Review of Economic Studies (2021) **88**, 378–405 doi:10.1093/restud/rdaa016

Advance access publication 19 May 2020

Heads or Tails: The Impact of a Coin Toss on Major Life Decisions and Subsequent Happiness

STEVEN D. LEVITT
University of Chicago and NBER

First version received November 2017*; Editorial decision October* 2019*; Accepted April* 2020 (*Eds.*)

Little is known about whether people make good choices when facing important decisions. This article reports on a large-scale randomized field experiment in which research subjects having difficulty making a decision flipped a coin to help determine their choice. For important decisions (*e.g.* quitting a job or ending a relationship), individuals who are told by the coin toss to make a change are more likely to make a change, more satisfied with their decisions, and happier six months later than those whose coin toss instructed maintaining the status quo. This finding suggests that people may be excessively cautious when facing life-changing choices.

Key words: Quitting, Happiness, Decision biases

JEL Codes: D12, D8

1. INTRODUCTION

In every life, there arise difficult decisions with potentially far-reaching consequences on lifetime utility: whether to quit a job, seek more education, end a relationship, quit smoking, start a diet, etc. Expected utility maximization is the workhorse economic model for thinking about such choices. Behavioural economics offers a host of alternative descriptive models of decision-making, *e.g.* prospect theory, hyperbolic discounting, and the sunk cost fallacy. Yet, from an empirical perspective, economics has almost nothing to say about whether or not people are actually making good choices when it comes to their most important decisions.[1]

1. There is, of course, a rich experimental literature exploring individual decision making under uncertainty. For surveys of this enormous literature, see Camerer (1995), Smith (1994), and Chaudhuri (2011). A notable recent contribution to decision-making under uncertainty is Gneezy *et al.* (2015). Most of this literature focuses on low-stakes decisions. Slonim and Roth (1998) and Andersen et al. (2011) explore decision-making in a high-stakes dictator game. In recent years, field experiments exploring decision making in natural environments have become more common (Bowles *et al.*, 2001; Gneezy and List, 2006; Dellavigna, 2009; Levitt and List, 2009), but most of these have investigated relatively

The editor in charge of this paper was Nicola Gennaioli.

One reason that so little is known about these important decisions is that researchers do not generally have the power to randomize people into treatments that compel them to, say, quit their jobs or leave their spouses. Even if it were possible to choose 1,000 married couples from the general population and randomly force 500 of those couples to divorce, it would not be particularly informative. Such a study would tell us about the average treatment effect of divorce. What we really care about, however, is the impact on the *marginal* decision maker. It would not be surprising if getting a divorce would have a devastating impact on the infra-marginal married person. A much more interesting question is whether divorce, *ex post*, will be the right choice for someone teetering on the edge of ending a relationship.[2]

Even if one found such a group of individuals who are close to indifferent between remaining married and getting divorced, an *ex post* comparison of the happiness of those who do and do not make a change still would not have an easy causal interpretation, because the people who make a change will systematically differ from those who do not on many dimensions. To convincingly answer the question, a researcher would not only need to find large numbers of these marginal individuals, but also, through some sort of randomization, influence their important life choices.

That is what I do in this study. I created a website called FreakonomicsExperiments.com. On the website, individuals who are having a difficult time making a life decision are asked to answer a series of questions concerning the decision they are struggling with. Users are presented with a wide range of questions to choose from (see Supplementary Appendix A for the full set of questions offered) or invited to create their own question. One choice (*e.g.* "go on a diet") is assigned to heads and the other choice (in this case "don't go on a diet") is assigned to tails. The outcome of the coin toss is randomized and the user is shown the outcome of the coin toss. The coin flippers are then re-surveyed two and six months after the initial coin toss. Additionally, prior to the randomization, coin flippers are encouraged to identify a third party (a friend or family member) to verify their outcomes. The third parties are also surveyed two and six months after the coin toss.

While it might seem implausible that anyone would come to such a website and flip a coin, much less follow the dictate of the coin toss, the results obtained speak to the contrary. In the year of data collection, over 20,000 coins were flipped. A number of results emerge from the analysis.

First, two months into the study participants show a bias towards the status quo, in the sense that people report making a change less frequently than they predicted they would before the coin toss. Six months after the coin toss, however, this bias is gone.

Second, those who report making a change in follow-up surveys are substantially happier than those who do not make a change, and they are more likely to say they would make the same decision if they were to choose again. This is true for virtually every question asked both two and six months later. This correlation does not, of course, necessarily imply causality. Those who make a change differ from those who do not make a change on many dimensions.

Third, the outcome of the coin toss appears to influence the actions taken. Those who flipped heads were approximately 25% more likely to report making a change than those who got tails. The coin toss had a roughly equal impact on decisions across the entire range of self-stated *ex ante* likelihoods of making a change (*i.e.* the coin toss matters whether before the toss the coin-flipper says he/she has a 20%, 50%, or 90% likelihood of making the change). The coin toss was roughly equally influential on men and women, the old and the young, and across income levels. The coin

minor decisions [*e.g.* what quality of baseball card to offer (List, 2002), whether to respond to a solicitation letter from a charity (Falk, 2007), and when to make mail-order catalogue purchases (Anderson and Simester, 2003)].

2. To answer questions like that, previous research has typically had to rely on correlational studies (*e.g.* Kalmijn *et al.*, 2009; Pedersen and Schmidt, 2014) or natural experimental variation (*e.g.* Gruber and Mullainathan, 2005; Meier and Stutzer, 2007), with the usual challenges to causal inference.

toss, not surprisingly, had the biggest impact on relatively unimportant decisions like whether or not to go on a diet, but also influenced much more important choices like job quitting and ending relationships. The coin toss only influenced decisions made within the first two months of the coin toss; later changes were unrelated to the outcome of the toss.

Fourth, when it comes to "important" decisions (*e.g.* job quitting, separating from your husband or wife), making a change appears to be not only correlated with increased self-reported happiness, but also causally related, especially six months after the coin toss.[3] Those who were instructed by the coin toss to make a change were both more likely to make the change (as noted above) and, on average, report greater happiness on the follow-up surveys. This finding is inconsistent with expected utility theory; those who are on the margin should, on average, be equally well off regardless of the decision they make. This result provides strong empirical support for the notion of a status quo bias (Samuelson and Zeckhauser, 1998; Kahneman *et al.*, 1991). There is suggestive evidence that the coin toss outcome on "less important" decisions (*e.g.* going on a diet, dying one's hair, quitting a bad habit) influences future happiness in a similar, but more muted, fashion.

Fifth, for all decisions—not just the most important ones—there appears to be a causal impact of making a change on how satisfied the subject is *ex post* with the decision. Those who were instructed to make a change by the coin toss are substantially more likely to report that they made the correct decision and that they would make the same decision again if given the chance.

All of these results are subject to the important caveats related to using self-reported happiness as a proxy for utility, a research subject pool that is far from representative, potential sample selection in which coin flippers complete the surveys, and responses that might not be truthful. I consider a wide range of possible sources of bias and where feasible explore these biases empirically, concluding that it is likely that the first-stage estimates (*i.e.* the effect of the coin toss on decisions made) represent an upper bound. There is less reason to believe, however, that there are strong biases in the 2SLS estimates (*i.e.* the causal impact of the decision on self-reported happiness).

The structure of the remainder of the article is as follows. Section 2 describes in greater detail the experiment and how it was carried out. Section 3 reports the results of the experiment. Section 4 explores how a variety of potential biases might influence the inferences drawn from the study and also considers how likely those biases are to be important. Because this study differs in substantial ways from standard experimental interventions by economists, the issues of bias that arise are not the typical ones economists are used to thinking about. Section 5 concludes.

2. EXPERIMENTAL DESIGN

The experiment was carried out online at the website www.FreakonomicsExperiments.com.[4] Users who arrived at the site were greeted with the home page shown in Figure 1, which offered

3. Richard Easterlin was one of the first economists to be widely recognized for work with self-reported happiness data, and since his contribution in 1974 on the link between income and subjective happiness many others have made use of such data. Dolan *et al.* (2008) and Frey and Stutzer (2002) provide overviews of the use of self-reported happiness data in the economics literature. Additional applications of happiness data in the field are outlined by Di Tella and Macculloch (2006) who conclude that, treated with caution, the data have the potential to add value to empirical work. Researchers differ in their level of optimism regarding the validity of such data—Kahneman and Krueger (2006) note that the cleanest use of self-reported happiness data would "avoid effects of judgment and of memory as much as possible" but acknowledge that subject to these limitations such data can add important contributions to the field, while Bertrand and Mullainathan (2001) offer skepticism in noting that the use of a dependent variable that relies on self-reported happiness data can be problematic because "the measurement error appears to correlate with a large set of characteristics and behaviours."

4. For a further description of the experiment and preliminary results, written for a popular audience, see Dubner and Levitt (2014).

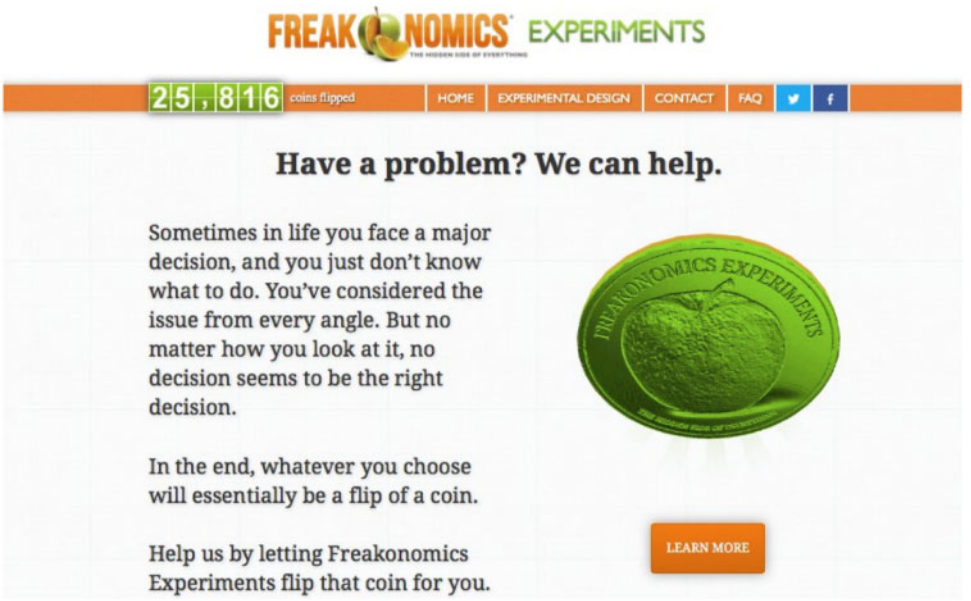

FIGURE 1
Website home page.

to help people make decisions through the use of a coin flip. Those individuals who clicked "Learn More" saw the screen-shot presented in Figure 2. If they proceeded further, they were shown a menu of life decisions over which to flip a coin from which they could choose; they were also given the option of designing their own customized question. After selecting a question relevant to their particular dilemma, subjects filled out a short survey that collected basic demographic data, asked them to rate their current level of happiness, probed them about the decision they were having trouble making, and gave them the opportunity to identify a third party, typically a friend or family member, who could be surveyed in the future regarding their decision.[5] Approximately 30% of subjects provided the name and email address of a third party. This sub-sample of the data is of particular interest for two reasons. First, naming a third party may signal greater commitment to following the coin toss. Second, the existence of a third party provides an independent source of information to verify later participant responses, as well as a source when the subject fails to respond to follow-up surveys.

