○	○
I get satisfaction from competing with others	○	○	○	○	○
I find competitive situations unpleasant	○	○	○	○	○

>>

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	○	○	○	○	○
The things I own say a lot about how I'm doing in life	○	○	○	○	○
Buying things gives me a lot of pleasure	○	○	○	○	○
I like a lot of luxury in my life	○	○	○	○	○
My life would be better if I owned certain things I don't have	○	○	○	○	○
I'd be happier if I could afford to buy more things	○	○	○	○	○

C.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,

C.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]

C.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]

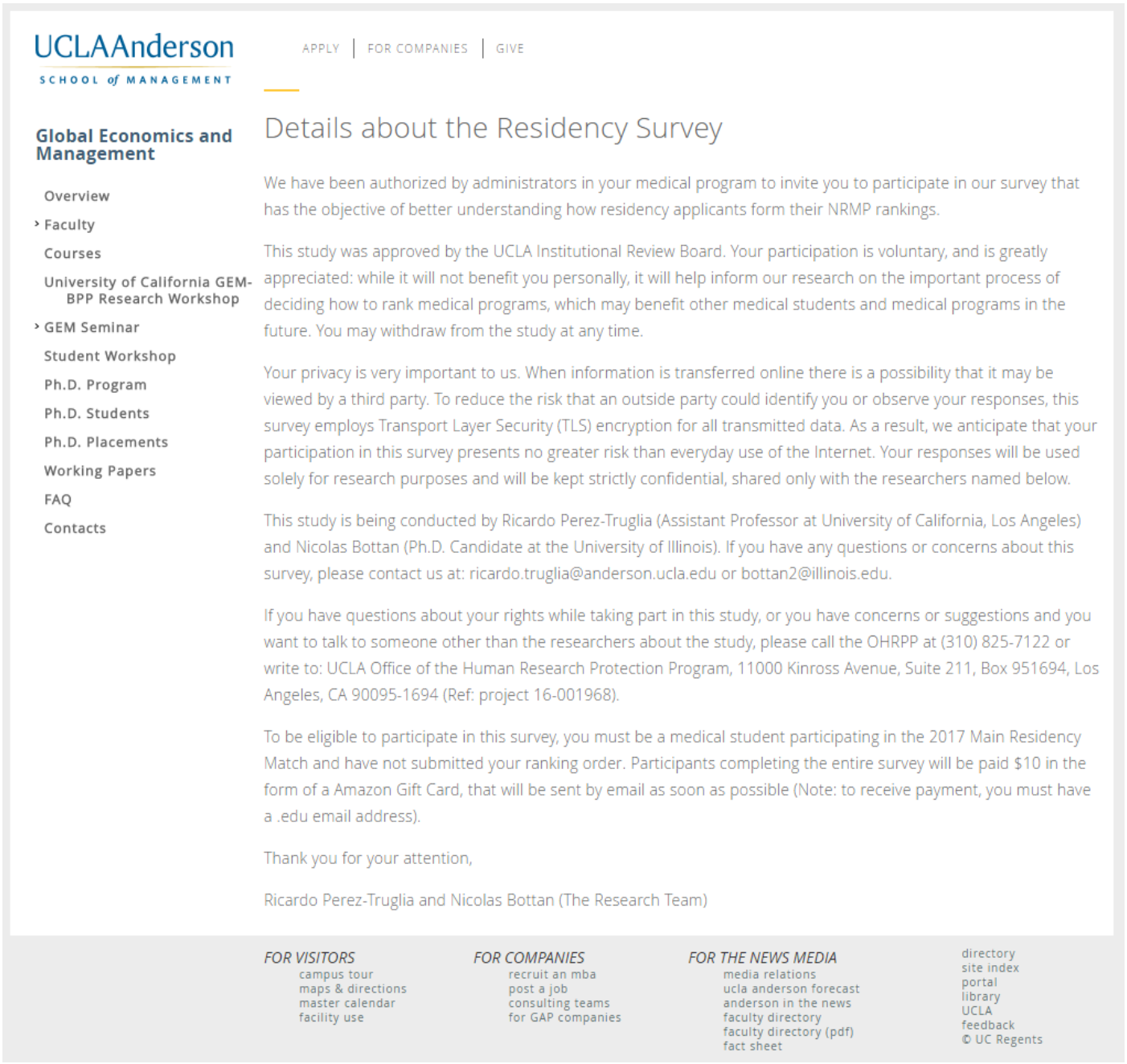
Or copy and paste the URL below into your internet browser: [URL]

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]

C.6 Project's Website



UCLAAnderson
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

C.7 Sample Questionnaire: Auxiliary Experiment

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?

Age

What is your relationship status?

- ○ Single
- ○ In a long-term relationship
- ○ Married

How many children do you have?

1

How many children do you have?

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

>>

Consider the following hypothetical scenario:

Think about two cities you know well but that you do not currently live in.

Now imagine that you are offered a job where you will be paid **an annual gross salary of $54,000** and have to move to one of those two cities.

