Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Highlights:

- We test the external validity of laboratory measures of preferences for risk and guilt using COVID-19 pandemic behavior.
- We estimate risk aversion parameters for CRRA utility functions as in Holt and Laury (2002) and guilt parameters for other-regarding preferences as in Fehr and Schmidt (1999).
- Our estimates of subject-level risk-aversion are positively correlated with the frequency of handwashing and negatively correlated with the frequency of dining in at restaurants and socializing.
- Our estimates of guilt are not statistically significantly related to pandemic behavior.
- Our results provide evidence for the external validity of laboratory-measured risk preferences.
Using Pandemic Behavior to Test the External Validity of Laboratory Measurements of Risk Aversion and Guilt

Trevor Collier†  Stephen Cotten*  Justin Roush§

Abstract
We test whether laboratory measures of individual preferences for risk and guilt relate to risk-connected behaviors in a pandemic, such as socializing, dining in at restaurants, and hand washing. We utilize a survey administrated to a nationally representative subject pool in the United States in April, 2020 - the month following the declaration of a national state of emergency in response to the global outbreak of COVID-19. We find that higher levels of risk aversion are associated with risk-reducing behaviors during the COVID-19 pandemic. Meanwhile, we do not find strong evidence that guilt relates to the same behavior.

Keywords: pandemic, risk aversion, inequality aversion, guilt, experiment; external validity

JEL Codes: D9, I12, I18

Funding: This research was funded through Xavier University annual faculty development funds and a Williams College of Business Intellectual Contributions Committee grant.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Availability of data and code: Roush, Justin (2021), “Data and Code from Collier, Cotten, and Roush”, Mendeley Data, V1, doi: 10.17632/d8zw6x3vxf.3

† Professor of Economics; University of Dayton, Dayton, OH; ORCID: 0000-0001-5675-7203
* Associate Professor of Economics; University of Houston-Clear Lake, Houston, TX; ORCID: 0000-0003-0572-8167
§ Assistant Professor of Economics; Xavier University, Cincinnati, OH; ORCID: 0000-0002-8704-2849;

Corresponding Author Dr. Justin Roush
Economics
Xavier University
3800 Victory Parkway
Cincinnati, Ohio 45207
United States: roushj2@xavier.edu
1. Introduction

In December 2019, a novel coronavirus named SARS-CoV-2 (COVID-19) was discovered in Wuhan, China and began spreading globally, reaching the United States by January 21st, 2020. Nearly two months would pass before the federal government declared a State of Emergency and local governments enacted socially restrictive executive orders (e.g., shelter-in-place orders). Therefore, individuals were exposed to frequent reporting on non-negligible probabilities of long-term morbidity and mortality due to COVID-19 and were left to respond, without constraint, according to their preferences with social distancing and sanitization practices. This paper harnesses the natural experiment created by the COVID-19 pandemic to test the external validity of traditional laboratory measures of risk aversion and guilt from experimental economics.

Conceptually, risk preferences and other regarding preferences (chiefly guilt) should be important factors guiding pandemic behavior. Contracting the virus can lead to both pecuniary (e.g., medical) and non-pecuniary (e.g., pain and worry) losses. Risk aversion may guide people away from environments where the disease can be contracted or increase their engagement in risk-averting activities like sanitizing surfaces and handwashing. At the same time, guilt may guide individuals to the same avoidance behaviors. Research in infectious disease and medical outlets estimate 17.9% to 30.8% of COVID-19 cases were asymptomatic (e.g. Mizumoto et. al., 2020; Nishiura and Kobayashi, 2020). Moghadas et. al. (2020) estimate that the majority of infections may be caused by asymptomatic and pre-symptomatic individuals. Consequently, socializing or frequenting congested retail establishments creates the opportunity for all persons (both symptomatic and otherwise healthy-feeling individuals) to pass the virus to a vulnerable acquaintance. Individuals more heavily impacted by advantageous inequality (i.e., more guilty persons) may more intensely seek to avoid passing the virus.

We present the results from a survey administered to a nationally representative subject pool in the United States. Embedded in the survey are replications of cash-incentivized risk preference and inequity aversion parameter elicitation tasks often used in laboratories. From these we estimate subject-level coefficients of relative risk aversion and of the marginal disutility from advantageous inequality (the so called “guilt” parameter from Fehr and Schmidt’s (1999) model of inequality aversion). Then, using questions on subjects’ current and past behavior during the pandemic (visiting a grocery in person, using curbside grocery pickup, dining in at a restaurant, carrying out from a restaurant, handwashing, disinfecting, and socializing) we build a panel dataset which allows us to test for relationships between risk-reducing behavior, risk aversion, and guilt. We find that higher levels of risk aversion are associated with risk-reducing behaviors: increased handwashing, decreased dining in at restaurants, decreased carrying out at restaurants, and decreased socializing. We do not find strong evidence that guilt explains COVID behavior.

Our results are important on two dimensions: experimental economics methodology and policy opportunities. The methodological contribution of this paper regards the external validity of economic experiments, particularly with regards to the application of individual preference measurements. A collection of experimental economics critiques from Levitt and List (2006, 2007a, 2007b, and 2009) raised concern on whether lab results could be extrapolated to the real
world. In the context of risk and social preference measurement, two questions arise: i) do risk-averse subjects in a lab hold an aversion to risk in the real world? and ii) do guilty or envious laboratory subjects act with guilt or envy when making decisions in which social comparisons are possible?

For example, as Levitt and List (2006) suggest, the small monetary stakes typically used within the lab may frustrate the generalizability of results to real-world scenarios with higher stakes. In this study, risk preferences were elicited with choices between $5 or a gamble with two possible payoffs: $0 and $11. Similarly, social preferences were created after subjects made decisions over how to divide $10. But risk-taking during the COVID-19 pandemic involved the chance of hospitalization or even loss of life (own or others). We find, despite the potentially large discrepancy in the stakes of our laboratory tasks and COVID behavior, risk preference measurements align intuitively with pandemic behavior.

We add to an existing methodological literature with mixed results in general regarding external validity, especially in risk and social preference elicitation experiments. Some external-validity work considers the performance of the risk-elicitation mechanisms in explaining other within-lab behaviors. For example, Dasgupta et al. (2019) find that both a lottery choice task and a risky investment task perform well in explaining behavior in a competitive game involving uncertain outcomes. A recent paper by Charness et al. (2020) also finds that even simple laboratory measures of risk attitudes (as in Dohman et al., 2011) explain behavior in laboratory financial risk-taking tasks: portfolio choice, insurance, and mortgages. In sum, there is some support for lab-to-lab external validity.

Work on lab-to-field tests of external validity of risk preference measures do not robustly find a correspondence between lab and field behavior. Galizzi et al. (2016) explores the external validity of three methods of risk preference measurement on health and investment decisions. For most of the behaviors they study, no relationship is found. Charness et al. (2020) study the external application of risk measures to financial decisions (insurance, employment, and investment decisions) and find no relationship. However, several studies in health economics find a correspondence between the multiple price list mechanism of Holt and Laury (2002) and risky health behaviors. For example, Anderson and Mellor (2008) and Meredith et al. (2013) find, between them, negative relationships between risk aversion and binge drinking, smoking, obesity, and seatbelt non-use.

The literature exploring the external validity of social preference games is also inconclusive. These studies largely focus on measures of altruism, reciprocity, and trust through the dictator, ultimatum, trust, and public goods games. They explore the explanatory power of these measures in pro-social behaviors in the field, such as giving money, volunteering, and helping others. For example, Galizzi and Navarro-Martines (2019) conduct an experiment and perform a meta-analysis of the literature. Their experiment finds little evidence that behavior in social preference laboratory experiments predicts behavior in field settings. In their review of the literature they find that less than 40-percent of the studies in this area report statistically significant

---

1 See Kessler and Vesterlund (2015) for an excellent summary of external validity concerns.
relationships between results in the laboratory and results in the field. However, Franzen and Pointner (2013) find that subjects who displayed other-regarding behavior in a laboratory experiment were more likely to return money from a misdirected letter mailed weeks after the experiment. Related to the COVID-19 pandemic, Campos-Mercade et al. (2020) follow subjects from a pre-COVID pro-sociality experiment and find that pro-social behavior also drives social distancing behavior, mask-wearing, and shelter-in-place behavior.

Our results have implications for public policy as well, such as public health policy during COVID-19. Researchers have examined the relationship between shelter-in-place orders, social distancing, and trust. Dave et al. (2020) show that shelter-in-place orders increased the number of residents in a state who remained in their home full-time by almost 10%. Similarly, Courtemanche et al. (2020) find that shelter-in-place orders were effective in reducing the spread of COVID-19.

While overall shelter-in-place policies were effective, policymakers possessed other options to reduce the spread of COVID-19: chiefly, information dissemination. But public reaction to information depends upon underlying preferences. For example, Brodeur et al. (2020) and Bargain and Aminjonov (2020) discover that shelter-in-place orders reduce mobility and are more effective in high-trust areas. Maloney and Taskin (2020) suggest that fear or a sense of social responsibility are the primary drivers of social distancing behavior. Our study suggests risk aversion is correlated with reduced socialization and restaurant patronization and increased handwashing. As such, information campaigns which increases the saliency of risk among the public may have heterogeneous effects on behavior. Given the potential association between demographics and risk preferences, policymakers may find greater reductions in virus contraction in areas of where population risk aversion is greatest, ceteris paribus. For example, risk aversion has been found to be greater in relatively younger (e.g., Harrison and Ruststrom (2008)), more educated (e.g., Anderson et al (2010)) and black (Benjamin et al (2010)) individuals.