The participants were then led to a page where a simulated coin tied to a randomizing algorithm was flipped and came up either heads or tails.[6] Subjects were reminded of what action the coin toss directed them to take, and if the coin toss said to make a change, they were encouraged to

5. Users were also shown, at random, a fact relevant to the decision they were about to make. For instance, those pondering whether to quit their job were told either "The number of job openings is on the rise—up by nearly 70% since 2009" or "Workers who dislike their jobs report lower levels of wellbeing than the unemployed. In fact, 81% of the unemployed report that they are happy every day compared to only 69% of the unhappily employed." There are no statistically significant differences in actions associated with having seen different facts.

6. Before the coin toss took place, subjects were asked how likely they were to make the change. If subjects indicated that they were very likely or very unlikely to make a change, they were taken to a page telling them that it seemed like they had already made up their mind. Those subjects then had the option of proceeding to the coin toss or exiting. All users were given the choice of having their outcome determined by a single coin toss, or could opt for a "best two out of three." Approximately 56% of users chose the "two out of three" option. In terms of subsequent behavior,

FIGURE 2
What potential study participants saw when they clicked "Learn More".

make that change within the next two months. In those cases where the coin toss said don't make a change, the subjects were told to maintain the status quo for at least the next two months (*e.g.* if the coin toss said not to quit one's job, the subjects were asked to remain at the job for at least two months). In most, but not all cases, heads was associated with making a change and tails was associated with maintaining the status quo. For simplicity in exposition, I refer to heads in what follows as meaning that the coin toss recommended a change.

Subjects were aware that they were part of an experiment and were required to explicitly give their informed consent. Both the subjects and the third parties provided by the subjects were then surveyed two and six months after the coin toss. Survey reminders were sent via email and included a link to an online survey site where the follow-up surveys were done. In order to encourage survey completion, those who filled out the surveys were provided with small gifts that took the form of exclusive content from Freakonomics podcasts. It should be noted, however, that I intentionally made it difficult for subjects to determine the precise objective of the study. Subjects were told that their participation would "help us gain important insights into decision-making." The initial survey, prior to the coin toss, asked many questions about motivations and feelings surrounding the decision. The follow-up surveys also asked a number of questions unrelated to the actual purpose of the study.

The website FreakonomicsExperiments.com was launched on 23 January 2013. Recruiting was done through a variety of online and traditional media avenues including reddit.com, the Freakonomics podcast, the Freakonomics blog, Marginal Revolution, and articles published in The Financial Times and Forbes. Data collection at the site remained active for roughly a year, after which a scaled down version of the site remained operational, but all survey activity ended.

there are no clear differences between those who went for the single coin versus best of three option. In what follows, I use the shorthand of a coin toss to refer to both of these options.

TABLE 1
Question attributes

Question	Number of tosses	Important? question?	Choice between action and Status Quo?
Should I quit my job	2,186	Yes	Yes
Should I break up	1,686	Yes	Yes
Should I go back to school	1,203	Yes	Yes
Should I start my own business	893	Yes	Yes
Should I move	762	Yes	Yes
Should I quit smoking	499	Yes	Yes
Should I have a child	415	Yes	Yes
Should I propose	220	Yes	Yes
Should I retire	120	Yes	Yes
Should I adopt	42	Yes	Yes
Create your own question	3,485	No	No
Should I splurge	1,491	No	Yes
Should I go on a diet	1,134	No	Yes
Should I break my bad habit	984	No	Yes
What should I major in	959	No	No
Should I get a tattoo	876	No	Yes
Should I try online dating	699	No	Yes
What college should I go to	656	No	No
Should I join a gym	630	No	Yes
Should I dye my hair	514	No	Yes
Should I sign up for a running event	431	No	Yes
Where should I move to	425	No	No
Should I grow facial hair	424	No	Yes
Should I quit drinking	401	No	Yes
Should I ask for a raise	385	No	Yes
Should I start volunteering	364	No	Yes
Should I rent or buy	295	No	No
What school should I send my child to	130	No	No
Should I get a roommate	106	No	Yes
Which house should I buy	96	No	No

Notes: This table presents summary information by question. The first column displays the number of coins tossed for each question. The second column indicates whether the question is considered an important question, where important questions are displayed in the top panel of the table. The third column indicates whether a question represents a choice between action or maintaining one's status quo (Yes) as opposed to a choice between two possible actions (No).

During the time of the study, there were approximately 165,000 unique visitors. Roughly 23,500 coin tosses took place. Excluded from the analysis are coin tosses with technical problems (primarily as a result of the user providing a faulty email address), leaving 22,511 usable coin tosses.

The distribution of these coin tosses across questions is presented in Table 1. Questions are divided into two categories corresponding to the importance of the decision for a person's life. This classification is based on a survey of individuals who were not participants in the original experiment.[7] I use this classification to aggregate questions later in the article. "Important" questions are listed first in the table, followed by "less important" questions. Of the important questions, the single most popular was "Should I quit my job?" which attracted 2,186 coin tosses. The other "important" questions which yielded more than 1,000 coin flips were "Should I break up

7. These raters were asked to rate the importance of each life decision on a scale from 1 to 5. The correlation in rankings across individuals is quite high, with an average pairwise correlation of 0.707. The cutoff between "important" and "less important" is by necessity somewhat arbitrary. There was a large gap in ratings between "Should I move?" (average rating of 3.45) and "Should I go on a diet?" (average rating of 3.0), so I divided the sample there. The central findings of the paper are reproduced if instead a continuous measure of importance is utilized.

with my significant other?" and "Should I go back to school?" Among "less important" questions, over 3,000 individuals created their own questions. I mostly ignore these questions in the analysis that follows. Other popular choices related to splurging and going on a diet.

Online surveys of both the participants and the third parties were conducted two and six months after the coin toss. The surveys of coin flippers reminded the recipient which question had led to a coin being tossed (but did not remind them of the outcome of the coin toss), and then asked, among other questions, (1) whether an action had been taken since the coin toss and (2) about his/her overall happiness level and the degree of satisfaction with the specific decision on the coin toss question. Third parties were asked a parallel set of questions, appropriately rephrased.[8] For questions where a decision was essentially permanent (*e.g.* quitting a job), subjects were asked whether they had taken the action. On topics for which a change was potentially temporary (*e.g.* attempting to quit smoking which might succeed or fail), we asked subjects whether the attempt had been made.

Figure 3 reports the degree of success in obtaining follow-up surveys. There is at least one completed survey from roughly 58.34% of the coin flippers who did not name a third party. Those who named a third party before the coin toss were more likely (77.39%) to complete at least one survey, consistent with the conjecture that naming a third party signals commitment to the experiment. Adding in the surveys filled out by the third parties, I have at least one follow-up survey for 83.57% of the coin flippers who named a third party. Response rates were higher for the two-month survey (a total of 13,935 completed surveys) than the six-month survey (8,159 completed surveys). Throughout the analysis, except where noted, I analyse the two-month and six-month samples separately.

3. RESULTS

There are two questions of primary interest: (1) Did the coin toss influence behaviour? and (2) What can be learned about the impact of choices on subsequent happiness? I begin with an analysis of the first question before turning to the second question. In this section, I simply report the data generated by the experiment and the treatment effects that arise from those data. There are many potential sources of bias that might arise as a result of survey non-response and untruthful responses on the part of subjects. I defer careful consideration of these potential biases to Section 4.

3.1. *Did the outcome of the coin toss influence behaviour?*

Figure 4 presents data on the rate of coin toss adherence among survey respondents. The green bars correspond to two-month responses; blue represents data from the six-month survey. The values reported in the columns are the percentage of coin flippers whose actions correspond to the dictate of the coin toss, *i.e.*, making a change if heads came up and maintaining the status quo if tails was the outcome.[9] If the coin toss has no impact on behaviour, then 50% of the actions taken should match the coin's dictate. The first two bars in Figure 4 reflect data from all coin tosses. After two months, roughly 63% of the respondents' actions match the recommendation of the coin toss. This implies that 13% of all actions were affected by the coin toss, *i.e.*, that someone

8. Third parties were only asked about the general happiness level of the coin flipper, not about the specific choice (*e.g.* if the coin flipper could go back in time and make the decision again, would they make the same choice). *Ex post*, this is a research design decision that I regret.

9. For those cases where I have survey responses from both the coin flipper and the third party, and they disagree as to what action was taken, I use the stated action of the coin flipper.

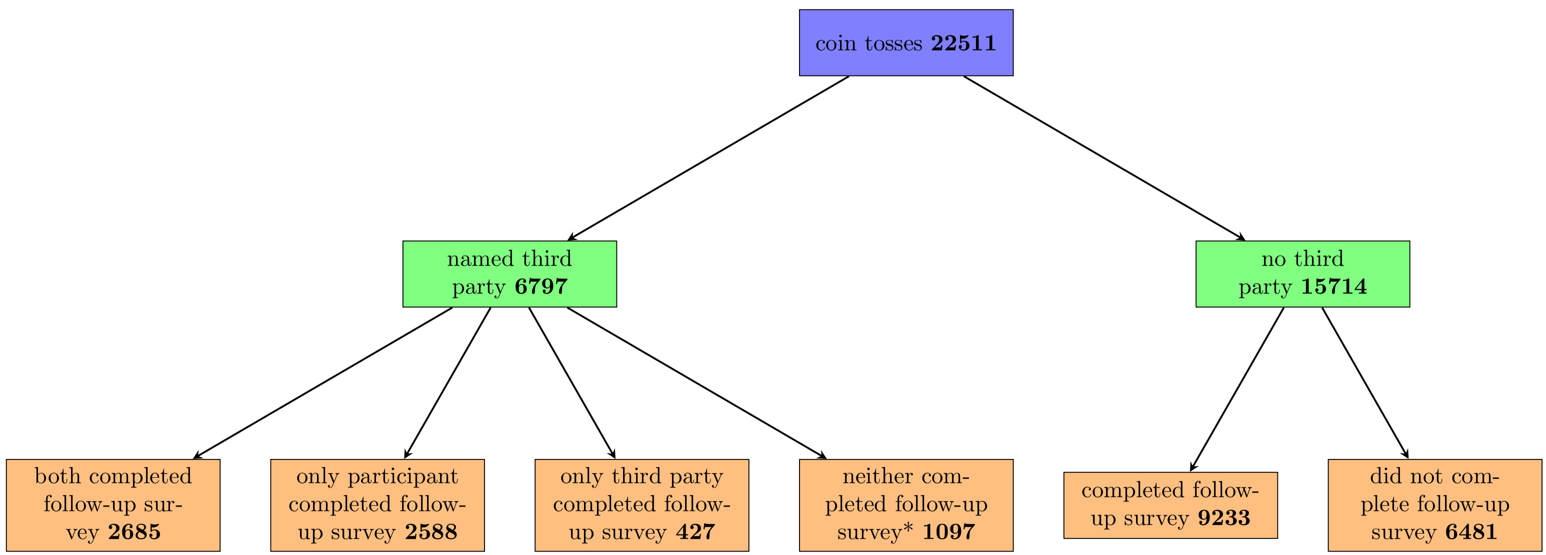

FIGURE 3

Follow-up survey response rates

Notes: This figure presents the number of total tosses and the number of completed surveys according to whether a third party was named. Note that the category consisting of participants who did not complete a follow-up survey includes those who did not complete their first follow-up survey and who never received a second follow-up survey because the experiment ended before the participant would receive this second survey. 3,655 participants never received their second follow-up survey. 798 participants did not receive their first follow-up survey due to the experiment's end date and were thus excluded from our analysis.

