Subjective beliefs and economic preferences during the COVID-19 pandemic

Glenn W. Harrison[1,2,3] ⬤ · **Andre Hofmeyr**[3,4] ⬤ · **Harold Kincaid**[3,4] ⬤ ·
Brian Monroe[5] ⬤ · **Don Ross**[2,3,4,6] ⬤ · **Mark Schneider**[2] ⬤ · **J. Todd Swarthout**[2,7] ⬤

Abstract

The COVID-19 pandemic presents a remarkable opportunity to put to work all of the research that has been undertaken in past decades on the elicitation and structural estimation of subjective belief distributions as well as preferences over atemporal risk, patience, and intertemporal risk. As contributors to elements of that research in laboratories and the field, we drew together those methods and applied them to an online, incentivized experiment in the United States. We have two major findings. First, the atemporal risk premium during the COVID-19 pandemic appeared to change significantly compared to before the pandemic, consistent with theoretical results of the effect of increased background risk on foreground risk attitudes. Second, subjective beliefs about the cumulative level of deaths evolved dramatically over the period between May and November 2020, a volatile one in terms of the background evolution of the pandemic.

1 Motivation

What happened to economic preferences and subjective beliefs during the COVID-19 pandemic in 2020? Figures 1 and 2 display the striking evolution of the pandemic in terms of daily and cumulative infections and deaths, respectively, in the United States.[1] While this unfolded, were risk and time preferences unconditionally stable,

---

[1] Figure A1 in the Online Appendix displays the raw and smoothed data side by side. All data and code to replicate our analyses in *Stata* can be obtained at http://www.cear.gsu.edu/gwh/. A larger archive at

✉ Glenn W. Harrison
  gharrison@gsu.edu

Extended author information available on the last page of the article

did they vary with the progress of the pandemic in some stable but conditional manner, or were they apparently disconnected? Did subjective beliefs about the prevalence and mortality of the pandemic track the actual progress of the pandemic, the projections of widely publicized epidemiological models, or neither? These are core questions we evaluated with a series of online experimental waves at monthly intervals between May and November 2020: the vertical gray lines in Figs. 1 and 2 show when our waves were conducted.[2] Subjects were sampled at random from the same population, with no subject asked to participate twice; 598 subjects participated.

Investigating the risk and time attitudes of individuals during COVID-19, and their beliefs about COVID-19 prevalence and mortality, is essential for public policy interventions. For example, if risk and time attitudes differ by demographics, then interventions need to take those differences into account. The same holds for beliefs, where forecasts of the path of the pandemic, relative confidence, and demographic differences are crucial for targeting educational interventions. In turn, rigorous, incentivized elicitation is important for reliable estimates of these attitudes and beliefs.

Section 2 reviews the preferences we elicited: atemporal risk preferences, time preferences, and intertemporal risk preferences. A novel feature of our experiment is a recognition that the volatility of the pandemic could impact *atemporal* and *intertemporal* risk preferences differently. Section 3 reviews the subjective beliefs we elicited: belief distributions, to allow measurement of bias and confidence, with respect to prevalence and mortality effects of the pandemic. Beliefs were elicited for a fixed one-month horizon, as well as for a fixed date of December 1, 2020 that implied declining horizons over the waves of our experiment. Another novel feature of our experiment is to track the extent to which the *bias* and *confidence* of beliefs changed as people lived through the striking evolution displayed in Fig. 1. Section 4 explains the experimental design. Section 5 presents the basic results, tracking the trends in preferences and beliefs over the course of the waves of the experiment. Section 6 discusses the main results in terms of the conditional stability of preferences and beliefs during the pandemic, as well as some comparisons with pre-pandemic preferences. Section 7 concludes with some general lessons. An Online Supplement at https://cear.gsu.edu/gwh/covid19/ documents additional results, elicitation interfaces, choice batteries, and instructions.

We have two major findings. First, atemporal risk preferences during the COVID-19 pandemic appeared to change significantly compared to before the pandemic, but only if one identifies the underlying structure of those preferences to be able to infer a shift from "global probability optimism" to "local probability optimism *and* local probability pessimism." The effect of that change in foreground risk attitude towards

---

Footnote 1 (continued)

https://doi.org/10.17605/OSF.IO/4ZVBG also contains the instructions and task parameters documented in Harrison et al. (2020a) as well as the software source code used to conduct the experiments in *oTree*.

[2] In addition to the incentivized elicitation of preferences and beliefs, we administered a complementary series of non-incentivized survey questions. Apart from the usual demographics, we asked about news sources and the reliability of information on the pandemic. We asked about anxiety and depression, using standard instruments, and about self-assessment of personal risk from the pandemic.