The paper proceeds as follows: section 2 describes our survey design, section 3 provides an overview of the data generated, section 4 presents an analysis of the results, and section 5 some additional discussion.

2. Survey Design

We use a within-subjects design with three different incentivized exercises to measure the risk preferences and social preferences of our subjects. Risk preferences are measured with a slight modification to standard paired lottery decisions. Social preferences are measured in two ways. First, subjects complete the ultimatum game, wherein they make choices as both the proposer and receiver. Then, subjects complete a modified version of the dictator game. Each of these instruments are presented separately and subjects are paid based on only one of the instruments, chosen randomly with equal probability at the conclusion of the experiment (the so-
called strategy method). Subsequent to each instrument, subjects are asked to indicate how well they understood the instrument prompt. After completing these instruments, subjects are asked about their behavior during the first few weeks of the COVID-19 pandemic in the United States. Lastly, subjects report a few key demographic characteristics. All subjects completed the survey over a 24-hour period on April 22 and April 23, 2020.

Importantly, the subjects are drawn from a pool of consumers in the United States by Prodege, LLC. Prodege is a marketing research company that generally helps other companies gather insights about their products or their consumers by maintaining a large pool of members. Prodege has over 100 million unique members who are accustomed to completing surveys, shopping online, playing games, and completing other tasks that reward them with SwagBucks (SB). SB can then be converted into gift cards to numerous large retail stores or exchanged for money through PayPal.

2.1 Risk Preference Measurement

Laboratory risk preference measures range from simple non-incentivized survey questions about subjects’ willingness to take risks in general (Dohmen et al., 2011) to complex mechanisms that allow researchers to estimate the utility-curvature and probability weighting functions of prospect theory (Tanaka et al., 2010). In between, there are several incentivized methods involving tradeoffs between safer and risky uncertain outcomes (as in Gneezy and Potters (1997), Holt and Laury (2002), Eckel and Grossman (2008), and Charness and Gneezy (2010)) by which experimenters vary the relative riskiness of the two decisions and record when subjects “accept” the risk. These are anticipated to reveal relative risk preferences among subjects that can be used to predict their behavior in other settings. Some literature compares subject performance on various risk-elicitation tasks and finds that they yield different results (e.g. Charness et al., 2013; Crosetto and Filippin, 2016) though there is still debate over whether these differences are evidence of inconsistency in the measurements (Holzmeister, 2020).

The first exercise in our survey involves making paired lottery choice decisions, similar to Holt and Laury (2002) (hereafter HL), but with a modification for simplicity that participants choose between: (i) receiving a guaranteed five dollars or (ii) a lottery that might pay $11 or might pay $0. Instead of making multiple decisions via a price list, participants are asked to use a slider-bar to choose the lowest probability of winning $11 that would make them prefer the lottery over the guaranteed $5. Participants are informed that if this exercise is chosen for payment, the computer will choose a probability of winning $11 for the lottery. If the computer’s choice is a lower probability than their slider bar, they do not play the lottery and get $5; if it is at or above their slider bar, they take the lottery (then the lottery is played out with the probability

---

3 The strategy method is designed to collect multiple observations from a single subject while maintaining incentives in each decision. Since each decision is equally likely to be used for payment, subjects should treat each decision as if it is in fact the only one being paid.

4 Instructions, preference-elicitation instruments, and survey questions can be found in Appendix B.

5 Payments in our experiment were described in dollars. Subjects were informed that dollars earned in the game would be paid in SB equivalent. For example, a subject earning $20 in our experiment would be paid 2000 SB by Prodege, which they could then redeem for $20 gift cards or for $20 cash through PayPal.
of winning $11 equal to the number chosen by the computer). In this way, our task also shares similarities with the Becker-DeGroot-Marschak (1964) procedure, as modified by Harrison (1986, 1990) and Loomes (1988), in which subjects identify how much they are willing to accept to give up a gamble assigned to them (or willing to pay to buy the gamble).

2.2 Other Regarding Preferences Measurement

The second exercise is a modification of the ultimatum game presented in Blanco et al (2011) (hereafter, BEN), which is similar to the “easy game” in Guth et al. (1982). We do not use the data from this exercise in this project but provide the description for a complete picture of what subjects did. Subjects participate in two, sequential stages in this game. In the first stage, Player A proposes how to share $10 between herself and Player B. In the second stage, Player B chooses whether to accept Player A’s proposal. If Player B accepts the proposal, they are each paid the amount proposed by Player A. If Player B rejects the proposal, then both players receive no money. We use the strategy-elicitation method, like BEN, where every participant is asked to make decisions as Player A and as Player B. Subjects are told they will be randomly paired with another person participating in the study. For the Player A role, the participant must choose how much to keep for herself and how much to leave for Player B. These two amounts must sum to ten before the participant can continue. For the Player B role, they are asked to report the smallest amount they are willing to accept from Player A. If they are paired with a Player A who offers a larger amount, then Player B accepts the offer. If Player B is paired with a Player A who offers a smaller amount, then Player B rejects the offer.

The third (and final incentivized) exercise is a modified dictator game. In our modified dictator game, which follows BEN, the dictator is offered a series of choices between two payoff vectors. The payoff vectors reflect payments that will be made to the dictator and another, randomly chosen participant in the study. One of the choices is always ($10,0), meaning the dictator gets $10 and the other participant receives nothing. The alternative payoff vector is always an even amount of money for the dictator and the other participant. This amount varies from ($0, $0), ($0.50, $0.50), ($1, $1) all the way up to ($10, $10). All participants in our study make this decision, but like the ultimatum setting, they are randomly assigned roles ex post. The unequal payoff vector ($10, $0) is always listed as the left-hand choice and the equal payoff vector (e.g. $1,$1) is always listed as the right-hand choice. Participants are asked to choose one pair (if any) at which they want to switch from choosing the unequal payoff vector (left) to choosing the equal payoff vector (right). If this exercise is chosen for payment, one of the 21 pairs of payoff vectors is chosen at random and the decision of the participant selected to be the dictator determines the payoff.

2.3 Baseline and COVID Behavior Survey

The fourth exercise asks participants to report the frequency with which they engaged in the following activities: physically entered a grocery store, utilized curbside pickup from a grocery store, dined in a restaurant, picked up carry-out or had food delivered from a restaurant, used
household disinfectants to clean surfaces at home, socialized in person with friends/family not living in their home, washed hands or used hand sanitizer. They are asked to report these frequencies for a normal two-week period (or a normal day for washing hands and sanitizing surfaces) in 2019. Next, all participants are asked to report the frequency of these same behaviors during each of three separate two-week periods, starting at the beginning of March 2020, relative to their behavior in 2019. Recall, all subjects completed the survey over a 24-hour period on April 22 and April 23, 2020.

*Figure 1: Sample Pandemic Behavior Survey Questions*

**Panel A**

| Question | Not at all | Much less | Moderately less | Slightly less | About the same | Slightly more | Moderately more | Much more |
|----------|------------|-----------|-----------------|---------------|---------------|---------------|----------------|-----------|
| FIRST two weeks of March 2020 |           |           |                 |               |               |               |               |           |
| LAST two weeks of March 2020 |           |           |                 |               |               |               |               |           |
| FIRST two weeks of April 2020 |           |           |                 |               |               |               |               |           |

*Note: This figure depicts samples of our pandemic behavior survey questions. See Appendix B for the full survey.*
For example, Figure 1 shows the questions pertaining to entering grocery stores. Shown in Panel A, participants were presented with the following question: How many times did you physically enter a grocery store in a normal two-week period in 2019? They indicated by radio button the frequency of their 2019 behavior and repeated this for the remaining six behavior categories. Next, as shown in Panel B, subjects were asked to compare the change in frequency in which they entered a grocery store in the present year (2020) for the first two weeks of March, the last two weeks of March, and the first two weeks of April, relative to the frequency with which they entered a grocery store in a normal two-week period in 2019. This was again repeated for the remaining six behaviors.

Prior research has shown that subjects are better at recollecting relative changes in behavior than recollecting absolute changes (Ariely et al., 2003). Therefore, we captured relative frequencies in 2020 instead of repeating the 0-10 or more scale from our 2019 baselines. The 2019 values were collected in part to prime subjects to reflect upon their behavior before the pandemic for better comparison with their behaviors in 2020. Capturing 2019 frequencies also allow us to control for baseline behavior across subjects. For example, subjects who largely used grocery pickup in 2019 cannot greatly increase it in 2020, so controlling for their baseline behavior allows us to better compare across subjects.

The fifth exercise provides participants with a statement from the CDC then asks them to report on the probabilities of certain events occurring that would make them feel “comfortable returning to normal daily life.” We do not use the data from the fifth exercise in this study. Lastly, the sixth exercise asks participants to answer several questions about their demographics, background, political views, and other personal preferences. The full instructions for this study are available in Appendix B.