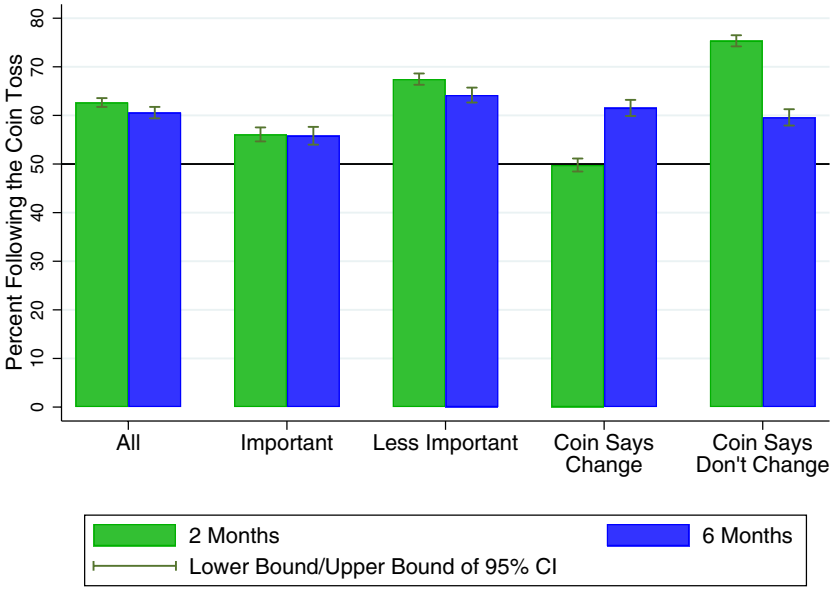

FIGURE 4

Coin toss adherence among survey respondents

Notes: This figure presents coin toss adherence based on two- and six-month survey responses. The vertical axis reflects the percent following the coin toss. The horizontal axis categorizes response rates by question type and survey.

who got heads was 26 percentage points more likely to have made a change than someone who got tails. The corresponding numbers, here and in the remainder of the article, are slightly lower at six months. This implies that some part of the impact of the coin toss is to accelerate changes that would have happened anyway, but at a later date.[10]

The next two sets of columns in Figure 4 divide the sample between "important" and "less important" questions, as defined above. On "important" questions, the rates of reported coin-toss adherence are much lower than for the full sample (56.1% at two months; 55.8% at six months), but still above 50%. For "less important" questions, more than 67% of the subjects report following the coin toss at two months. The final two sets of columns parse the data according to whether the coin says to make a change or recommends maintaining the status quo. At two months, there is a bias towards the status quo. Only half of the respondents told to make a change do so, whereas 75% of those told to maintain the status quo do so. At six months, roughly 60% of participants follow the coin toss whether it comes up heads or tails.

Prior to the coin toss, participants were asked to report how likely they believed they were *ex ante* to take the action associated with their coin toss, *e.g.*, to propose to their significant other. They were given a menu of choices ranging from 0% to 100% at 10% intervals.[11] Figure 5 plots the impact of the coin toss as a function of these *ex ante* likelihoods. The horizontal axis

10. The fact that 13% of actions were affected by the coin toss has several implications. First, as hypothesized earlier, it indicates that many people are on the margin when making a decision. More interestingly, it means some people would prefer to give up control of their decision-making, even to something as arbitrary as a randomization device. One potential mechanism could be regret aversion—regret is a product of decisions that one has control over, so by giving up control, one minimizes regret.

11. The average predicted probability of taking the action across the research subjects was 41.94%. 8.38% predicted that there was no chance of changing; 2.58% thought they would change for sure. The most popular response was 50%.

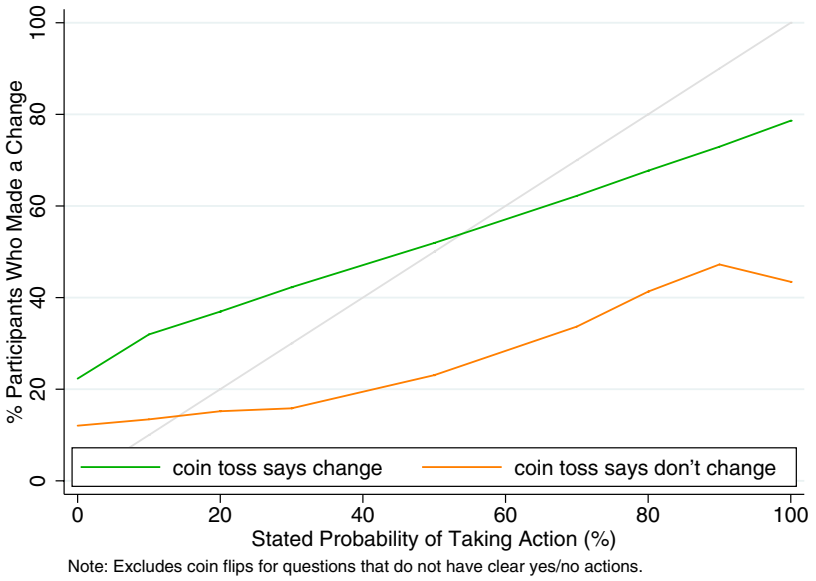

FIGURE 5

Likelihood of taking action as a function of ex ante stated probabilities, two-month survey

Notes: This figure presents the percent of participants who make a change by the two-month survey mark according to their stated probability of changing and the result of the coin flip. The vertical axis reflects the percent of respondents who reported making a change. The horizontal axis groups respondents according to to their stated ex ante likelihoods of making a change. Responses are categorized according to whether the coin came up heads (make a change) or tails (no change).

corresponds to the participants' stated likelihood of taking an action, prior to tossing the coin. The vertical axis is the percentage of subjects who report taking the action on the two-month survey. The two lines plotted in the figure correspond to those whose coin tosses came up heads and tails respectively. A number of insights emerge from the figure. First, the outcome of the coin toss exerted influence across the entire distribution of *ex ante* probabilities. This can be seen in the fact that the line corresponding to heads is above the line for tails across the entire span of the graph by an average of roughly 20 percentage points. The coin toss had the smallest impact (*i.e.* the two lines are closest together) when the self-proclaimed likelihood of a change was small. A second fact that emerges from the figure is that the lines in the graph slope upward, meaning that the *ex ante* probabilities are correlated with actual actions. The predictions by the subjects are not particularly accurate, however, as the slopes of the lines are well below the 45 degree line. A non-trivial share of those who said that they would take a particular action (or non-action) with certainty did the opposite. Finally, there is some evidence of a bias towards inaction in the two-month survey data. Since roughly half the participants got heads and half tails, the overall likelihood of taking the action falls halfway between the two lines in the figure. For *ex ante* probabilities above 30%, the actual rate at which the action is taken is less than was predicted by the individuals. The gap is most extreme among those who predicted they would make a change with 100% certainty. In fact, only about 80% of those participants made a change in response to heads, and less than half actually changed when the coin came up tails.

Figure 6 is identical to Figure 5, except that it shows results for the six-month survey rather than the two-month survey. The general patterns observed are similar, with one notable difference. Any evidence of a bias towards inaction has disappeared. Overall, after six months, the action is

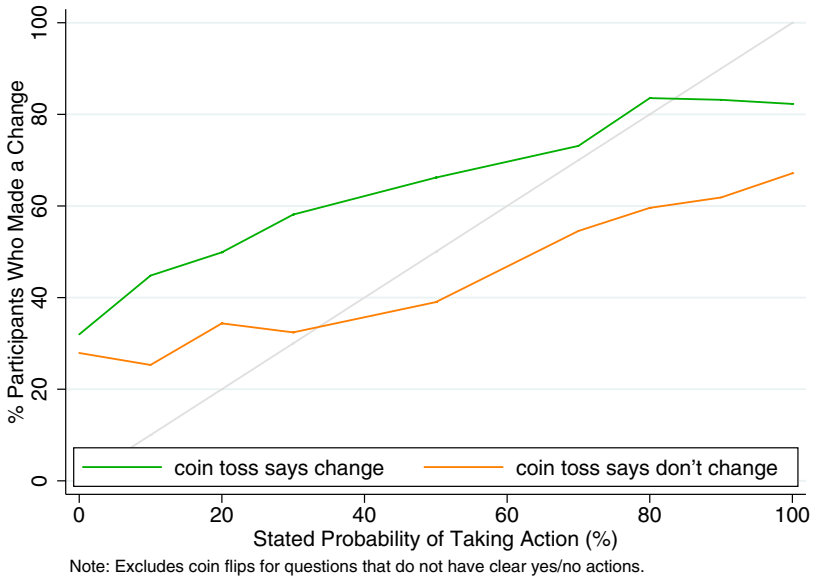

FIGURE 6

Likelihood of taking action as a function of ex ante stated probabilities, six-month survey

Notes: This figure presents the percent of participants who make a change by the six-month survey mark according to their stated probability of changing and the result of the coin flip. The vertical axis reflects the percent of respondents who reported making a change. The horizontal axis groups respondents according to to their stated ex ante likelihoods of making a change. Responses are categorized according to whether the coin came up heads (make a change) or tails (no change).

taken slightly more frequently than predicted *ex ante* by the participants.[12] It should be noted, however, that the *ex ante* probabilities refer to the likelihood of making a change within two months, not within six months.

Figure 7 shows the impact of the coin toss on actions across individual questions. Included in the figure are the results for every question with at least 150 responses. The top portion of the figure reports findings for the questions deemed "important;" the bottom part of the figure corresponds to "less important" decisions. The values reported in the figure are the percentage of all respondents to the two-month survey who report taking the action that corresponds to the coin outcome. With the exception of "Should I move?" which shows no impact of the coin toss, for all the other "important" choices between 55% and 60% of the subjects report following the suggestion of the coin on the two-month survey. Decisions on "less important" questions, as might be expected, are more affected by the coin toss, with the highest compliance rate on "Should I break my bad habit" (over 80%), "Should I go on a diet," "Should I quit drinking," and "Should I try online dating." Supplementary Appendix Figure 1 is identical to Figure 7, except that it shows results for the six-month survey rather than the two-month survey. The patterns are similar.