Daily Infections and Deaths in the United States

Raw data from Johns Hopkins Coronavirus Resource Center
Display shows smoothed data
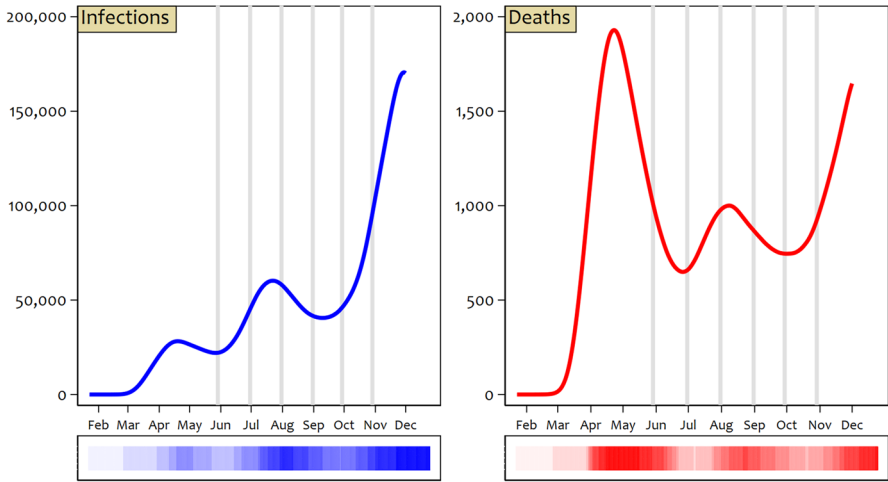Vertical grey lines are dates of experimental waves

Fig. 1 Daily infections and deaths in the United States

Cumulative Infections and Deaths

Raw data from Johns Hopkins Coronavirus Resource Center
Infections were 13,721,304 and deaths were 270,642 as of December 1, 2020
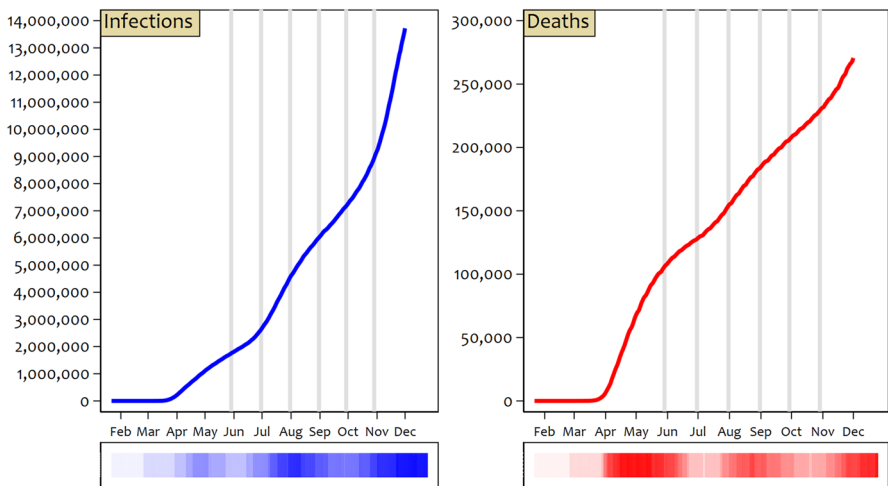Vertical grey lines are dates of experimental waves

Fig. 2 Cumulative infections and deaths

Springer

probabilities is to increase the atemporal risk premium, and is consistent with theoretical results of the effect of increased background risk on foreground risk attitudes.

Second, subjective beliefs about the cumulative level of deaths evolved dramatically over the period between May and November 2020, a volatile one in terms of the background evolution of the pandemic. Specifically, we observed a marked increase between waves 3 and 4 in the confidence with which beliefs were held about cumulative COVID-19 deaths by December 1, 2020. We also found statistically significant evidence of biased beliefs, in the sense that the mean of the estimated belief distributions was *below* the actual number of COVID-19 deaths in waves 1, 2, and 6, but *above* the actual number of deaths in waves 3, 4, and 5. The degree of (statistically significant) bias varied across waves, reaching a high in wave 5 and a low in wave 6. But given the diffuse nature of beliefs, all estimated belief distributions include significant probability density around the actual number of COVID-19 deaths by December 1, 2020, allowing us to assess the "economic significance" or "policy significance" of these biases in subsequent work.