3. Data

Recall that after completing incentivized instruments to measure subjects’ aversion to risk and aversion to advantageous inequality (guilt), we asked all subjects to report how their behavior has changed from their typical behavior in the previous year across several dimensions: entering a grocery store (Grocery), using curbside grocery pickup (GroceryPickup), dining inside a restaurant (DineIn), carrying out food (Carryout), general socializing in person with people outside their household (Socialize), handwashing or using hand sanitizer (Handwash), and disinfecting surfaces in their home (Disinfect). We captured this information across three periods during the early stages of the COVID-19 pandemic: first two weeks in March, last two weeks in March, and first two weeks in April. The survey was conducted on April 22 and 23, proximate to the three periods of study.

These periods were chosen intentionally. Knowledge and information about the disease grew significantly over the month of February. The World Health Organization (WHO) named the disease (COVID-19) on February 11. An outbreak of the virus emerged in Italy in late February, with officials closing 10 towns near Milan on February 23. The first death caused by COVID-19

\footnote{Note that questions regarding handwashing and disinfecting requested frequencies on the average day, not average week.}
in the United States was reported on February 29. And finally, a national state of emergency in the U.S. was declared on March 13th. Our three periods begin immediately following the first reported death, span the declaration of a national state of emergency, and include days after most local shelter-in-place orders were implemented in the United States. As such, it is possible that behavior after the first two weeks in March converges in response to federal and local laws.

Further, research has already found that shelter-in-place orders affect risk-taking behavior (Courtemanche et al., 2020). We recorded ZIP codes to match subjects to state and local behavioral restrictions, as prepared by researchers at Johns Hopkins University (Killeen et al., 2020). This allows us to observe when individuals faced a shelter-in-place order relative to the three periods in our study. Prior to March 19th, no state had issued a shelter-in-place order. Some cities and counties (and Puerto Rico), however, had their own orders. Figure A1 (in the appendix) summarizes the variation in shelter-in-place orders in our sample. It depicts the counties represented in our sample as well as the number of days a subject from that county faced a shelter-in-place order during our period of study. In sum, period dummies and shelter-in-place information allows us to fully control for when subject behavior became exogenously limited by law.

We also collected demographics which could correlate both with pandemic behavior but also risk-aversion and guilt. For example, some experimental and observational evidence suggest that risk preferences and inequality aversion depend upon sex, race, and other demographics (see, e.g., Eckel and Grossman, 2008; Benjamin et al., 2010). Insofar as social norms create differences by sex or race in grocery store visits, socializing, or cleaning and disinfecting, for example, our estimates may be biased without their inclusion. We also collect variables on pandemic employment status: whether the subject is an essential worker and whether they have been unemployed/lost hours because of COVID-19.

Our full sample consists of 184 subjects from 36 states. The survey took an average of 20.5 minutes to complete. We reduced the sample to account for disingenuous study participants, which is commonly required when conducting online surveys (see, e.g., Horton et al (2011) who recognize a need for data filtering with online experiments and argue for transparency of the process, which we provide here). Nine subjects rushed through the survey in less than eight minutes, which is the fastest 5% of observations. We further filtered five “straight-liners”, or subjects that clicked through the survey using the same response (e.g., selected 10 for every frequency category for normal 2019 behavior). Six more observations drop out of both the full and filtered sample when demographic controls are used, as a handful of subjects did not complete demographic survey questions. We refer to this sample as our “filtered” sample. The statistical summaries, analysis, and our results that follow are based on this sample of 164 subjects. We also briefly describe a set of results from a subsample built by additionally filtering...

---

7 We followed Courtemanche et al. (2020) in making minor modifications to the SIPO dates.
8 The United States never issued a national shelter-in-place order.
9 For ZIP codes that overlapped county-lines, counties were designated based on majority ZIP residential population proportions. The U.S. Department of Housing and Urban Development's (HUD's) Office of Policy Development and Research (PD&R) provides ZIP-to-County crosswalk data.
subjects who indicated they did not understand the HL or BEN tasks (146 subjects). A participants understanding of tasks is subjective and susceptible to noise as those who were confident in their responses may have misunderstood the prompt. For that reason, we focus our discussion on the more conservative sample of 164 subjects. Lastly, for transparency, we also present results on the full, unfiltered sample.

3.1 COVID Behavior
Each panel in Figure 2 charts proportions of responses across periods for one of seven categories: Socialize, Grocery, GroceryPickup, DineIn, CarryOut, Handwash, and Disinfect. Grey indicates individuals held their behavior constant compared to a normal period in 2019 (see the legend). Relatively darker bars represent reductions in behavior relative to 2019 while relatively lighter bars represent increases in behavior relative to 2019.10 We find variation in behaviors across subjects within each period. Most variation occurs in Period 1, in which subjects were aware of the pandemic, but a Federal State of Emergency had not been declared and states had not implemented shelter-in-place orders. In other words, Period 1 represents a time in which subjects were not strongly guided by policy and instead made decisions based on their own preferences. However, behavior does not completely converge after Period 1. Our study seeks to examine whether the differences in behavior in each period can be explained by laboratory estimates of risk and guilt preferences.

---

10 Some subjects who did not do an activity in 2019 and also did not do the activity during pandemic chose “Not at all”. We recoded these subjects’ responses to “About the Same” in the pandemic since their behavior, in fact, did not change.
Figure 2: Response Proportions by Category across Periods

Note: This figure shows the distribution behavior frequencies across behavior types and periods for the filtered sample. Respondents reported how their behavior changed from their typical behavior in the previous year across several dimensions: entering a grocery store (Grocery), using curbside grocery pickup (GroceryPickup), dining inside a restaurant (DineIn), carrying out food (Carryout), general socializing in person with people outside their household (Socialize), handwashing or using hand sanitizer (Handwash), and disinfecting surfaces in their home (Disinfect). Period 1 refers to the first two weeks in March 2020, Period 2 includes the last two weeks in March 2020, and Period 3 is the first two weeks in April 2020. Grey indicates individuals held their behavior constant compared to a normal period in 2019 (see the legend). Relatively darker bars represent reductions in behavior relative to 2019 while relatively lighter bars represent increases in behavior relative to 2019. Shelter-in-place orders began in some areas in Period 2. By Period 3, nearly all subjects face shelter-in-place orders.

3.2 Demographics and Pre-COVID Behavior
Table 1 summarizes the demographics and control variables for our sample. The sample is approximately half female. We restricted our sampling to those older than 23, so the average age in our sample is 43.8 with a minimum of 23 and a max of 81. Approximately 60% of the sample is married, and 55% have children. There is a wide range of educational backgrounds and political leanings. About 10% had unemployment attributed to COVID-19 and 20% remained employed but lost income or hours due to COVID-19. Of those still employed, 18.3% were essential workers. About 72% of our subjects were working prior to COVID-19, which is expected with a February 2020 labor force participation rate of 63.4% and a 25-54 subset participation rate of 82.9%. These numbers correspond closely to a May 2020 survey (Kaiser Family Foundation, 2020) on employment impacts of COVID.
Reported values in 2019 should be viewed as a noisy control for baseline behavior as we request year-old memories of socially desirable behavior. We likely have both a form of recall bias (projecting present behavior onto past behavior), social desirability bias (not wanting to

| Table 1. Summary of Control Variables |
|--------------------------------------|
| **Basic Demographics**               |
| Male                                 |
| 49.2% (50.0%)                       |
| Age                                  |
| 42.8 (13.2)                          |
| Married                               |
| 57.1% (50.0%)                       |
| Have Children                        |
| 52.7% (50.0%)                       |
| **Household Income**                 |
| Under $20,000                        |
| 7.8% (26.9%)                         |
| $20,000 to $39,999                  |
| 17.4% (39.0%)                       |
| $40,000 to $59,999                  |
| 18.0% (38.4%)                       |
| $60,000 to $79,999                  |
| 13.5% (34.2%)                       |
| $80,000 to $99,999                  |
| 16.2% (37.0%)                       |
| Above $100,000                      |
| 27.0% (44.4%)                       |
| **COVID-19 Job Impact**              |
| Essential Worker                     |
| 19.1% (39.3%)                       |
| Job Reduced                          |
| 20.2% (40.2%)                       |
| Job Lost                             |
| 10.7% (30.9%)                       |
| Not Working Prior                    |
| 27.5% (44.7%)                       |
| **Education**                        |
| High School or Less                  |
| 12.9% (33.6%)                       |
| Some College                         |
| 25.2% (43.5%)                       |
| Four Year Degree                     |
| 37.1% (48.3%)                       |
| Masters                              |
| 18.0% (38.4%)                       |
| Professional/Doctorial               |
| 6.7% (25.1%)                        |
| **Political Views**                  |
| Left                                 |
| 24.1% (43.2%)                       |
| Moderate Left                        |
| 16.9% (37.4%)                       |
| Central                              |
| 27.0% (44.4%)                       |
| Moderate Right                       |
| 7.3% (26.0%)                        |
| Right                                |
| 24.7% (43.1%)                       |
| **Behavior in 2019**                 |
| Grocery                              |
| 3.949 (2.507)                       |
| Grocery Pickup                       |
| 0.972 (2.316)                       |
| Dine In Restaurant                   |
| 2.725 (2.460)                       |
| Carryout                             |
| 2.438 (2.513)                       |
| Disinfect                            |
| 3.961 (3.323)                       |
| Handwash                             |
| 5.742 (2.967)                       |
| Socialize                            |
| 3.511 (2.863)                       |
| Observations                         | 183 | 164 |

Note: Our filtered data set results from dropping nine subjects who rushed through the survey in less than eight minutes, five "straight-liners", or subjects that clicked through the survey using the same response, and six subjects who did not complete demographic survey questions.