All of the numbers presented thus far are raw data. Table 2 demonstrates that the impact of the coin toss is both robust to the inclusion of covariates and is highly statistically significant. Each column of Table 2 reports the results of a linear probability model in which the dependent variable is a dichotomous variable corresponding to whether the survey respondent says a change was made. Included as right-hand side variables are the result of the coin toss, how likely the

12. Supplementary Appendix Figures 2–5 mirror Figures 5 and 6, but divide the sample into "important" and "less important" questions. The same patterns are present, except that the gap between the lines for "important" questions is smaller throughout because of the reduced influence of the coin toss.

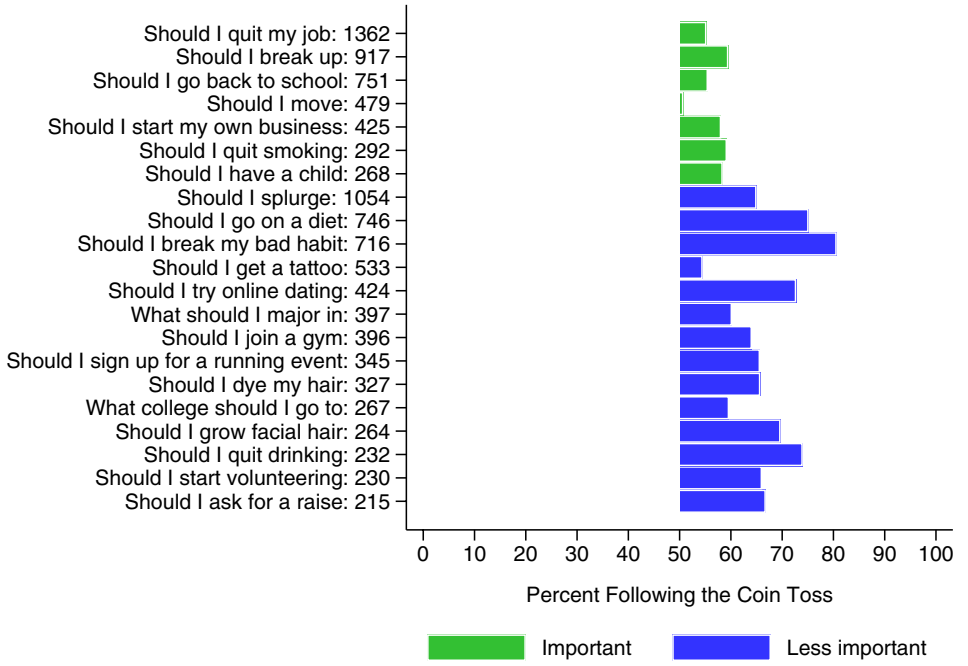

FIGURE 7

Percentage following the coin toss, two-month survey

Notes: This figure presents the percentage of all respondents to the two-month survey who report taking the action that corresponds to the result of the coin toss. The questions are listed on the vertical axis and are divided into "important" and "less important" groupings. Questions with fewer than 150 responses were excluded from this figure.

subject said they were to change *ex ante*, a range of demographic variables, whether the subject opted for the "best two out of three coin toss" option, and an indicator variable for the particular question for which the coin was tossed. Columns 1 and 4 reflect the whole sample. Columns 2 and 5 are the subset of "important" questions, and columns 3 and 6 correspond to the "less important" questions. The top row is the coefficient on the coin toss coming up heads. For all questions on the two-month survey, individuals who got heads report being 24.9 percentage points more likely to have made a change than those who got tails. This result is highly statistically significant. The point estimate at six months is slightly smaller (0.211), implying that some of the impact of getting heads operates through accelerating the timing of a change. Comparing important questions (columns 2 and 5) to less important questions (columns 3 and 6), the impact of the coin toss is only about one-third as large for important questions, but is still highly statistically significant. The coin-flipper's *ex ante* assessment of how likely he or she is to make a change is also highly informative about whether a change is eventually made. If the subjects made unbiased forecasts, the coefficient on this variable would be one; in actuality it ranges between 0.279 and 0.597. Subjects are better predictors of their own behaviour on important questions than on less important ones. The only other variable which has a strong and consistent relationship to making a change is age. Older subjects are less likely to make changes, especially on important questions.

3.2. *Is there a causal impact of making a change on happiness and satisfaction with the decision?*

The results above suggest that the outcome of the coin toss affected the behaviour of some participants. Consequently, the coin toss has the potential to shed light on the question of whether

TABLE 2
The impact of the coin toss on subsequent behavior

	Two months after coin toss			Six months after coin toss		
	All	Important	Less important	All	Important	Less important
Heads	0.249***	0.111***	0.364***	0.211***	0.112***	0.295***
	(0.009)	(0.012)	(0.012)	(0.012)	(0.017)	(0.016)
Prob of change	0.445***	0.594***	0.279***	0.476***	0.597***	0.341***
	(0.017)	(0.023)	(0.024)	(0.023)	(0.032)	(0.033)
Male	0.012	0.005	0.018	−0.001	−0.001	−0.003
	(0.009)	(0.013)	(0.012)	(0.012)	(0.018)	(0.017)
Age	−0.002***	−0.003***	−0.002*	−0.002**	−0.006***	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Married	0.002	−0.014	0.013	−0.014	−0.037	0.015
	(0.011)	(0.015)	(0.016)	(0.015)	(0.021)	(0.021)
US resident	0.033***	0.020	0.039**	0.008	0.002	0.007
	(0.010)	(0.014)	(0.013)	(0.014)	(0.020)	(0.018)
Black	0.006	−0.029	0.041	−0.046	−0.127*	0.042
	(0.027)	(0.035)	(0.041)	(0.039)	(0.053)	(0.055)
Asian	0.004	0.008	−0.012	−0.022	−0.041	−0.023
	(0.013)	(0.019)	(0.018)	(0.018)	(0.027)	(0.025)
Hispanic	0.019	0.007	0.027	−0.011	−0.013	−0.008
	(0.017)	(0.024)	(0.023)	(0.023)	(0.034)	(0.032)
Race-other	0.004	0.002	0.004	0.010	0.051	−0.038
	(0.022)	(0.031)	(0.030)	(0.032)	(0.046)	(0.043)
4-year college	−0.007	−0.013	−0.008	−0.003	0.008	−0.023
	(0.010)	(0.016)	(0.014)	(0.015)	(0.023)	(0.019)
Income > 50K	0.000	−0.015	0.016	−0.008	−0.011	0.003
	(0.010)	(0.014)	(0.014)	(0.014)	(0.020)	(0.019)
Live in a city	0.007	−0.003	0.013	−0.008	−0.003	−0.011
	(0.009)	(0.013)	(0.012)	(0.012)	(0.018)	(0.016)
Pre-toss happiness	−0.002	0.002	−0.003	−0.006	−0.004	−0.006
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
Best 2 of 3 flip	−0.009	−0.006	−0.010	−0.013	−0.006	−0.019
	(0.011)	(0.016)	(0.014)	(0.016)	(0.024)	(0.021)
Include question indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,094	4,607	5,487	6,131	2,874	3,257

Notes: This table explores the impact of the coin toss on participants' subsequent behavior. Each column reports the results of a linear probability model in which the dependent variable is a dichotomous variable that corresponds to whether the survey respondent says a change was made. Columns 1 and 4 reflect two- and six-month survey responses, respectively, from the entire sample. Columns 2 and 5 present the same information for the subset of important questions, and Columns 3 and 6 correspond to the less important questions. Standard errors are reported in parentheses. *, **, *** denote significance at the 5, 1, and 0.1% levels.

making a particular change (*e.g.* going on a diet) has a positive or negative impact on self-reported happiness and other proxies for whether the right choice was made. Before the coin toss, those who will get heads are, in expectation, identical in all respects to those who will get tails. If the only channel through which the coin toss operates is to increase the likelihood that the particular change in question is made, then the coin toss can serve as an instrumental variable.

More formally, let H represent happiness, which is influenced by the choice of whether or not to take some binary action A. Additionally, let the set of all other factors that influence H be captured by some vector of variables X. For instance, relevant X's might include the salary of one's current job, what city one lives in, the level of education, how happily married one is, etc. Some of these X's might be observable, but many would not be. A simple comparison of happiness amongst those who take the action ($A=1$) versus those who do not ($A=0$), *i.e.*,

$$E[H|A=1]-E[H|A=0]$$

is unlikely to have a causal interpretation because X is not held constant across those who do and do not switch jobs. Empirically, those who make a change are statistically significantly younger, less likely to be married, less educated, and lower income than those who do not make a change. While it is possible to control for these observable factors, it is likely that these two groups differ substantially on unobservable dimensions as well. *A priori*, the sign of the bias in OLS is not obvious.

OLS suffers from a second weakness: a simple comparison of *everyone* who quits their job to *everyone* who does not quit their job does not answer the economically interesting question. When considering the impact of making a change, it is the *marginal* actor who is of primary interest. There are many happily married couples and a few that are so disastrously unhappy that divorce is certain. A comparison of these two sets of couples tells us nothing about how getting divorced will affect the happiness of the couples who are truly marginal.

The outcome of the coin toss, used as an instrumental variable, potentially solves both of those problems. Let C represent an indicator variable corresponding to 1 if the coin comes up heads and 0 otherwise. Under the assumptions that

$$E[A|C=1]-E[A|C=0]\neq 0 \text{ and}$$
$$E[X|C=1]-E[X|C=0]=0$$

then a simple Wald estimator provides an estimate of the causal impact of action A on happiness H

$$\hat{B}_{\text{Wald}}=\frac{E[H]|C=1-E[H]|C=0}{E[A]|C=1-E[A]|C=0}.$$

As long as the only channel through which the coin toss operates is via influencing the likelihood that the action in question is taken, then the Wald estimator represents a local average treatment effect on H of taking the action A, for that group whose behaviour is influenced by the coin toss, *i.e.*, the people who are so marginal that they are willing to have their action swayed by a coin toss.

Turning to the empirical findings, participants were asked five questions designed to ascertain their satisfaction with life as a whole: (1) general level of happiness on a seven-point scale, (2) how the subject believes friends would rate himself/herself on a seven-point happiness scale, (3) whether the subject is better off, worse off, or the same relative to the point in time when the coin was tossed. Two further questions focused more specifically on the decision for which they flipped a coin: (4) does the subject feel he/she made the correct decision on the choice for which the coin was tossed, and (5) if the subject could go back in time, would he/she make the same decision again. Questions 1 and 2 were asked on both the two-month and six-month surveys. Question 3 was only asked on the six-month survey, and questions 4 and 5 were only asked at two months.