Investigating the risk and time attitudes of individuals during COVID-19, and their beliefs about COVID-19 prevalence and mortality, is essential for public policy interventions. For example, if preferences or beliefs differ by demographics, then interventions need to take this heterogeneity into account. This clearly matters for evaluating the risk perceptions that led people to take certain actions or, in the case of vaccination, not take those actions. Similarly, in §6.A we ask whether the massive "background risk" of the pandemic has effects on "foreground risk preferences," as claimed by some literature and a lot of casual empiricism. Knowledge about beliefs is likewise important for public policy interventions, because forecasts of the path of the pandemic, relative confidence, and demographic differences are crucial for targeting educational interventions. Much of §5.D is about whether the subjective beliefs of individuals between May and November track the likely state of the pandemic as of December 1, 2020: do individuals correctly foresee what is to come? A key insight there is that one must account for the understandable lack of confidence that our subjects had about those beliefs, particularly earlier in the year, rather than just track average beliefs. Of course, rigorous, incentivized elicitation is important for reliable estimates of these preferences and beliefs.

2 Risk and time preferences

We are interested in three broad types of preferences. One is atemporal risk aversion, measuring aversion to stochastic variability of outcomes *at* some point in time. Another is time preference, measuring discounting of time-dated, non-stochastic outcomes. And the third is intertemporal risk aversion, measuring aversion to stochastic variability of outcomes *over* time. Each of these is connected as a matter of theory, and intertemporal risk aversion is literally a theoretical interaction of atemporal risk preferences and time preferences as defined here.

We are also interested in the structural decomposition of these preferences. For atemporal risk aversion, different theories agree on what defines the risk premium, but then decompose it differently. Expected Utility Theory (EUT) attributes all of

the risk premium to aversion to variability of outcomes, measured by the non-constant marginal utility of outcomes as the level of the outcomes vary. If the risk premium is positive, and there is risk aversion, this is diminishing marginal utility and a concave utility function. Rank-Dependent Utility (RDU), due to Quiggin (1982), adds to this account of the risk premium some allowance for various forms of probability weighting, leading to decision weights on the utilities of outcomes that can differ systematically from observed or subjective probabilities. These are just two of the most important structural models, and the ones we consider here.[3]

Similarly, for time preferences, different theories agree on the definition of the discount factor, but then decompose it differently. Exponential discounting models attribute all of the discount factor to a constant variable (utility) cost of time delay, where the variability derives from the time horizon. Quasi-hyperbolic discounting models in addition attribute some of the discount factor to a fixed (utility) cost of any time delay. These are just two of the more important structural models, and the ones we consider here.[4]

Intertemporal risk preferences are currently modeled in terms of several sharply contrasting structural theories. One imposes intertemporal risk neutrality by assuming an additively separable intertemporal utility function. This assumption also ties atemporal risk preferences and time preferences at the hip, in the sense that they cannot be independent of each other. The other theories allow for some non-additivity,

---

[3] We know of six incentivized experiments eliciting atemporal risk preferences during the COVID-19 pandemic, and expect that there are others. Li et al. (2020) had subjects complete a single multiple price list task, and Shachat et al. (2021a, b) had subjects complete a similar task defined over gains as well as one defined over losses from other earnings. Angrisani et al. (2020) had subjects complete the "bomb elicitation task." Guenther et al. (2021) had subjects complete the "bomb elicitation task" and the Binswanger lottery selection task. None of these elicitation methods are designed to allow one to infer structural models of risk preferences such as RDU, and allow only limited inferences about intervals within which risk attitudes lie if one assumes EUT and a single-parameter utility function. Drichoutis and Nayga (2021) elicited risk preferences (and time preferences) using two multiple price lists, with one holding constant the lottery outcomes and the other holding constant the probability. They estimated structural models of EUT *and RDU* risk preferences on a pooled basis, following Harrison and Rutström (2008). This elicitation procedure is not designed to make rich inferences at the individual level. Gassmann et al. (2020) had subjects complete a multiple price list task, with outcomes ranging between €10 and €112. Unfortunately, beyond a €3 participation fee, their subjects only had a 1% chance of being selected to be paid for one of many choices; average payments were only €0.27, although expected earnings were €0.71. Although formally incentivized, we view these as hypothetical choices for practical purposes. There have also been non-incentivized survey questions about risk preference during the pandemic.