Reported values in 2019 should be viewed as a noisy control for baseline behavior as we request year-old memories of socially desirable behavior. We likely have both a form of recall bias (projecting present behavior onto past behavior), social desirability bias (not wanting to
reveal you eat out “too much” or that you have lacking hygienic practices), or affective motivated beliefs (actually believing you had better hygienic practices than you actually did – Benabou and Tirole (2006)). For example, in a normal day in 2019 subjects reported washing their hands 5.7 times and disinfecting surfaces 3.9 times on average. Furthermore, subjects visited the grocery almost four times, dined in more than twice, carried out food from a restaurant twice, and socialized just over three times in a normal two-week period in 2019. We note that our results are robust with the omission of 2019 baseline controls, which is somewhat anticipated given the Ariely et al. (2003) result suggesting subjects do poorly when tasked with recalling absolute changes in behavior but do well at recollecting relative changes in behavior (our dependent variable is denominated in relative changes).

3.3 Risk Preferences
We used participants’ responses to our modified HL risk instrument to calculate their coefficient of relative risk aversion (CRRA). Our strategy is to study whether variation across the distribution of risk preferences maps to variation in observed risky behaviors during COVID. Table 2 summarizes our computed CRRA in comparison to those measured in HL and Figure 3 depicts the entire distribution of our estimation sample. The central tendency of our data falls close to HL. Recall the slider-bar allows subjects to choose from 1% to 100% probabilities in increments of 1 percentage point. The mean point of indifference for our participants was a 53.3% chance (54.0% full) of earning $11 versus a guaranteed $5, equating to a CRRA of 0.1948. The modal choice was 50% which equates to a CRRA of 0.1209 (25 out of 164 in filtered, and 26 out of 184 in full).

Table 2. Comparison of risk aversion with Holt and Laury (2002)

| CRRA  | HL Classification  | Filtered Proportion | Full Proportion | HL Proportion |
|-------|---------------------|---------------------|-----------------|--------------|
| r < -0.49 | Very risk loving | 44 (26.8%) | 50 (27.2%) | 7% |
| -0.49 < r < -0.15 | Risk loving | 14 (8.5%) | 16 | 8% |
| -0.15 < r < 0.15 | Risk neutral | 30 (18.3%) | 31 (16.8%) | 29% |
| 0.15 < r < 0.68 | Risk averse | 37 (22.6%) | 40 (21.7%) | 41% |
| 0.68 > r | Very risk averse | 39 (23.7%) | 47 (25.5%) | 15% |

11 We assume utility is defined by the following constant relative-risk aversion specification, \( u(x) = \frac{x^{1-r}}{1-r} \), where \( x \) is the payoff received. The minimum probability (\( p \)) of winning the larger prize ($11) that the subject selects should be their indifference point between that and the $5 certainty equivalent. Equalizing their utility in each case, it must be true that \( \frac{x^{1-r}}{1-r} = p \frac{x^{11-r}}{1-r} \) so solving for \( r \) gives us \( r = \frac{\ln(11p) - \ln(5)}{\ln(11) - \ln(5)} \). We denote \( r \) as CRRA in text.
Our sample was much more likely to test in the extremes than the subjects in HL and other studies who use similar risk elicitation tasks. This may be due to task confusion or differences between subject samples in our study. Task confusion is anticipated given our study is not conducted in a lab but via online survey. It may be plausible that subjects believed choosing a probability of 1 indicated a low chance at receiving the gamble, when in fact it was the opposite. Or, perhaps by indicating a choice of 100 they believed they were increasing their chance at receiving $11. Confusion has been established in the literature in, for example, Cason and Plott (2014) and Burfurd and Wilkening (2018).

Our sample composition differs greatly from typical studies of risk preferences. For example, subjects in HL were recruited from universities and were mostly undergraduates and MBA students. Kachelmeier and Shehata (1992), who use a BDM certainty equivalent
elicitation, studied university students from China and also did not produce the extremes observed here. Our sample is collected to be nationally representative and, as such, demonstrates significant heterogeneity in demographic background. Our results are most related to Anderson et al (2010) who utilize a multiple price list in a CRRA comparison study between a representative Danish sample and a laboratory sample with students at the University of Copenhagen. Like our study, their nationally representative sample distribution of CRRA measures slightly more risk seeking on average and has longer tails into the extremes.

The greater background heterogeneity of our nationally representative sample is likely important. For example, variation in numeracy across subjects may produce greater variation in our distribution of CRRA estimates: while college-aged samples are usually trained in math, the general population may possess lower numeracy leading to more task confusion as discussed prior. Additionally, variation in perception of the stakes may produce extreme risk preferences; Kachelmeier and Shehata (1992), Fehr-Duda et al (2010), and others show risk-aversion is increasing in the size of the stake. Low-stakes gambles can produce high risk-lovingness. Lastly of note is that our sample is on average older than most studies since we collected surveys for subjects 24 and older. The literature on risk preferences supports a result that older subjects are on average more risk seeking (see, e.g., Harrison and Rutstrom (2008)). By comparison, the representative sample of Danish subjects in Anderson et al (2010) start at age 19.

There may be other explanations for the differences between our distribution of CRRA and other studies, but altogether the presence of extreme risk preferences among a few subjects is only concerning if the causes of their abnormal CRRAs also systematically correlate with COVID behavior. For example, if confusion error on the task is correlated with high or low levels of socialization. It is not theoretically obvious that this is the case. As such, the extreme metrics add noise to our estimates. We build evidence for this in our analysis by presenting in Appendix Table A5 estimation results with an 80% Winsorization of the CRRA metric. This resets the very risk love subjects to have CRRAs at the 10th percentile (-1.04). Our results are robust to this treatment of the data suggesting that outliers are not accounting for the relationships between risk preferences and pandemic behavior that we discover.

3.4 Guilt
We use a modification of the dictator game to capture $\beta$, the aversion to advantageous inequality, which is also known as a guilt parameter in Fehr and Schmidt (1999). We follow BEN and use a Fehr and Schmidt (1999) utility function to compute $\beta$. For two player games, utility is given by

$$U_i(x_i, x_j) = \begin{cases} x_i - a_i(x_j - x_i), & \text{if } x_i \leq x_j \\ x_i - \beta_i(x_j - x_i), & \text{if } x_i > x_j \end{cases}$$

The monetary payoffs to players $i$ and $j$ are $x_i$ and $x_j$, respectively.

We obtain the switching points at which subjects move from preferring to keep the entire endowment ($10,$0) to preferring the egalitarian outcome ($x_i$, $x_j$). Because the subject indicates a preference for ($x_i$, $x_j$), that implies indifference occurs at some point before the
previous step $0.50 lower. We estimate indifference occurs at the midpoint, so we calculate $\beta \text{ (guilt)}$ as

$$\beta_i = 1 - \frac{x_i - 0.25}{10}$$

Subjects who choose ($0,0$) over ($10,0$) may actually be willing to pay to reduce inequality and $\beta_i > 1$. However, we do not observe a switching point for these subjects and, like BEN, assign them $\beta_i = 1$. Likewise subjects who prefer ($10, 0$) to ($10, 10$) may be willing to give up money to increase inequality and have $\beta_i < 0$, but again we do not have a switching point and assign them $\beta_i = 0$.

In our modified dictator game, the average switching point amongst subjects who switched at all from ($10, 0$) to an equal split was approximately $3.82 (3.61 \text{ full})$, or approximately $38\%$ of the “selfish” amount. This is fairly close to the $45\%$ observed in BEN. $4.3\%$ of our sample (7 of 164 in filtered, 10 of 184 in full) chose ($10,0$) over ($10,10$) and a further $8.5\%$ (14 of 164 in filtered, 14 of 184 in full) only switched to the egalitarian option at ($10,10$) when it was costless. 17 of 164 or $10.4\%$ in filtered (23 of 184 in full) chose ($0,0$) over ($10,50$), which is substantially higher than the $2$ in $61 \text{ (3\%)}$ in BEN. $53\%$ of subjects ($54\%$ in full) switched to the egalitarian outcome in the range of ($0,0$) to ($4.50, 4.50$), compared to $43\%$ in BEN. In both of our samples and in BEN, the modal switching point was $50\%$.

**Table 3. Comparison of Aversion to Advantageous Inequality (Guilt) with BEN**

| $\beta$ | Fehr-Schmidt Proportion | BEN Proportion | Filtered Proportion | Full Proportion |
|---------|-------------------------|----------------|---------------------|----------------|
| $< 0.235$ | 30% | 29% | 16.5% | 16.3% |
| $0.235 \leq \beta < 0.5$ | 30% | 15% | 6.7% | 6.5% |
| $\geq 0.5$ | 40% | 56% | 76.8% | 77.2% |

Note: This table reproduces the $\beta$ categorizations from BEN with their results, our full-sample results, and our filtered results. Our subjects were more likely to possess $\beta \geq 0.5$.