Table 3 shows the degree of within-respondent correlation across these various outcomes. The top and middle panels of Table 3 report results for the two-month and six-month surveys, respectively. The bottom panel correlates responses across the two-month and six-month surveys. On the two-month survey, the two questions addressing happiness (*i.e.* the standard measure of self-reported happiness and how the subject thinks friends would rate his/her happiness) have a correlation of 0.666. These two happiness measures are relatively weakly correlated with whether someone reports having made the correct decision or whether they would have made the same choice with perfect foresight. At six months, the happiness measures and reporting being better or worse off now compared to the time of the coin toss are all relatively highly correlated. The bottom

TABLE 3
Correlations across self-reported outcome measures within and across surveys

Panel A: Two-month survey

	Happiness	Appear happy	Correct decision	Perfect foresight
Happiness	1.000			
Appear happy	0.666	1.000		
Correct decision	0.177	0.117	1.000	
Perfect foresight	0.104	0.054	0.278	1.000

Panel B: Six-Month survey

	Happiness	Appear happy	Better/worse off
Happiness	1.000		
Appear happy	0.701	1.000	
Better/worse off	0.485	0.353	1.000

Panel C: Correlation across the two- and six-month survey

		Two-month survey			
		Happiness	Appear happy	Correct decision	Perfect foresight
Six Month	Happiness	0.465	0.360	0.102	0.031
	Appear happy	0.390	0.466	0.066	0.014
	Better/worse off	0.143	0.097	0.132	0.036

Notes: Panel A reports pairwise correlations in responses for study participants on the two-month survey. Panel B presents parallel correlations, but for the six-month survey. Panel C reports correlations across time for participants who completed both two-month (columns) and six-month surveys (rows). The results in this table include responses for both important and less important questions.

panel reports correlations between the two-month outcomes (columns) and six-month outcomes (rows). The direct happiness measures are much more strongly correlated than the others.[13]

Table 4 presents the basic empirical findings regarding the link between choice and subsequent life satisfaction outcomes. Columns 1–8 correspond to the two-month survey; columns 9–14 reflect six-month survey responses. For each outcome question asked, we report both OLS estimates (odd columns) and 2SLS estimates (even columns). The OLS estimates reflect differences in outcomes across those who made a change and those who maintained the status quo. The OLS estimates are explicitly correlational—to the extent that people who do and do not make a change differ systematically, the OLS estimates will not have a causal interpretation. In contrast, under the assumption that the only channel through which the outcome of the coin toss affects happiness is through the choice made, the instrumental variable estimates in the even columns capture the causal impact of the action on subsequent outcomes. The first panel of the table presents results aggregated across all the questions. The second and third panels also report aggregated data, but classifying questions as either "important" or "less important."[14] Each entry in the table is from a different regression. Only the key coefficient of interest is presented in the

13. Limiting Table 3 to the most important questions leads to somewhat higher correlations between the happiness measures and the questions that more narrowly relate to the decision surrounding the coin toss. This would be expected, since those decisions carry more significant life implications.

14. I limit the sample of questions to those in which the coin flippers are making a choice between a change and the status quo. This eliminates questions like "Should I attend college A or college B?" Since colleges A and B are different across people, it is difficult to know how to evaluate such questions. The same is true with the widely varying "create your own" questions, which are also excluded.

TABLE 4
The link between choices and self-reported happiness (all outcomes)

	Two months after coin toss								Six months after coin toss					
	Happiness		**Appear happy**		**Correct decision**		**Perfect foresight**		**Happiness**		**Appear happy**		**Better/worse off**	
Question	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
All	0.449	0.041	0.309	0.236	0.173	0.325	0.079	0.235	0.584	0.476	0.442	0.149	0.109	0.167
	(0.039)	(0.139)	(0.038)	(0.134)	(0.006)	(0.024)	(0.007)	(0.027)	(0.048)	(0.214)	(0.046)	(0.207)	(0.009)	(0.038)
	[μ=6.837]		[μ=7.161]		[μ=0.593]		[μ=0.852]		[μ=7.059]		[μ=7.312]		[μ=0.756]	
Important	0.782	0.554	0.588	1.070	0.151	0.456	0.034	0.285	1.011	2.153	0.717	1.418	0.146	0.412
	(0.066)	(0.495)	(0.064)	(0.491)	(0.010)	(0.085)	(0.010)	(0.082)	(0.076)	(0.652)	(0.073)	(0.619)	(0.013)	(0.112)
	[μ=6.566]		[μ=6.943]		[μ=0.630]		[μ=0.892]		[μ=6.932]		[μ=7.207]		[μ=0.777]	
Less important	0.213	−0.073	0.111	0.038	0.186	0.290	0.107	0.218	0.190	−0.077	0.168	−0.266	0.075	0.087
	(0.047)	(0.119)	(0.045)	(0.115)	(0.008)	(0.022)	(0.010)	(0.027)	(0.061)	(0.194)	(0.059)	(0.189)	(0.012)	(0.038)
	[μ=6.999]		[μ=7.291]		[μ=0.571]		[μ=0.828]		[μ=7.139]		[μ=7.378]		[μ=0.743]	

Notes: This table presents regression results exploring the link between choices and various metrics of happiness. The first column within each metric, "OLS", shows the extent to which those who make a change are more or less happy (as measured by that metric) than those who maintain the status quo. The second column is the mean value of the outcome variable. The third column in each metric, "2SLS", are the instrumental variable estimates. The left-hand side panels corresponds to the two-month survey; the right-hand side panel corresponds to the six-month survey. "Happiness" refers to self-reported happiness. "Appear happy" refers to a participant's guess of how happy their friend would say the participant is. "Correct decision" equals 1 if the subject feels they made the correct decision two months ago, equals 0 if they feel they made the wrong decision, and .5 otherwise. "Perfect foresight" equals 1 if the subject, given perfect foresight, would have made the same decision two months ago, equals 0 if they would have made a different decision, and .5 otherwise. "Better/worse off" equals 1 if the subject thinks they are better off than they were six months ago, equals 0 if they think they are worse off than six months ago, and .5 otherwise. Standard errors are reported in parentheses.

table. In all specifications, I include a basic set of control variables mirroring those included in the first-stage regressions reported earlier. Full results are available in an Supplementary Appendix. For each question, the mean of the outcome variable is displayed in square brackets.

The OLS results carry a positive and statistically significant coefficient in all 21 possible cases. This means that those who make a change report increased happiness/satisfaction with the choice made relative to those who maintain the status quo. In five of the seven columns, the coefficient is larger for important decisions than for less important decisions. The magnitude of the coefficients is substantial. For instance, on happiness, those who make a change are roughly 0.5 points higher on a 10 point scale, or nearly one-fifth of a standard deviation. As argued above, however, these OLS coefficients need not imply causality.

Indeed, the instrumental variable estimates tell a more nuanced story than do the OLS estimates. At two months, there is only weak evidence that making a change affects the happiness measures (the only coefficient that is borderline significant at the 0.05 level is "appearing happy" for the important questions), but there are large and highly statistically significant impacts on feeling that the correct decision was made and whether he/she would follow the same path with perfect foresight. At six months, making a change is associated with large and statistically significant increases on the happiness measures for important questions, but not for less important questions. On both categories of questions, but especially important ones, the 2SLS estimates imply that those participants making a change are more likely to be better off relative to six months ago. For important questions at six months, the 2SLS estimates are two to three times larger than the OLS estimates.

Table 5 reports results for individual questions, only for the happiness measure. Parallel results for the other outcome measures are presented in Supplementary Appendix Tables 2–5. The OLS estimates on the individual questions classified as important are uniformly positive and often statistically significant. Most, but not all, of the less important questions carry a positive OLS coefficient. The 2SLS estimates are imprecise. Job quitting and breaking up both carry very large, positive, and statistically significant coefficients at six months. Going on a diet is positive and statistically significant at two months, but has a small and insignificant impact by six months. Online dating is positive and significant at the 0.10 level at two months, but turns negative by six months. Splurging is negative and significant at the 0.10 level at two months, but has no discernible impact by six months. Attempting to break a bad habit is negative with a t-stat of 1.5 at both points in time, perhaps because breaking bad habits is so hard. For those subjects who reported trying to break a bad habit, third parties said the bad habit had actually been broken only 20.93% of the time at two months and only 24.49% at six months.[15]

Table 6 explores the sensitivity of the estimates on the happiness outcome across subsamples of the data. The columns in Table 6 match those of Table 4. The top row of Table 6 replicates the full sample results as a baseline. Relatively few strong patterns emerge in Table 6. With respect to the first stage, the most pronounced result that emerges is that (as expected) those who report being likely to follow the coin toss are, indeed, three to four times more likely to follow the coin flip. Those who name a friend (signalling greater commitment to the experiment) are also more likely to follow the coin flip. For the OLS estimates, older subjects have a greater increase in

15. As shown in the Supplementary Appendix, the results for the outcome of how happy one appears is broadly similar to those for the happiness measure. Stronger results are obtained on the question of whether the correct decision was made: the 2SLS coefficient is positive and statistically significant for breaking up, starting a new business, quitting smoking, going on a diet, breaking a bad habit, joining a gym, signing up for a running event, quitting drinking, asking for a raise, and starting to volunteer. On the perfect foresight question, quitting smoking, going on a diet, and breaking a bad habit are all positive and significant, while making a splurge is negative and significant. With respect to being better off relative to six months earlier, breaking up and joining a gym are both positive and significant.