Please tell us (in no particular order) the first location you would consider moving to:

State: Illinois

Metropolitan Area: Champaign-Urbana, IL

How well do you know this place? [0 - not at all; 10 - Extremely well]

0 1 2 3 4 5 6 7 8 9 10

>>

Please tell us the other location you would consider moving to:

State: California

Metropolitan Area: Los Angeles-Long Beach-Anaheim, CA

How well do you know this place? [0 - not at all; 10 - Extremely well]

0 1 2 3 4 5 6 7 8 9 10

>>

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 **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?

[▲▼]

[>>]

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?

>>

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 **Champaign-Urbana, IL**. With your individual annual earnings of **$ 54,000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s individual earners?

Richer than 70% of individual earners

>>

Imagine that you chose to work in **Los Angeles-Long Beach-Anaheim, CA**. With your individual annual earnings of **$ 54,000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

Richer than 45% 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 **Champaign-Urbana, IL** metro area is **6.6% cheaper** than the U.S. average.

The **Los Angeles-Long Beach-Anaheim, CA** metro area is **17.0% more expensive** 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 **$ 54,000**, you would be richer than **60.3%** of **Champaign-Urbana, IL**'s population.

With your individual annual earnings of **$ 54,000**, you would be richer than **57.9%** of **Los Angeles-Long Beach-Anaheim, CA**'s population.

Source: based on most recent data from the American Community 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 **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?

6% ▲▼

>>

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?

Imagine that you chose to work in **Los Angeles-Long Beach-Anaheim, CA**. With your individual annual earnings of **$ 54,000**, you would be richer than what percentage of **Los Angeles-Long Beach-Anaheim, CA**'s individual earners?

>>

Imagine that you chose to work in **Champaign-Urbana, IL**. With your individual annual earnings of **$ 54,000**, you would be richer than what percentage of **Champaign-Urbana, IL**'s 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 Champaign-Urbana, IL, you would be able to afford 19.8% more than if you chose to live in Los Angeles-Long Beach-Anaheim, CA.

- If you chose to live in Champaign-Urbana, IL, your earnings ranking would be 9.3% higher than if you chose to live in Los Angeles-Long Beach-Anaheim, CA.

If given the choice, where would you choose to live?

- ○ Very likely Champaign-Urbana, IL
- ○ Likely Champaign-Urbana, IL
- ○ Leaning Champaign-Urbana, IL
- ○ Leaning Los Angeles-Long Beach-Anaheim, CA
- ○ Likely Los Angeles-Long Beach-Anaheim, CA
- ○ Very likely Los Angeles-Long Beach-Anaheim, CA

>>

Next, we want to ask you about your perceptions regarding other aspects of these two cities .

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that **healthcare** (such as quantity and quality of hospitals and doctors) is better or worse than the US average?

- ○ better
- ● worse

How much worse is healthcare than the U.S. average?

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that **healthcare** (such as quantity and quality of hospitals and doctors) is better or worse than the US average?

- ● better
- ○ worse

How much better is healthcare than the U.S. average?

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that **schools** would be better or worse than the US average?

- (●) better
- ( ) worse

How much better are schools than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that **schools** would be better or worse than the US average?

- ( ) better
- (●) worse

How much worse are schools than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that the **quality of the environment** (such as the air and water purity) is better or worse than the US average?

- (●) better
- ( ) worse

How much better is the quality of the environment than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that the **quality of the environment** (such as the air and water purity) is better or worse than the US average?

- (●) better
- ( ) worse

How much better is the quality of the environment than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that **public spaces** (such as the number and quality of parks and recreational areas) are better or worse than the US average?

- ○ better
- ● worse

How much worse are public spaces than the U.S. average?

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that **public spaces** (such as the number and quality of parks and recreational areas) are better or worse than the US average?

- ● better
- ○ worse

How much better are public spaces than the U.S. average?

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that the **share of population with a college degree** is higher or lower than the US average?

- (●) higher
- ( ) lower

How much higher than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that the **share of population with a college degree** is higher or lower than the US average?

- ( ) higher
- (●) lower

How much lower than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that the **quality of entertainment** (such as the number and quality of cinemas, theaters and bars) is better or worse than the US average?

- ○ better
- ● worse

How much worse is the quality of entertainment than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that the **quality of entertainment** (such as the number and quality of cinemas, theaters and bars) is better or worse than the US average?

- ● better
- ○ worse

How much better is the quality of entertainment than the U.S. average?

[ ]

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that the **share of population who voted for Donald Trump in the 2016 general election** is higher or lower than the US average?

- ○ higher
- ● lower

How much lower than the U.S. average?

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that the **share of population who voted for Donald Trump in the 2016 general election** is higher or lower than the US average?

- ○ higher
- ● lower

How much lower than the U.S. average?

>>

Imagine that you chose to work in the **Champaign-Urbana, IL** metro area. Do you think that **crime** is higher or lower than the US average?

- ○ higher
- ● lower

How much lower is crime than the U.S. average?

>>

Imagine that you chose to work in the **Los Angeles-Long Beach-Anaheim, CA** metro area. Do you think that **crime** higher or lower than the US average?

- ● higher
- ○ lower

How much higher is crime than the U.S. average?

>>