We compare our results with that of BEN in Table 3 and Figure 4 to gauge the quality of our guilt estimates. Overall, the subjects in our experiment show higher measures of guilt: $76.8\%$ of our subjects possess $\beta \geq 0.5$ as compared to $40\%$ in BEN. As shown in Figure 4, this comes mostly by way of a larger proportion of $\beta_i = 1$. Otherwise, the kernel density estimates for our sample in the top panel of Figure 4 and BEN’s sample in bottom panel of Figure 4 are similar. We elicited this data during a pandemic in which people were asked (even forced through shelter-in-place orders) to behave pro-socially and were sanctioned, informally or formally, for failure to comply. The pandemic also aggravated underlying inequalities, and this might give ground for stronger-than-normal feelings of guilt and larger-than-normal giving. Although subjects are on average guiltier, there is still variation across the $76.8\%$ of subjects within those
measuring $\beta \geq 0.5$ which we exploit to identify the relationship between relatively higher guilt and pandemic behavior.

**Figure 4: Beta Distributions from Estimation Sample and BEN**

![Histogram and Kernel Density Plot for Beta](image1)

![Histogram and Kernel Density Plot for Beta from Blanco et al (2011)](image2)

*Note: This figure shows the distribution of $\beta$ coefficients estimated from a modified dictator game to measure the aversion to advantageous inequality (guilt). The top panel displays coefficients from our filtered sample versus those from BEN in the bottom panel. In the absence of a relatively larger group of $\beta = 1$, the kernel density estimates of the $\beta$ distributions are similar across studies.*

4. **Analysis and Results**

In this section we analyze our data to inform two broad hypotheses:

**Hypothesis 1:** Risk averse individuals (higher CRRA) more heavily engage in risk avoidance behaviors in a pandemic than less risk averse individuals, ceteris paribus.

In theory, more risk averse individuals should engage in behaviors that reduce the risk of getting infected with COVID-19. We expected CRRA to be negatively related to DineIn and Socialize since social distancing reduces the chance of contracting the virus. For the same reason we expected positive relationships for Handwashing and Disinfecting because they too reduce the odds of contracting the virus. GroceryPickup may have a positive relationship to risk-
aversion as it provides a lower-risk means to obtain food than grocery or dining in. However, with three quarters of our sample never having used grocery pickup services in 2019, and roughly three quarters also reporting a 2020 use-frequency of “Not at all” (see Figure 2), there is likely not enough variation in GroceryPickup to identify a relationship between risk-aversion and guilt in our study.

Our intuition for CRRA’s relationship with (in-person) Grocery and (restaurant) Carryout is less clear. Given that physical distancing is limited in restaurants and grocers, the question remains as to where subjects will obtain food should they seek to avoid the risk of these establishments. When available, we expect more risk-averse people to switch more from Grocery to GroceryPickup or Carryout, but at the same time Grocery could be getting an inflow of people switching from the even more risky activity of dining in restaurants. A similar story exists with carryout orders from restaurants, as individuals could be substituting from that to dining at home (Grocery) even as others are substituting into carrying out from dining in. These competing substitutions between “food sources” could result in failure to reject the null.

Table 4 reports odds ratios for the relationships between CRRA and behavior estimated with a random effects ordered logit estimator. Since we created rules for eliminating some subjects from analysis, we report results from four different filtering schemes: none (None); only straight-liners (SLine); straight-liners and speeders (SLineXSpeed); and finally straight-liners, speeders, and those who indicated they did not understand the lottery or modified dictator game (SLSXSpeedXNoUnderstand). Another concern with the use of the ordered logit estimator is that it may produce biased slope estimates when too many response categories are used or, more chiefly, there are small counts in a few categories. Murad et al (2003) note that the slope parameters are invariant to collapsing adjacent categories and doing so can improve ordered logit asymptotics. Figure 2 reveals rare responses for decreases in handwashing and disinfecting, as well as increases in socializing, dining in, entering a grocery or restaurant, and using carryout. We collapse these rare categories into larger “Decrease” or “Increase” categories, respectively, and present their results Table A4 of Appendix A. We note that the results we discuss next are mostly robust to samples and collapsing schemes.

Consistent with earlier sections of the paper, we discuss results from our full sample (None) and from a sample that filters out only straight-liners and speeders (SLinexSpeed). We focus on results significant at the 5% confidence level. For dining in and socializing, the odds of reporting a marginally higher frequency is 0.54 and 0.59 times lower, respectively, for a one-unit increase in CRRA. Alternatively, we can say that a one-unit increase in CRRA increases the odds of...
reporting a marginally lower frequency category by 1.8 times for DineIn and 1.7 times for Socialize. For Handwash, a one unit increase in CRRA increases the odds of reporting a marginally higher frequency by 1.8 times. The odds ratios for CRRA in the models for Grocery, GroceryPickup, Carryout, and Disinfecting are not statistically different from zero at the 5% level of significance.

Table 4: Ordered Logit Output (Odds Ratios)

| Filter          | Covariate | Grocery | Grocery Pickup | Dine In | Carryout | Disinfect | Handwash | Socialize |
|-----------------|-----------|---------|----------------|---------|----------|-----------|----------|-----------|
| None            | CRRRA     | 1.047   | 0.997          | 0.656   | 0.736    | 1.440     | 1.650**  | 0.680*    |
| None            | Beta      | [0.216] | [0.275]        | [0.212] | [0.151]  | [0.378]   | [0.380]  | [0.152]   |
| None            | Beta      | [1.117] | 0.423          | 1.006   | 0.334    | 1.753     | 0.696    | 0.780     |
| SLine           | Beta      | [1.114] | [0.445]        | [1.627] | [0.292]  | [1.839]   | [0.643]  | [0.896]   |
| SLine           | CRRRA     | 0.977   | 0.931          | 0.542** | 0.693**  | 1.508     | 1.757**  | 0.576***  |
| SLine           | Beta      | [0.198] | [0.253]        | [0.157] | [0.138]  | [0.405]   | [0.402]  | [0.115]   |
| SLine           | Beta      | [1.315] | 0.521          | 1.500   | 0.324    | 2.183     | 0.971    | 1.083     |
| SLine           | Beta      | [1.302] | [0.554]        | [2.145] | [0.307]  | [2.269]   | [0.903]  | [1.115]   |
| SLineXSpeed     | CRRRA     | 1.005   | 0.918          | 0.541** | 0.720*   | 1.550     | 1.778**  | 0.594***  |
| SLineXSpeed     | Beta      | [0.206] | [0.255]        | [0.151] | [0.143]  | [0.434]   | [0.411]  | [0.118]   |
| SLineXSpeed     | Beta      | 0.757   | 0.436          | 0.926   | 0.172*   | 1.521     | 0.659    | 0.418     |
| SLineXSpeed     | Beta      | [0.700] | [0.470]        | [1.312] | [0.162]  | [1.707]   | [0.633]  | [0.390]   |
| SLSXSpeedXNoUnderstand | CRRRA     | 1.063   | 0.877          | 0.544** | 0.657**  | 1.547     | 1.932*** | 0.529***  |
| SLSXSpeedXNoUnderstand | Beta    | [0.243] | [0.289]        | [0.153] | [0.141]  | [0.440]   | [0.491]  | [0.111]   |
| SLSXSpeedXNoUnderstand | Beta    | 0.976   | 0.395          | 0.585   | 0.089**  | 0.728     | 0.362    | 0.327     |
| SLSXSpeedXNoUnderstand | Beta    | [0.305] | [0.484]        | [0.793] | [0.087]  | [0.861]   | [0.387]  | [0.316]   |

Notes: Panel is balanced with 164 subjects and 3 periods each. Odds ratios reported. Robust standard errors in brackets. Statistical Significance: *** p<0.01, ** p<0.05, * p<0.10. This table organizes the coefficient estimates of CRRA and Beta across data cleaning strategies. For Filter, SLine represents a filter for straight-liners, Speed is a filter for speeders, and NoUnderstand is a filter for those indicating they did not understand either HL or BEN tasks.

We can get a clearer picture of how risk aversion impacts these behaviors by looking at Figure 5 which shows plots of the average marginal effects of CRRA within frequency categories for each behavior. Estimates are shown with whiskers indicating their 95% confidence intervals. Table 5 shows the marginal effect estimates plotted in Figure 5. The figure and table allow us to explore movements between the response categories. For example, while Table 4 suggests more risk averse subjects socialized less, it does not suggest the magnitude of the change (e.g. did they move from “Same” to “Slightly Less” or further to “Not at All”?). Further, Figure 5 and Table 5 allow us to detect if the overall effects in Table 4 are consistent (i.e., risk aversion correlates to movements out of high-risk behaviors and into low-risk behaviors) or spurious (i.e., risk aversion correlates to competing changes in behavior, with one side larger in magnitude than the other). Across all statistically significant odds-ratios, we find evidence for consistent behavioral movements.