TABLE 5
The link between choices and self-reported happiness

	Question	Two months after coin toss			Six months after coin toss		
		1st stage	OLS	2SLS	1st stage	OLS	2SLS
	All	0.249	0.449	0.041	0.211	0.584	0.476
		(0.009)	(0.039)	(0.139)	(0.012)	(0.048)	(0.214)
	Important	0.111	0.782	0.554	0.112	1.011	2.153
		(0.012)	(0.066)	(0.495)	(0.017)	(0.076)	(0.652)
	Less important	0.364	0.213	−0.073	0.295	0.190	−0.077
		(0.012)	(0.047)	(0.119)	(0.016)	(0.061)	(0.194)
Important	Should I quit my job	0.059	1.643	0.905	0.070	1.890	5.203
		(0.022)	(0.127)	(1.774)	(0.031)	(0.137)	(2.313)
	Should I break up	0.167	0.356	0.639	0.157	0.278	2.698
		(0.030)	(0.159)	(0.818)	(0.040)	(0.192)	(1.259)
	Should I go back to school	0.119	0.595	−0.583	0.133	0.949	0.007
		(0.030)	(0.168)	(1.162)	(0.042)	(0.190)	(1.280)
	Should I start my own business	0.168	0.399	0.000	0.077	0.520	5.256
		(0.046)	(0.185)	(1.014)	(0.074)	(0.307)	(5.707)
	Should I move	0.004	0.795	56.326	0.087	0.823	3.176
		(0.034)	(0.233)	(450.597)	(0.053)	(0.239)	(2.775)
	Should I quit smoking	0.129	0.160	1.417	0.147	0.313	−1.096
		(0.051)	(0.225)	(1.498)	(0.078)	(0.304)	(1.995)
	Should I have a child	0.195	0.471	−1.598	0.193	0.395	−0.450
		(0.046)	(0.261)	(1.083)	(0.073)	(0.270)	(1.288)
	Should I propose	0.183	0.362	1.021	−0.041	1.529	−5.125
		(0.075)	(0.506)	(1.862)	(0.124)	(0.640)	(19.881)
Less important	Should I splurge	0.303	0.197	−0.555	0.204	0.458	0.163
		(0.029)	(0.096)	(0.312)	(0.035)	(0.136)	(0.594)
	Should I go on a diet	0.488	0.413	0.754	0.471	0.146	0.154
		(0.032)	(0.126)	(0.252)	(0.044)	(0.176)	(0.361)
	Should I break my bad habit	0.607	0.146	−0.325	0.384	−0.001	−0.597
		(0.030)	(0.123)	(0.199)	(0.044)	(0.168)	(0.420)
	Should I get a tattoo	0.111	0.524	−0.775	0.139	0.669	−1.123
		(0.024)	(0.255)	(1.270)	(0.044)	(0.260)	(1.478)
	Should I try online dating	0.465	0.043	0.611	0.269	0.129	−0.429
		(0.044)	(0.180)	(0.377)	(0.060)	(0.249)	(0.846)
	Should I join a gym	0.236	0.690	0.369	0.288	0.292	0.970
		(0.045)	(0.188)	(0.686)	(0.065)	(0.209)	(0.671)
	Should I dye my hair	0.315	0.327	0.266	0.148	0.664	1.863
		(0.052)	(0.188)	(0.553)	(0.068)	(0.261)	(1.623)
	Should I sign up for a running event	0.265	0.437	−0.790	0.347	−0.234	−0.395
		(0.049)	(0.192)	(0.675)	(0.065)	(0.216)	(0.572)
	Should I grow facial hair	0.390	−0.137	−0.275	0.234	−0.726	−0.624
		(0.053)	(0.209)	(0.467)	(0.079)	(0.334)	(1.149)
	Should I quit drinking	0.446	−0.309	−0.427	0.278	−0.083	1.150
		(0.059)	(0.246)	(0.507)	(0.087)	(0.316)	(1.068)
	Should I ask for a raise	0.356	−0.037	−0.689	0.425	0.115	−1.116
		(0.064)	(0.276)	(0.712)	(0.087)	(0.375)	(0.833)
	Should I start volunteering	0.303	0.037	−0.135	0.478	0.090	0.129
		(0.054)	(0.274)	(0.714)	(0.071)	(0.303)	(0.510)
	Observations	10,094	10,094	10,094	6,131	6,131	6,131

Notes: This table presents regression results exploring the link between choices and self-reported happiness. Columns 1 to 3 correspond to the two-month survey; Columns 4 to 6 correspond to the six-month survey. Columns 1 and 4 are first-stage estimates and describe the degree to which the coin toss affected the action taken. Columns 2 and 5 are OLS estimates, which show the extent to which those who make a change are more or less happy than those who maintain the status quo. Columns 3 and 6 are the instrumental variable estimates. The row "Observations" reflects the number of observations in the regression that includes all questions. Questions with fewer than 150 respondents were included in the first panel but are not presented as separate regressions. Standard errors are reported in parentheses.

TABLE 6
Sensitivity analysis for all questions (dependent = happiness)

	Two months after coin toss				Six months after coin toss			
Question	**1st stage**	**OLS**	**2SLS**	**Observations**	**1st stage**	**OLS**	**2SLS**	**Observations**
All	0.249	0.449	0.041	10,094	0.211	0.584	0.476	6,131
	(0.009)	(0.039)	(0.139)		(0.012)	(0.048)	(0.214)	
Female	0.259	0.537	0.299	4,400	0.212	0.655	0.857	2,697
	(0.013)	(0.060)	(0.207)		(0.018)	(0.071)	(0.315)	
Male	0.242	0.382	−0.149	5,694	0.211	0.522	0.230	3,434
	(0.011)	(0.051)	(0.186)		(0.016)	(0.066)	(0.290)	
Younger than 30	0.265	0.335	−0.016	5,777	0.214	0.452	0.547	3,469
	(0.011)	(0.050)	(0.170)		(0.016)	(0.062)	(0.270)	
30 or Older	0.225	0.599	0.121	4,317	0.205	0.748	0.433	2,662
	(0.013)	(0.062)	(0.239)		(0.018)	(0.077)	(0.350)	
No friend named	0.214	0.427	0.178	6,368	0.185	0.564	0.750	3,752
	(0.011)	(0.051)	(0.208)		(0.015)	(0.062)	(0.315)	
Friend named	0.311	0.480	−0.116	3,726	0.251	0.624	0.178	2,379
	(0.014)	(0.060)	(0.175)		(0.019)	(0.076)	(0.284)	
Income below 50K	0.254	0.416	−0.173	5,504	0.201	0.451	0.237	3,289
	(0.012)	(0.053)	(0.188)		(0.016)	(0.065)	(0.304)	
Income above 50K	0.242	0.482	0.287	4,590	0.219	0.735	0.729	2,842
	(0.012)	(0.057)	(0.207)		(0.017)	(0.072)	(0.305)	
Report unlikely to follow toss	0.097	0.571	−0.220	3,947	0.064	0.786	1.871	2,420
	(0.013)	(0.070)	(0.598)		(0.019)	(0.081)	(1.187)	
Report likely to follow toss	0.344	0.374	0.060	6,125	0.306	0.458	0.295	3,698
	(0.011)	(0.047)	(0.125)		(0.015)	(0.060)	(0.187)	
Below average pre-toss happiness	0.222	0.670	0.418	4,357	0.166	0.938	1.009	2,604
	(0.013)	(0.067)	(0.268)		(0.018)	(0.083)	(0.464)	
Above average pre-toss happiness	0.271	0.273	−0.170	5,737	0.246	0.315	0.262	3,527
	(0.011)	(0.045)	(0.148)		(0.015)	(0.057)	(0.216)	

Notes: This table presents a sensitivity analysis for all questions. Columns 1 to 3 correspond to the two-month survey; Columns 4-6 correspond to the six-month survey. Columns 1 and 4 are first-stage estimates and describe the degree to which the coin toss affected the action taken. Columns 2 and 5 are OLS estimates, which show the extent to which those who make a change are more or less happy than those who maintain the status quo. Columns 3 and 6 are the instrumental variable estimates. The top row of this table replicates the second row of Table 5, which serves as the baseline specification against which the other results of this table can be compared. The remaining rows categorize the participants by gender, age, and the like and evaluate the robustness of the results presented in Table 5. Standard errors are reported in parentheses.

reported happiness from changes the younger subjects, as do people who reported being unlikely to follow the coin toss, and whose baseline happiness is low. On the age dimension, this pattern is interesting because older subjects are less likely to make changes than younger ones. There are few discernible patterns in the 2SLS comparisons, in large part because of imprecision. There is weak evidence of higher 2SLS of making a change for women, those with higher incomes, and those with low pre-experiment happiness.

4. POTENTIAL BIASES

There are many potential biases in the results presented above. The sources of bias fall into three broad categories: non-representativeness of the subject pool, selective response to the surveys, and untruthful answers to the survey questions. I tackle these three sets of concerns in turn, in each instance considering how the biases might affect the first-stage estimates (*i.e.* the willingness to follow the coin toss), the OLS estimates of the partial correlation between actually making a change and future happiness, and the instrumental variable estimate of the causal impact of taking an action on future happiness.[16] It is important to note, that with respect to the causal impact of the decision, many stories that might at first blush seem likely to bias the results (*e.g.* happy respondents are more likely to complete surveys, people who change are more likely to respond) in fact do not have a first-order impact on any of the estimates because there is randomization. In order for a factor to bias the 2SLS results, it must distort either the numerator or the denominator in the equation characterizing the Wald estimator above. Factors that do not differentially impact those who got heads versus tails wash out of that equation. I limit the discussion below to sources of bias which, if present, will have a first-order impact on the estimates. I focus the bias discussion on the seven-point happiness outcome that is the mainstay in the literature. The underlying logic extends to all the outcome measures.[17] Because there is no clear impact of less important decisions on happiness empirically, I focus the bias analysis on the set of important questions; it is only for these questions that the biases explored will affect the conclusions of the article.[18]

4.1. *Non-representative subject pool*

There can be no doubt that the subject pool participating in this study is highly unusual. The great majority of the recruitment for the study was done through social media associated with Freakonomics, so participants are likely to both be aware of my prior research and favourably inclined towards it. Participants tended to be young, male, and highly educated. Secondly, in the recruiting for the study, I emphasized that I was only interested in people who were having a difficult time making a life decision. This was true both in the marketing to get subjects to the website, and in the messaging once subjects arrived at the site. Consequently, individuals who are on the margin are highly over-represented, intentionally, in the subject pool. Finally, this is a group which is apparently attracted to the idea of using a coin toss to potentially resolve major life dilemmas. It is unclear whether that is a trait that is widespread in the population. Finally, because fans of Freakonomics are over-represented in this group, they might be especially likely to be responsive to my requests that they should abide by the outcome of the coin toss.

16. While OLS is of less interest than either the first stage or instrumental variable estimates, I also discuss the impact of these biases on the OLS estimates.

17. The only happiness-related outcome asked of the third parties was the standard seven-point happiness question. That is an important reason why I focus on that question in the bias analysis. I felt the third parties might not be well situated to answer the other outcomes, although in retrospect I regret the decision not to ask the other questions.

18. For each bias analyzed, a parallel table for less important questions is presented in the appendix for completeness.

All of these factors suggest that subjects in this sample are far more likely to have been influenced by the coin toss than would a randomly drawn sample, *i.e.*, the first stage is much stronger in this group than would be the case more generally.

It is less clear, however, precisely how or why this sample selection would bias the paper's estimates of causal effects of decisions. One possible channel would be that the people who participated in this study are particularly bad at making decisions on their own. So, for instance, they might tend to have difficulty making changes and wait far too long to make changes when it is obvious that a change needs to be made, and thus accrue large improvements to happiness once change occurs. However, if that were true I would have expected to see strong positive casual effects on happiness of making a change in the two-month survey, but that does not occur.

4.2. *Selective survey responses*

The results presented throughout this article are based on the subset of study participants who completed surveys. If survey respondents are not a random sample of the coin flippers, a number of different biases may be introduced, depending on the nature of the selection. The presence of the third parties identified by the subjects potentially allow me to assess both the size and direction of these possible biases.

Selective survey response can potentially affect each of the estimates presented in this article: whether people follow the coin toss, the OLS estimates of changes on happiness, and the 2SLS estimates that use the coin toss as an instrument. I deal with these three cases in turn.