Table 4 suggested more risk averse subjects dined in less and socialized less. Figure 5 suggests that more risk averse subjects were significantly less likely to choose approximately any choice except for “Not at All”. From Table 5, a one unit increase in CRRA is associated with a
5.0% increase in likelihood of choosing “Not at All” for DineIn, with 3.3% reduced chance of maintaining behavior. A one unit increase in CRRA relates to a 5.1% increase in the likelihood of “Not at All” for Socialize, with a 3.0% smaller chance of maintaining behavior from 2019. The only increase in choice likelihood that we observe for Dine In and Socialize is for “Not At All”, indicating more risk averse subjects tended to stop eating in restaurants and socializing altogether. Table 4 suggested more risk averse subjects increased handwashing. Figure 5 shows more risk averse subjects had a higher likelihood of reporting “Moderately More” or “Much More,” with reduced probability of choosing “Slightly More” or “About the Same”. From Table 5, a 1-unit increase in CRRA is associated with being 5.1% more likely to report “Much More” and a 4% reduction in the likelihood of reporting “About the Same”. Although not statistically significant overall, Figure 5 and Table 5 demonstrate a pattern for Carryout and Disinfect that is consistent with an inverse relationship with CRRA. Altogether, Figure 5 and Table 5 demonstrates a statistically significant and positive relationship between risk avoidance and risk aversion during the pandemic in three of four categories with unambiguous theoretical predictions (Dine In, Handwash, and Socialize) and a statistically significant but supportive pattern on the fourth (Disinfect), attesting to the external validity of our laboratory measures.

**Figure 5: Average Marginal Effects of CRRA on Behaviors (95% Confidence Intervals)**

*Note: Average marginal effects of CRRA are plotted along with 95% confidence intervals across all frequency levels of each behavior. Negative coefficients indicate risk averse subjects were less likely to choose a given frequency, while positive indicate they were more likely to choose a given frequency. At the*
5% level, subjects were statistically significantly less likely to choose “About Same” for Dine In and Socialize and more likely to choose “Not At All”. At the 5% level, subjects were statistically significantly less likely to choose “About Same” for Handwash and more likely to choose “Much More”. The pattern is similar for Disinfect, but statistically insignificant.

Table 5: Average Marginal Effects of CRRA on Behavioral Frequency Responses

|                | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       |
|----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Grocery        | -0.000    | 0.002     | 0.050**   | 0.023*    | -0.003    | -0.003    | 0.051***  |
| GroceryPickup  | (0.013)   | (0.007)   | (0.022)   | (0.014)   | (0.003)   | (0.002)   | (0.019)   |
| Dine In        | -0.000    | 0.001     | -0.003    | 0.006*    | -0.001    | -0.001    | -0.000    |
| Carryout       | (0.009)   | (0.002)   | (0.002)   | (0.004)   | (0.001)   | (0.001)   | (0.001)   |
| Disinfect      | -0.000    | 0.01      | -0.003*   | 0.002*    | -0.002    | -0.002    | -0.004**  |
| Handwash       | (0.000)   | (0.002)   | (0.001)   | (0.001)   | (0.001)   | (0.001)   | (0.002)   |
| Socialize      | 0.000     | 0.001     | -0.003*   | 0.001     | -0.002    | -0.003    | -0.006**  |
|               | (0.002)   | (0.002)   | (0.001)   | (0.001)   | (0.002)   | (0.002)   | (0.002)   |
| About the Same | 0.000     | 0.002     | -0.033**  | -0.004    | -0.028    | -0.040**  | -0.030*** |
|               | (0.011)   | (0.007)   | (0.015)   | (0.003)   | (0.017)   | (0.016)   | (0.011)   |
| Slightly More  | 0.000     | -0.002    | -0.003    | -0.009*   | -0.003    | -0.010**  | -0.004*   |
|               | (0.004)   | (0.006)   | (0.002)   | (0.005)   | (0.003)   | (0.004)   | (0.002)   |
| Moderately More| 0.000     | -0.001    | -0.001    | -0.010*   | 0.009     | 0.007*    | -0.003*   |
|               | (0.003)   | (0.004)   | (0.001)   | (0.006)   | (0.006)   | (0.004)   | (0.002)   |
| Much More      | 0.000     | -0.003    | -0.004*   | -0.009    | 0.029     | 0.051***  | -0.004**  |
|               | (0.004)   | (0.010)   | (0.002)   | (0.006)   | (0.018)   | (0.019)   | (0.002)   |
| N              | 492       | 492       | 492       | 492       | 492       | 492       | 492       |

Note: Average marginal effects coefficients for CRRA across all frequency levels of each behavior. Standard errors in parentheses: * p < .10, ** p < .05, *** p < .01. Negative coefficients indicate risk averse subjects were less likely to choose a given frequency, while positive indicate they were more likely to choose a given frequency. Subjects were statistically significantly (p < 0.05) more likely to choose “Not At All” for Socialize and Dine In (and less likely to report other categories, particularly “About the Same”). For Handwash, subjects were statistically significantly (p < 0.05) less likely to choose “About Same” and more likely to choose “Much More”. The pattern is similar for Disinfect, but statistically insignificant.

In the appendix, Figure A2 shows the average marginal effects from the same model in Table 4 but with period-by-CRRA interactions to study whether CRRA has dynamic relationships across periods. Table A2 shows the average marginal effect estimates and standard errors used in the plot. Conceptually, behavior can differ across periods for a few reasons. First, localities enacted increasingly stricter regulations over time that limited person-to-person contact. Additionally, information about COVID-19 grew over time. In the first period, the US had not yet declared a state of emergency. Although contraction rates were growing, risks may have been unclear during this period. In periods 2 and 3, information was growing and risks were better-known. In the results that follow, we are not able to test for the specific channels through which risk aversion depends upon pandemic timing and present them as candidate explanations for behavior differences over time.
By parsing the results this way, we see that the marginal effect of risk aversion on Socialization and Handwashing is dynamic. CRRA’s relationship with Socialization grows over time: a 1-unit increase in CRRA is associated with a 3.2% increase in the likelihood of choosing “Not at All” in the first period, a 5.7% increase in the second period, and a 7.2% increase in the third period. This perhaps results from changes in the understanding of risk as the pandemic continued. The marginal effect for CRRA and handwashing is strongest in the first period: a 1-unit increase in CRRA is associated with a 6.3% increase in the likelihood of choosing “Much More” in the first period, but statistically insignificant at the 5% level in the second and third periods. Prior to shelter in place orders, individuals were potentially more willing to frequent public spaces which enhances the need to sanitize more often. Once sheltering (either by rule of government or personal preference), the need for handwashing and sanitizing surfaces may fall for many households that do not venture into public.

**Hypothesis 2:** Guilty individuals (higher $\beta$) increase sanitization and more heavily avoid environments where they may spread the virus during a pandemic, ceteris paribus.

Subjects with more guilt (higher $\beta$) demonstrate greater marginal disutility from advantageous inequality (e.g., they are guilty when ahead of others). In short, we hypothesize that guilty individuals will avoid opportunities to spread the virus to others. We expect $\beta$ is negatively related to DineIn, and Socialize, as a potentially asymptomatic person may pass on the virus in social settings. We expect $\beta$ to be positively related to Handwash and Disinfect. Though similar channels discussed earlier for CRRA, we expect the relationship between guilt and GroceryPickup to be positive, but ambiguously related to Grocery and Carryout. An additional layer with guilt, however, is that subjects may feel guilty about others putting themselves at risk while working in food service or at the grocer (although subjects are not personally transporting the virus to another, they may feel responsible contributing to others’ exposure to it).

Guilt is not statistically significantly related to any pandemic behaviors at the 5% confidence level or better. The results for Carryout are statistically significant at the 10% level. A one-unit increase in $\beta$ reduces the odds of reporting a marginally higher frequency category by 0.172 times. Alternatively, we can say that a one-unit increase in $\beta$ increases the odds of reporting a marginally lower frequency category by 5.8 times for Carryout. For context, a one-unit increase in $\beta$ is the entire range of $\beta$: subjects who chose equality at ($0, 0$) versus subjects who always chose the selfish option, even when it offered equal total monetary gain ($10, 10$). A 0.25 increase in $\beta$ (e.g., a subject accepts equality at $0.50$ lower offer) increases the odds of reporting a marginally lower frequency category by approximately 1.6 times.

Figure 6 plots the average marginal effects of $\beta$ on behavior across each response level from the random effects ordered logit estimation shown in Table 4 (with no collapsing and SLineXSpeed filter). Table 6 shows the average marginal effects and standard errors used in the Figure 6 plot. For Carryout, a one unit increase in $\beta$ increases a subject’s likelihood of choosing
“Not at All” by 12.5%, “Much Less” by 3.5%, and “Moderately Less” by 1.1%. Conversely, a one unit increase in $\beta$ maps to about 5% decreases in choosing “Much More”, “Moderately More”, and “Slightly More” (all significant at the 10% level). Although the relationship between Carryout and guilt is significant at the 10%, this is a considerably weaker result than that of CRRA. Socialization shows a pattern consistent with a positive relationship between guilt and avoiding activities where one could spread the virus, but there is a large confidence interval surrounding each estimate. In total, we do not have enough evidence to support the external validity of the guilt parameter at explaining pandemic behavior.