4.2.1. Selective response biasing the first stage: are those who follow the coin toss more likely to report?. The measured impact of the coin toss on making a change will be exaggerated if those who follow the coin toss are more likely to respond to the survey than those who go against it. Given that the website made it clear to participants that following the coin toss was important to me, it seems plausible that those who followed the coin toss would be more likely to respond. Those who make a change might tend to fill out the survey more often if they get heads, and those who do not make a change might complete the survey with a higher probability if they get tails.[19]

To measure the actual degree of sample selection on this dimension requires some group of research subjects for whom I know the action they took, even if they do not complete the survey. The third parties are critical in this dimension. Conditional on a third party having completed a questionnaire, I am able to compare the likelihood the subject completes a survey as a function of whether or not they followed the coin toss (using as a proxy the third party's assessment of whether the coin toss was followed). Table 7 does precisely this. Entries in the first two columns of the table are the percentage of subjects who complete a survey, conditional on the third party's opinion as to whether the subject followed the coin toss (column 1) or did not follow the toss

19. The effect of this type of selection on estimates of the causal link between making a change and subsequent happiness is more subtle. As long as the coin toss has some real impact on behavior, then the 2SLS estimates will be a mixture of that causal, randomization-induced variation and variation induced by the sample selection. If, for instance, the extra individuals who are induced to respond are (as good as) randomly drawn from the underlying subject distribution, then the 2SLS will be a mix of the true causal impact and the OLS estimate of the correlation between change and future happiness. But, it is also possible that the kind of people who are very sensitive to pleasing or disappointing the experimenter are different, on average, than the other subjects. These subjects might feel guilty after making a change, and be worse off after the change than other participants, leading the 2SLS estimate to be too small. One could tell equally compelling stories as to how the bias could go the other direction as well.

TABLE 7
Are coin flippers who follow the toss more likely to report?

	Third party says coin flipper followed toss	Third party says coin flipper did not follow toss	Difference
Two-month survey	0.856	0.807	0.049
	(0.017)	(0.021)	(0.027)
	$N=443$	$N=357$	
Six-month survey	0.752	0.681	0.071
	(0.028)	(0.032)	(0.043)
	$N=234$	$N=207$	

Notes: This table explores whether the survey response rate for important questions is affected by whether the coin flipper follows the result of the flip. Questions which did not match up between the participant's and the third-party's survey were excluded. Columns 1 and 2 present coin flipper response rates according to whether the third party reported that the coin flipper did or did not follow the toss. Column 3 reports the resulting difference between the first two columns. The rows divide the results by two- and six-month survey responses. Standard errors are reported in parentheses.

(column 2).[20] The third column is the difference between the first two columns. Standard errors are in parentheses. The rows of the table correspond to the two-month and six-month surveys, respectively. Starting in the upper left corner, when the third party completes a two-month survey and says the action taken matches the coin toss, approximately 86% of the subjects also complete the survey. The second entry in the top row shows that when the third party says the subject did not follow the coin toss, reporting rates are roughly 81%, or 5 percentage points lower as shown in column 3. All the reporting rates are lower at the six-month survey, but the relative patterns are similar, with those who followed the coin toss 7 percentage points more likely to report. Thus, there does appear to biased reporting along this dimension. To the extent that third parties have imperfect knowledge of the actual actions taken by the coin flippers,[21] the numbers above actually understate the degree of selection due to attenuation bias.

A fair bit of algebra is required to ascertain the magnitude of the bias implied by the values in Table 7. Assuming the same degree of sample selection observed among this set of subjects holds across the whole population and factoring in measurement error as well, back-of-the-envelope calculations suggest that, for important decisions, about one-fifth of the estimated first-stage impact might be due to this bias on the two-month survey, and 25–30% of the six-month first-stage impact.[22]

4.2.2. Selective response biasing OLS: are happy changers especially likely to report?. It is possible that those who make a change feel particular pride if things turn out well and greater shame if the change feels like a mistake *ex post*. If that is the case, and pride leads to reporting and shame to non-reporting, then the OLS estimates of the benefit of a change will be exaggerated.[23]

20. Note that I did not ask the third parties whether the coin toss was followed, but rather, what action the subject took, which I then compare to the recommendation of the coin.

21. One way of measuring whether third parties accurately observe the actions taken is to compare responses of the coin flippers and the third parties when both complete the survey. For important questions, the two sources agree on the action taken roughly 90% of the time. For less important questions that number is roughly 83% of the time. Those numbers represent a lower bound on accurate assessment by third parties because some of the discordance may come from false reports on the part of the coin flipper.

22. See the Supplementary Appendix for the algebra underlying these calculations.

23. Although it might seem like this type of selection would be very damaging to the interpretation of the 2SLS results as well, in actuality, it is not likely to affect things much. It has no obvious impact on the first-stage estimates, because the selection is operating on the happiness dimension, not on whether a subject made a change or not. And because this type of selection affects both those who got heads and those who got tails, the overall level of reported happiness for those who flipped heads and tails—which determines the numerator of 2SLS – is not obviously biased.

TABLE 8
Are happy changers especially likely to report?

	Third party says coin flipper made a change	Third party says coin flipper did not make a change	Difference
	Two-month survey		
Third party says coin flipper is happier than average	0.894 (0.026) $N=142$	0.832 (0.028) $N=185$	0.062 (0.038)
Third party says coin flipper is less happy than average	0.819 (0.036) $N=116$	0.815 (0.021) $N=351$	0.004 (0.041)
	Six-month survey		
Third party says coin flipper is happier than average	0.823 (0.032) $N=147$	0.688 (0.044) $N=112$	0.136 (0.054)
Third party says coin flipper is less happy than average	0.689 (0.060) $N=61$	0.641 (0.045) $N=117$	0.047 (0.075)

Notes: This table explores whether the survey response rate for important questions is higher among happy coin flippers who make a change. Questions which did not match up between the participant's and the third-party's survey were excluded. The percent of coin flippers who completed a survey is presented in the cells. The first two columns divide responses according to whether the third party reported that the coin flipper made a change. The third column takes the difference between the first two columns. Rows divide the sample by whether the third party reported that the coin flipper's happiness was above- or below- average. The two panels reflect the two- and six-month survey responses, respectively. Standard errors are reported in parentheses.

Table 8 explores this possible bias. The top panel of the table corresponds to the two-month survey; the bottom panel reflects the six-month survey. In both cases, the sample is restricted to those subjects for whom a third party survey is completed. I divide the sample of subjects according to whether the third party says the subject is above or below the average level of happiness at the time of the follow-up survey. The columns of the table reflect whether the third party believes that the subject made a change. The entries in the table are the percent of subjects in that category who complete a survey. The parameter of interest is the difference-in-difference: are changers disproportionately likely to report when happy relative to non-changers. Focusing first on the top row of the top panel of the table, among subjects judged by their third party to be above average on happiness, reporting rates are six percentage points higher (89.4% versus 83.2%) when a change is made than when no change occurred. For subjects who are below average on happiness in the eyes of the third party, the gap in reporting rates is only 0.4 percentage points. This suggests that, indeed, there is potentially substantial bias at two months towards "happy changers" reporting, although the estimates are imprecise so that the t-stat on the difference is roughly equal to one. The same pattern, even stronger, appears in the bottom panel of the table which reflects the six-month survey. The difference is nearly nine percentage points (although again with a t-stat close to one because of imprecise estimates).

Back-of-the-envelope calculations imply that these differences in reporting will exaggerate the OLS estimates of making a change by roughly 10% on the two-month survey and roughly 20% on the six-month survey.

4.2.3. Selective response biasing 2SLS: are happy heads and sad tails especially likely to report?. It is not obvious why people who get heads would be disproportionately likely to report if happy, whereas those who get tails would do the opposite. If they do, however, it

TABLE 9
Are happy heads and sad tails especially likely to report?

	Heads result	Tails result	Difference
Two-month survey			
Third party says coin flipper is happier than average	0.849	0.864	-0.015
	(0.026)	(0.025)	(0.037)
	$N=185$	$N=184$	
Third party says coin flipper is less happy than average	0.794	0.823	-0.029
	(0.026)	(0.023)	(0.035)
	$N=243$	$N=283$	
Six-month survey			
Third party says coin flipper is happier than average	0.771	0.752	0.019
	(0.037)	(0.036)	(0.051)
	$N=131$	$N=149$	
Third party says coin flipper is less happy than average	0.663	0.630	0.033
	(0.048)	(0.049)	(0.068)
	$N=98$	$N=100$	

Notes: This table explores whether the survey response rate for important questions is higher among happy coin flippers who flip heads and sad coin flippers who flip tails. The percent of coin flippers who completed a survey is presented in the cells. The first two columns divide responses according to whether the coin flippers flipped heads versus tails. The third column takes the difference between the first two. Rows divide the sample by whether the third party reported that the coin flipper's happiness was above- or below- average. The two panels reflect the two- and six-month survey responses, respectively. Standard errors are reported in parentheses.

will greatly bias the 2SLS estimates. Consequently, I explore whether this bias is present in the data in Table 9.[24] This table has the same structure as Table 8. The only difference is that the columns of this table correspond to whether the subject got heads or tails. Once again, the difference-in-difference is the parameter of interest: if this bias is present, then happy heads should disproportionately report.

The numbers in Table 9 show no evidence of this form of bias. On both the two-month and six-month surveys, happy subjects are more likely to respond, but in neither case is there a notable difference between those who got heads versus those who got tails.[25]

4.3. *Untruthful answers from the subjects*

In the cases considered above, sample selection is induced by differences in survey response rates across participants, but the maintained assumption is that the research subjects truthfully answer the questions that are asked. If respondents lie, this will also affect the estimates. In what follows, I consider three different types of lies that subjects might tell in follow-up surveys: (1) claiming to have followed the coin toss when that is not true, (2) exaggerating the degree of happiness after making a change, and (3) exaggerating how happy they are if they follow the coin toss. I address these three concerns in turn.

In testing for untruthful answers on the part of subjects, my approach is always the same: I compare the answers participants give to those of the third parties, under the assumption that the third parties have no reason to lie, unlike the subjects, who may be embarrassed about their actions

24. One possible story is experimenter demand effects. If sophisticated subjects managed to (correctly) infer that the purpose of this study was to use the coin toss as a randomizing device to estimate a causal impact of making a change on future happiness, and additionally guessed (incorrectly) that I was hoping to find that change is beneficial, then in order to please me, they might have differentially reported along these dimensions.

25. Interestingly, for less important questions at six months this bias is present, as shown in Supplementary Appendix Table 18.

TABLE 10
Do coin flippers claim to have followed the toss when they have not actually done so? conditional on having a response from both the coin flipper and third party

	Flipper	Third party
Two-month survey	0.611	0.572
	(0.019)	(0.019)
	$N=661$	$N=661$
Six-month survey	0.548	0.557
	(0.028)	(0.028)
	$N=314$	$N=314$

Notes: This table explores whether coin flippers overreport having followed the toss for important questions. Questions which did not match up between the participant's and the third-party's survey were excluded. Column 1 presents the rate at which coin flippers report following the toss, while Column 2 presents the same information based on third party reports. The rows show results from the two- and six-month surveys, respectively. Standard errors are reported in parentheses.

or the consequences of their actions. Not all disagreements imply lying—third parties might not be fully informed—but to the extent that there are systematic patterns to the disagreements, this may be a sign of lying.