One reason we may not pick up statistically significant relationships between our utility parameter for advantageous inequality could be that we cannot observe the reference point subjects are using for inequality. Subjects may view being healthy while others are ill as disutility-invoking inequality. For others, they may view stable employment versus others’ pandemic-invoked job loss as unequal. In the later view, guilty subjects may in fact patronize businesses more in order to pass their income to others and resolve the inequality. The wide confidence intervals from Figure 6 suggest this heterogeneity of guilt may be important.

Figure 6: Average Marginal Effects of $\beta$ on Behaviors (95% Confidence Intervals)

Note: Average marginal effects of $\beta$ are plotted along with 95% confidence across all frequency levels of each behavior. Negative coefficients indicate risk averse subjects were less likely to choose a given frequency, while positive indicate they were more likely to choose a given frequency. At the 5% level, no $\beta$ are statistically significant, although Carryout and Socialize show a pattern fitting a negative relationship between guilt and creating the potential to pass the virus to others.
Table 6: Average Marginal Effects of Beta on Behavioral Frequency Responses

|                | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  |
|----------------|------|------|------|------|------|------|------|
| Grocery        | 0.018| 0.021| 0.006| 0.125*| -0.003| 0.002| 0.085|
| GroceryPickup  | (0.059)| (0.027)| (0.114)| (0.068)| (0.008)| (0.005)| (0.090)|
| Dine In        | 0.012| 0.007| -0.000| 0.035*| -0.001| 0.001| -0.001|
| Carryout       | (0.040)| (0.010)| (0.008)| (0.018)| (0.003)| (0.002)| (0.003)|
| Disinfect      | 0.000| 0.007| -0.000| 0.011*| -0.001| 0.001| -0.006|
| Handwash       | (0.001)| (0.009)| (0.006)| (0.006)| (0.004)| (0.003)| (0.007)|
| Socialize      | -0.002| 0.006| -0.000| 0.007| -0.002| 0.002| -0.010|
| (0.008)        | (0.008)| (0.007)| (0.005)| (0.006)| (0.005)| (0.005)| (0.010)|
| Not at All     | -0.015| 0.019| -0.004| -0.023| -0.026| 0.029| -0.049|
| Moderately Less| (0.048)| (0.027)| (0.076)| (0.015)| (0.070)| (0.067)| (0.052)|
| Slightly Less  | -0.005| -0.019| -0.000| -0.050*| -0.003| 0.007| -0.007|
| (0.016)        | (0.025)| (0.006)| (0.026)| (0.009)| (0.016)| (0.007)|
| About the Same | -0.003| -0.012| -0.000| -0.051*| 0.009| -0.005| -0.005|
| (0.012)        | (0.016)| (0.003)| (0.029)| (0.024)| (0.012)| (0.006)|
| Slightly More  | -0.005| -0.029| -0.000| -0.051*| 0.028| -0.037| -0.007|
| (0.016)        | (0.038)| (0.008)| (0.029)| (0.074)| (0.086)| (0.008)|
| Moderately More|         |     |       |       |       |       |       |
| Much More      |         |     |       |       |       |       |       |
| N              | 492   | 492  | 492  | 492  | 492  | 492  | 492  |

Note: Average marginal effects coefficients for across all frequency levels of each behavior. Standard errors in parentheses: * p < .10, ** p < .05, *** p < .01. Negative coefficients indicate risk averse subjects were less likely to choose a given frequency, while positive indicate they were more likely to choose a given frequency. At the 5% level, no β are statistically significant, although Carryout and Socialize show a pattern fitting a negative relationship between guilt and creating the potential to pass the virus to others.

In the appendix, A3 plots the average marginal effects from the same model in Table 4 but with period-by-β interactions to study whether β has dynamic relationships. Table A3 shows the average marginal effect estimates and standard errors. Conceptual, timing can impact the relationship between guilt and behavior through the same information channel alluded to with prior with regards to CRRA: risks become more salient as the pandemic continues. Guilty individuals may feel the chance of passing the virus to another (perhaps asymptptomatically) is greater in later periods. Or they may feel that patronizing businesses puts workers at greater risk in later periods. However, the relationship between guilt and all behaviors except Carryout is statistically insignificant at the 5% level overall. Guilt’s impact on Carryout appears concentrated to the first period of the pandemic: subjects were 14.9% more likely to avoid Carryout (Not at All) in period 1, significant at the 5% level. This result disappears in periods 2 and 3. This is perhaps driven by guilt of patronizing restaurants in which the workers are placing themselves at risk to deliver food, but behavior among guilty and guiltless subjects converges as shelter-in-place orders reduce the amount of trips taken for and availability of prepared food.

5. Discussion

We administered this survey to a nationally representative subject pool in the United States to explore how individual risk- and other regarding preferences relate to individual behaviors.
during a pandemic. This adds to the literature on laboratory-elicited preference measures in explaining real-world behavior. In support of the external validity of risk aversion preference elicitation, we find that risk-averse survey participants more heavily engaged in risk avoidance behaviors in a pandemic than less risk averse survey participants even though the stakes of the mechanism were very small. We discovered patterns in the data suggesting guilt may be related to pandemic behavior, but all estimates were statistically insignificant.

We add to an existing methodological literature with mixed results in general regarding external validity, especially in risk and social preference elicitation experiments. We find support for the external validity of risk preferences. One candidate explanation for the connection between risk preferences and pandemic behavior (in company with the relationship to risky health behaviors from the literature) and not real-world financial decision making is that lab risk elicitation may most accurately correspond to decisions where risk is imminent. That is, when completing a multiple price list, subjects are acutely aware of the risk. During a pandemic, it is salient that all social interactions and sanitary practices are directly linked to contracting the virus. By avoiding a seatbelt, subjects are aware they will eject from the seat upon severe impact. By comparison, the consequences of saving, investing, and purchasing insurance may not be acutely relevant; perhaps lab-elicited risks would map better to insurance scenarios where subjects are tasked with purchasing flood insurance the week before a hurricane. Related to the COVID-19 pandemic, researchers have found that people tend to adopt social distancing practices when there are personal risks to health and finances (Makris, 2020). We leave examination of this candidate explanation for future research.

There are several limitations to the study. The first is that the survey design entails a loss of control. While our filtered sample imposes a few rules to try and detect professional survey takers who are not responding faithfully, we cannot guarantee that others were fully paying attention. We would expect random behavior to inflate standard errors and not systematically bias our results.

Second, our data also relies on remembered behavior and may be subject to recall bias. Because subjects were asked about their pandemic behavior going back only six weeks, we do not expect that recall bias would be particularly large. Moreover, we request subjects report relative changes across each category. Ariely et al. (2003) shows that while experimental subjects are weak at recalling absolute behavior, they are strong at recalling relative changes in behavior.

Third, there may also be a social desirability bias or experimenter demand effects. Although the survey was completely anonymous, subjects may be averse to reporting they were not participating in pandemic-reducing behavior. If subjects that “behaved well” respond honestly, while those that “behaved poorly” inflate their behavior, this would lead to less variation in behavior and likely understate our treatment effects. Our results are biased if experimenter demand effects or social desirability bias influences reported behaviors as well as risk preferences and guilt. It is unlikely social desirability bias or experimenter demand effects are correlated with risk aversion (e.g., subjects who report more handwashing because they think it...
is the right thing to do are not likely to systematically take on less risk in our modified HL procedure because of social or experimenter pressures).

Fourth, we do not vary the order of exercises in our experiment. The social preference exercises may be subject to ordering effects (risk exercises come first and are not subject to ordering bias). For example, if they earned $0 in the lottery they may only accept $10 as the dictator in modified dictator game. We mitigate this ordering effect by withholding exercise outcomes until the end of the experiment so that this subject, for example, does not know the outcome of the lottery while making dictator choices. What we cannot explicitly control for, even with delayed earnings announcement, is that earlier behavior may bias future decisions. For example, a subject who took on a good deal of risk in the lottery may smooth their overall risk by guaranteeing themselves $10 as a dictator (should it be chosen). While this is intuitive, first note that Beta and CRRA is poorly correlated in our data (-0.085). Second, since a relationship between Beta and CRRA does not appear systematic, at worst it adds noise and not bias to our study; we do not anticipate ordering effects between cash incentivized tasks and reported pre-COVID and COVID behaviors.

Finally, our research focuses on mechanisms in the vein of the traditional HL measure of risk-aversion. A future direction of research may be to explore different types of risk-preference on behavior. While economics focuses on a subjective expected utility approach to risk-measurement, psychological research on risk-preferences has been dominated by the psychometric paradigm. According to Leppin and Åro (2009), the psychometric paradigm considers different perceptions to different classifications of risk, such as dread risk (involuntariness, uncontrollability, catastrophic dimensions) or unknown risk (new, unobservable, unknown to science, delayed effects). It may be that more precise risk-measures that incorporate the greater salience of dread risk have more explanatory power on pandemic behavior. It is also possible that, rather than behavioral preferences affecting how people respond during a pandemic, a pandemic changes behavioral preferences. Aguero and Beleche (2009), for example, found long-lasting behavioral effects from the H1N1 pandemic. With its wide scale, it is possible that COVID-19 may exert even more powerful force.
References
Aguero, Jorge M. and Trinidad Beleche. 2017. Health shocks and their long-lasting impact on health behaviors: Evidence from the 2009 H1N1 pandemic in Mexico. Journal of Health Economics 54: 40-55.

Akesson, Jesper, Sam Ashworth-Hayes, Robert Hahn, Robert D. Metcalfe, and Itzhak Rasooly. 2020. Fatalism, Beliefs, and Behaviors During the COVID-19 Pandemic. NBER Working Paper 27245. National Bureau of Economic Research.

Anderson, Lisa R. and Jennifer M. Mellor. 2008. Predicting health behaviors with an experimental measure of risk preference. Journal of Health Economics 27: 1260-1274.

Ariely, Dan, George Loewenstein, and Drazel Prelec. 2003. “Coherent Arbitrariness”: Stable Demand Curves Without Stable Preferences. The Quarterly Journal of Economics 118(1): 73-105.

Bargain, O., & Aminjonov, U. 2020. Trust and Compliance to Public Health Policies in Times of Covid-19. IZA Discussion Paper No. 13205. Institute of Labor Economics.

Benjamin, Daniel J., James J. Choi, and Joshua Strickland. 2010. Social Identity and Preferences. American Economic Review 100: 1913-1928.

Becker G. M., De Groot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. Behavioral Science, 9(3), 226–32.

Benabou, Roland and Jean Tirole. 2006. Incentives and Prosocial Behavior. American Economic Review, 96(5): 1652-1678.

Blanco, Mariana, Dirk Engelmann, and Hans Theo Normann. 2011. A within-subject analysis of other-regarding preferences. Games and Economic Behavior 72: 321-338.

Brodeur, Abel, Idaliya Grigoryeva and Lamis Kattan. 2020. Stay-At-Home Orders, Social Distancing, and Trust. IZA Discussion Paper No. 132234. Institute of Labor Economics.

Burford, Ingrid and Tom Wilkening. 2018. Experimental guidance for eliciting beliefs with the Stochastic Becker-DeGroot-Marschak mechanism. Journal of the Economic Science Association 4(1-2): 15-28.

Campos-Mercade, Pol, Armando Meier, Florian Schneider, and Erik Wengstrom. 2020. Prosociality predicts health behaviors during the COVID-19 pandemic. University of Zurich, Working Paper No. 346.
Cason, Timothy and Charles Plott. 2014. Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing. *Journal of Political Economy* 122(6): 1235-1270.

Charness, Gary and Uri Gneezy. 2010. Portfolio Choice and Risk Attitudes: An Experiment. *Economic Inquiry* 48(1): 133-146.

Charness, Gary, Uri Gneezy, and Alex Imas. 2013. Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization* 87: 43-51.

Charness, Gary, Thomas Garcia, Theo Offerman, Marie Claire Villeval. 2020. Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory. *Journal of Risk and Uncertainty.*

Courtemanche, Charles, Joseph Garuccio, Anh Le, Joshua Pinkston, and Aaron Yelowitz. 2020. Strong Social Distancing Measures in the United States Reduced the COVID-19 Growth Rate. *Health Affairs* 39(7): 1237-1246.

Crosetto, Paolo and Antonio Filippin. 2016. A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics* 19: 613-641.

Dave, Dhaval, Friedson, Andrew I, Matsuzawa, Kyutaro, and Joseph J. Sabia. 2020. When do shelter-in-place orders fight COVID-a9 best? Policy heterogeneity across states and adoption time. *Economic Inquiry*: early view online. https://doi.org/10.1111/ecin.12944

Dasgupta, Utteeyo, Subha Mani, Smriti Sharma, and Saurabh Singhal. 2019. Internal and external validity: Comparing two simple risk elicitation tasks. *Journal of Behavioral and Experimental Economics* 81: 39-46.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jurgen Schupp, and Gert G. Wagner. 2011. Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association* 9(3): 522-550.

Eckel, Catherine and Philip J. Grossman. 2008. Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices. *Journal of Economic Behavior and Organization* 68(1): 1-17.

Eckel, Catherine and Philip J. Grossman, 2008. Men, Women, and Risk Aversion: Experimental Evidence. In *Handbook of Experimental Economics Results*, Vol. 1: 1061-1073.
Fehr, Ernst and Klaus M. Schmidt. 1999. A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 108: 437-460.

Fehr-Duda Helga, Adrian Bruhn, Thomas Epper, and Renate Schubert. 2010. Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty* 40(2): 147-180.

Franzen, Axel and Sonja Pointner, 2013. The external validity of giving in the dictator game. *Experimental Economics* 16: 155-169.

Galizzi, Matteo M., Sara R. Machado, and Raffaele Miniaci. 2016. Temporal stability, cross-validity, and external validity of risk preferences measures: experimental evidence from a UK representative sample. The London School of Economics and Political Science, Department of Social Policy. Working Paper.

Galizzi, Matteo M. and Daniel Navarro-Martinez. 2019. On the external validity of social preference games: a systematic lab-field study. *Management Science* 65(3): 976-1002.

Gneezy, Uri and Jan Potters. 1997. An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics* 112(2): 631-645.

Guth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. An experimental anaylsis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3(4): 367-388.

Harrison, Glenn. 1986. An Experimental Test for Risk Aversion. *Economics Letters* 21(1): 7-11.

Harrison, Glenn. 1990. Risk Attitudes and First-Price Auction Experiment: A Bayesian Analysis. *The Review of Economics & Statistics* 72: 541-546.

Harrison, Glenn and E. Rutstrom. 2008. Risk aversion in the laboratory. *Research in Experimental Economics* 12: 41-196.

Holt, Charles A. and Susan K. Laury. 2002. Risk Aversion and Incentive Effects. *American Economic Review* 92(5): 1644-1655.

Holzmeister, Felix and Matthias Stefan. 2020. The risk elicitation puzzle revisited: Across methods (in)consistency? *Experimental Economics* 23(4)

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 14: 399-425.
Kachelmeier and Shehata. 1992. Examining Risk Preferences under High Monetary Incentives. *American Economic Review* 82(5): 1120-1141.

Kaiser Family Foundation. 2020. Taking Stock of Essential Workers. Accessed on April 24, 2022 at https://www.kff.org/policy-watch/taking-stock-of-essential-workers/

Kessler, Judd B. and Lise Vesterlund. 2015. The External Validity of Laboratory Experiments: the Misleading Emphasis on Quantitative Effects. In *Handbook of Experimental Economics Methodology*.

Killeen, Benjamin, Jie Ying Wu, Kinjal Shah, Anna Zapaishchykova, Philipp Nikutta, Aniruddha Tamhame, Shreya Chakraborty, Jinchí Wei, Tiger Gao, Marike Thies, and Mathias Unberath. 2020. County-level Dataset for Information the United States’ Response to COVID-19. Internet. Cited 2020 Nov 15. Available from https://arxiv.org/pdf/2004.00756v1.pdf.

Levitt, S. D., List, J. A., June 2006. What do laboratory experiments tell us about the real world? *Working Paper*.

Levitt, S. D., List, J. A., 2007a. What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives* 21, 153–174.

Levitt, S. D., List, J. A., 2007b. Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics* 40, 347–370.

Levitt, S. D., List, J. A., 2009. Field experiments in economics: The past, the present, and the future. *European Economic Review* 53, 1–18.

Leppin, Anja and Arja R. Aro. 2009. Risk Perceptions Related to SARS and Avian Influenza: Theoretical Foundations of Current Empirical Research. *International Journal of Behavioral Medicine* 16(1): 7-29.

Loomes, Graham. 1988. When Actions Speak Louder Than Prospects. *American Economic Review* 78(3): 463-470.

Makris, Miltos. 2020. Covid and Social Distancing with a Heterogenous Population. *School of Economics Discussion Paper* No. 2002. University of Kent.

Maloney, William F., and Temel Taskin. 2020. Determinants of Social Distancing and Economic Activity during COVID-19: A Global View. *The World Bank* No. WPS9242: 1–23.
Mizumoto, Kenji, Katsushi Kagaya, Alexander Zerbski, and Gerardo Chowell. 2020. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship. *Euro Surveillance* 25(10): 2000180.

Meredith, Jennifer, Jonathan Robinson, Sarah Walker, Bruce Wydick. 2013. Keeping the doctor away: Experimental evidence on investment in preventative health products. *Journal of Development Economics* 105: 196-210.

Moghadas, Seyed, Meagan Fitzpatrick, Pratha Sah, Abhishek Pandey, Affan Shoukat, Burton Singer, and Alison Galvani. 2020. The implications of silent transmission for the control of COVID-19 outbreaks. *Proceedings of the National Academy of Sciences of the United States of America* 117 (30) 17513-17515.

Murad, Havi, Anat Fleischman, Siegal Sadetzki, Orna Geyer, and Laurence S. Freedman. 2003. Small Samples and Ordered Logistic Regression: Does it Help to Collapse Categories of Outcome? *The American Statistician* 57(3): 155-160.

Nishiura, Hiroshi, and Testuro Kobayashi. 2020. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Disease* 94: 154-155.

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review* 100(1): 557-571.
Appendix B

See online supplemental materials for the survey instrument.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.