4.3.1. Do subjects claim to have followed the toss when they have not actually done so?. Subjects may feel pressure to say that they have followed the coin flip, especially because I so heavily emphasized the importance of doing so in advance of the coin being flipped.[26] An obvious impact of lying of this sort is that it will exaggerate the first-stage estimates. The most likely consequence for the 2SLS estimates will be to *understate* the true causal impact of a change. This is because the 2SLS estimate is the ratio of the difference in happiness of those flipping heads versus tails over the difference in the probability of making a change across heads and tails. The numerator is unaffected by this type of lying, but the denominator is exaggeratedly large, shrinking the 2SLS point estimate. OLS estimates of the value of making a change will also be biased towards zero because of attenuation bias associated with agents being misclassified.

Table 10 reports, for the set of subjects for whom I have survey responses from both the participant and the third party, the rate of coin toss following. Starting in the upper left corner of the table on the two-month survey, subjects report following the coin toss 61.1% of the time compared to 57.2% for third parties. The gap is smaller and reverses sign at six months. The data suggest some possibility that the two-month first stage may be exaggerated slightly (with the 2SLS estimates and OLS consequently understated), but do not support such a story for the six-month survey.

4.3.2. Do subjects exaggerate how happy they are when they make a change?. Although subjects do not have any particular reason to lie to the experimenter regarding how happy they are after a change, it is possible that they lie to themselves for psychological reasons. For instance, if making a change is costly (*e.g.* breaking up with a girlfriend), then it may be difficult for a person *ex post* to accept that the choice turned out poorly. A person may engage in self-deception not to have to feel the regret associated with the action. This sort of deception will have a first-order impact of exaggerating the OLS estimates of the impact of making a change. It will have no impact at all on the first-stage estimates, but will somewhat inflate the 2SLS estimates since a greater share of those who flipped heads will have made a change and exaggerated how happy they are.

26. Note, however, that both the two-month and six-month surveys emphasized that I only cared about the truth.

TABLE 11
Do participants who make a change exaggerate how happy they are?

	OLS 2M	Observations	OLS 6M	Observations
Coin flipper report of own happiness	0.828	4316	1.059	2708
	(0.068)		(0.079)	
Coin flipper report of own happiness	1.010	690	1.337	323
Conditional on having third party response	(0.172)		(0.233)	
Third party report of coin flipper happiness	1.006	690	1.407	323
Conditional on having coin flipper response	(0.180)		(0.261)	

Notes: This table explores whether coin flippers who made a change are likely to exaggerate how happy they are for important questions. Questions which did not match up between the participant's and the third-party's survey were excluded. The first row presents the coefficent on whether the individual made a change from OLS regressions with the flipper's self-reported happiness as the lefthand variable. The second row presents the same information but conditional on having a response from the third party. The third row replaces the lefthand variable with the third party's report of the flipper's happiness. Columns report OLS results by two- and six-month survey results. Standard errors are reported in parentheses.

To test for this source of bias, I estimate the basic OLS specifications of the table, but using the *third party* estimate of how happy the subject is as the dependent variable, rather than the subject's own report. The assumption underlying this approach is that third parties have no obvious reason to distort their responses.[27]

I report the results of this exercise in Table 11. For purposes of comparison, the first two rows of the table report results using the subject's own happiness report. The first row replicates the basic specifications reported in Table 3 for important questions. The second row is identical to the first row, except that it limits the sample to those subjects for whom there is also a third party survey. This second row is relevant because that same sample restriction is present in the third row, which uses third party assessments of happiness as the dependent variable. A comparison between the three rows shows that restricting the sample somewhat increases the measured impacts (*i.e.* making a change is associated with a greater increase in happiness in the subset of the population where both the subject and the third party respond), but that the results are not particularly sensitive to whether I use the subject's own happiness as the outcome or the third party's assessment. Consequently, there is little evidence that this bias is present empirically.[28]

4.4. *Summary of potential biases*

Summarizing the discussion above, it is likely that the first-stage estimates in this paper are exaggerated, both because of the selected sample participating in this study and reporting biases. There is also evidence that differential reporting may bias upward the OLS estimates of making a change on subsequent happiness by 10–20%. There is no obvious evidence for strong bias in the 2SLS, nor does it seem to be the case that lying (as opposed to differential reporting rates) is biasing the various estimates.

27. It is possible that the coin flipper misrepresents his or her happiness not just to the experimenter, but also to friends and family, in which case their assessment might also be biased. If that is the case, than using third party evaluations may not fully address the bias due to misrepresentation.

28. In principle, I can carry out the same exercise using the third party happiness reports as the dependent variable in the 2SLS estimates to test whether misreporting of happiness might bias the 2SLS estimates. In practice, however, the estimates are so imprecise that they are uninformative. The 2SLS standard errors when I restrict the sample to cases where both the subject and the third party report are roughly one in the two month survey and nearly three in the three-month survey. Thus, no reasonable hypothesis can be rejected by the data.

5. CONCLUSION

The results of this article suggest the presence of a substantial bias against making changes when it comes to important life decisions, as evidenced by that fact that those who do make a change report being no worse off after two months and much better off six months later. Stronger results, with the same implication, are found using related outcome measures, such as whether the participant is better off today than six months ago, whether he/she made the correct decision, and whether he/she would stick to that decision in a perfect foresight world. The results of this article are, of course, merely suggestive. If the results are correct, then admonitions such as "winners never quit and quitters never win," while well-meaning, may actually be extremely poor advice.

A reasonable question to ask is why so many study participants were willing to let major life decisions be dictated by a coin toss. One simple explanation is that many participants were truly on the margin. Consequently, very small benefits (*e.g.* furthering scientific knowledge, a desire to please the experimenter who made it clear that I hoped they would follow the coin toss) were sufficient to sway behaviour. Alternatively, more complex mechanisms such as regret aversion (Fehr *et al.*, 2013) may be responsible. If regret is a product of decisions that one has control over, giving up control to a randomizing device may, lessen possible regret, thus enhancing expected utility.

A large literature in psychology focuses on the "hedonic treadmill," which posits that happiness mean reverts to a relatively fixed, individual-specific set point in the long run (see, for instance, Lyubomirski, 2010). The results of my study suggest that this phenomenon does not appear to operate strongly at a six-month time horizon, at least for the sample I observe. Unfortunately, because the results and purpose of the coin flipping experiment are now public, it would be difficult to obtain reliable happiness responses from my participants in the future.

Empirical economists are increasingly moving from a role of consumers of data to producers of data. This article represents an extreme expression of that trend. It is difficult to imagine how one could hope to answer the questions addressed in this article without generating the data. As the prominence of social media grows, opportunities to recruit subject pools for randomized field experiments from broad swaths of the population will only increase.

Acknowledgments. I would like to thank Gary Becker, Stephen Dubner, Henry Farber, Lawrence Katz, Alan Krueger, John List, Susanne Neckermann, Chad Syverson, two anonymous referees, and the editor Nicola Gennaioli for valuable comments. Erin Robertson did an amazing job spearheading the project. Anya Marchenko, Ellen Murphy, and Mattie Toma provided outstanding research assistance.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ANDERSON, E. and SIMESTER, D. (2003), "Effects of $9 Price Endings on Retail Sales: Evidence from Field Experiments", *Quantitative Marketing and Economics*, **1**, 93–110.

BECKER, S. and BROWNSON, O. (1964), "What Price Ambiguity? Or the Role of Ambiguity in Decision-Making", *Journal of Political Economy*, **72**, 62–73.

BERTRAND, M. and MULLAINATHAN, S. (2001), "Do People Mean What They Say? Implications for Subjective Survey Data", *American Economic Review*, **91**, 67–72.

BOWLES, S., BOYD, R. CAMERER, C., *et al.* (2001), "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", *American Economic Review*, **91**, 73–78.

CAMERER, C. (1995), "Individual Decision Making", in Kagel, J. and Roth, A. (eds) *The Handbook of Experimental Economics* (Princeton, NJ: Princeton University Press).

CHAUDHURI, A (2011), "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature", *Experimental Economics*, **14**, 47–83.

DELLAVIGNA, S. (2009), "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, **47**, 315–372.

DI TELLA, R. and MACCULLOCH, R. (2006), "Some Uses of Happiness Data in Economics", *Journal of Economic Perspectives*, **20**, 25–46.
DOLAN, P., PEASGOOD, T. and WHITE, M. (2008), "Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-being", *Journal of Economic Psychology*, **29**, 94–122.
EASTERLIN, R. A. (1974), "Does Economic Growth Improve the Human Lot? Some Empirical Evidence", in David, P. and Rederm, M. (eds) *Nations and Households in Economic Growth* (New York and London: Academic Press).
FALK, A. (2007), "Gift Exchange in the Field", *Econometrica*, **75**, 1501–1511.
FOX, C. and TVERSKY, A. (1995), "Ambiguity Aversion and Comparative Ignorance", *The Quarterly Journal of Economics*, **110**, 585–603.
FREY, B. and STUTZER, A. (2002), "The Economics of Happiness", *World Economics*, **3**, 1–17.
GNEEZY, U., IMAS, I. and LIST, J. (2015), "Estimating Individual Ambiguity Aversion: A Simple Approach" (NBER Working Paper No. 20982).
GNEEZY, U. and LIST, J. (2006), "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments", *Econometrica*, **74**, 1365–1384.
GRUBER, J. and MULLAINATHAN, S. (2005), "Do Cigarette Taxes Make Smokers Happier", *The B.E. Journal of Economic Analysis & Policy*, **5**, 1–45.
KAHNEMAN, D., KNETCH, J. L.and THALER, R. (1991), "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias", *Journal of Economic Perspectives*, **5**, 193–206.
KAHNEMAN, D. and KRUEGER, A. (2006), "Developments in the Measurement of Subjective Well-being", *Journal of Economic Perspectives*, **20**, 3–24.
KALMIJN, M., LIEFBROER, A., SOONS, J. (2009), "The Long-Term Consequences of Relationship Formation for Subjective Well-Being", *Journal of Marriage and Family*, **71**, 1254–1270.
LEVITT, S. and DUBNER, S. (2014), *Think Like a Freak*. New York: William Morris.
LEVITT, S. and LIST, J. (2009), "Field Experiments in Economics: The Past, the Present, and the Future", *European Economic Review*, **53**, 1–18.
LIST, J. (2002), "Preference Reversals of a Different Kind: The "More Is Less" Phenomenon", *American Economic Review*, **92**, 1636–1643.
LYUBOMIRSKI, S. (2010), "Hedonic Adaptation to Positive and Negative Experiences", in Folkman, S. (ed.) *The Oxford Handbook of Stress, Health, and Coping* (Oxford: Oxford University Press).
MEIER, S. and STUTZER, A. (2007), "Is Volunteering Rewarding in Itself?", *Economica*, **75**, 39–59.
PEDERSEN, P. and SCHMIDT, T (2014), "Life Events and Subjective Well-Being: The Case of Having Children" (IZA Discussion Paper No. 8207).
SAMUELSON, W. and ZECKHAUSER, R. (1998), "Status Quo Bias in Decisionmaking", *Journal of Risk and Uncertainty*, **1**, 7–59.
SMITH, V. L. (1994), "Economics in the Laboratory", *Journal of Economic Perspectives*, **8**, 113–131.