[4] We know of only one incentivized experiment eliciting time preferences during the COVID-19 pandemic: Drichoutis and Nayga (2021). They provided subjects with a single multiple price list of choices between a smaller, sooner principal and a larger, later payoff, following Coller and Williams (1999) and Harrison et al. (2002). All of their choices involved a positive "front end delay" on the receipt of the sooner payoffs, limiting them from being able to estimate the Quasi-Hyperbolic discounting model as an alternative to Exponential discounting. They do, however, consider a smooth, hyperbolic model as an alternative to Exponential discounting. And they correctly infer time preferences jointly with inferences about curvature of the utility function based on models of responses to risky lotteries, following Andersen et al. (2008). Gassmann et al. (2020) elicited time preferences using two multiple choices lists, one with a front end delay and one without that delay, allowing estimation of Quasi-Hyperbolic time preferences. Unfortunately, as noted earlier, for all practical purposes their experiments were non-incentivized.

allowing aversion to stochastic variability over time or a preference for temporally correlated variability. The specific alternative that we consider to intertemporal risk neutrality only relaxes the additive separability assumption on the intertemporal utility function.[5]

The concept of intertemporal risk aversion may be less familiar, and arises from theoretical deviations from additively separable intertemporal utility functions. Define a lottery α as a 50:50 mixture of $\{x_t, Y_{t+\tau}\}$ and $\{X_t, y_{t+\tau}\}$, and another lottery ω at the other extreme as a 50:50 mixture of $\{x_t, y_{t+\tau}\}$ and $\{X_t, Y_{t+\tau}\}$, where $X > x$ and $Y > y$. Lottery α is a 50:50 mixture of both bad and good outcomes in time t and $t+\tau$; and ω is a 50:50 mixture of only bad outcomes or only good outcomes in the two time periods. These lotteries α and ω are defined over all possible "good" and "bad" outcomes. If the individual is indifferent between α and ω we say that she is neutral with respect to intertemporally correlated payoffs in the two time periods. If the individual prefers α to ω we say that she is averse to intertemporally correlated payoffs: it is better to have a given chance of being lucky in one of the two periods than to have the same chance of being very unlucky or very lucky in both periods. The intertemporally risk averse individual prefers to have non-extreme payoffs *across* periods, just as the atemporally risk averse individual prefers to have non-extreme payoffs *within* periods. One can also view the intertemporally risk averse individual as preferring to avoid correlation-increasing transformations of payoffs in different periods.[6] Another glance at Figs. 1 and 2 makes it clear why attitudes to risk over time should be important for risk management by individuals over the pandemic.

Finally, we are eventually interested in drawing inferences about preferences at the level of the individual. This leads us to develop designs and batteries that will allow estimation of Bayesian Hierarchical models. But for present purposes we focus on representative agent models, albeit conditional on several observable characteristics of individuals: basic demographics, and of course the wave in which the preferences were elicited. Conditioning on demographics, and interacting these with the study wave, allows for greater comparability across waves for tests of stability.

Our experimental elicitation method for atemporal risk preferences uses unordered binary lottery choices, popularized by Hey and Orme (1994). To elicit time preferences we employ the approach of Andersen et al. (2008, 2014), again with binary choices. To elicit intertemporal risk preferences, which are really the conceptual interaction of atemporal risk and time preferences, we follow Andersen et al. (2018). The econometric methods we use are explained by Harrison and Rutström (2008). Each of these references provide discussion of previous literature, and evaluations of alternative approaches.

---

[5] Epstein and Zin (1989) preferences *require* a specific, empirically-rejected, non-EUT structure on atemporal risk preferences.

[6] Given the extremes of correlations of the payouts in the α and ω lotteries, it is easy to see why this concept is often referred to as "correlation aversion," following Epstein and Tanny (1980), as well as intertemporal risk aversion.

3 Subjective beliefs

We are interested in eliciting subjective belief distributions for individuals with respect to the short-term and longer-term progress of the pandemic. We are specifically interested in beliefs about the levels of infections as well as the levels of deaths of the population of the United States,[7] because policy interventions need to take account of these beliefs and their demographic variation. In short, our focus is on beliefs about the outcomes that unfolded over the course of 2020, displayed in Fig. 1. The short-term horizon is always one month from the day of elicitation. The longer-term horizon was for December 1, 2020, hence a varying-length horizon over the waves of the experiment. Figure 2 displays the cumulative levels of infections and deaths, derived from the daily data of Fig. 1. Hence Fig. 1 should be viewed as the volatile data-generating process behind the cumulative totals about which we elicited beliefs.

A key feature of our elicitation method is that we can make statements about the bias of beliefs as well as the confidence of beliefs. Bias is just the familiar concept from statistical estimation: how different is the weighted average belief from the realized event, or some selected econometric or epidemiological model that might be influential, or the claims of political leaders? All are actually useful metrics for different reasons, so there is not just one measure of bias that is of interest. Confidence is just the familiar concept of imprecision from statistical estimation, most commonly captured by the variance of beliefs about their mean. We prefer to think of confidence more broadly as reflecting the variability of beliefs, so we can also consider skewness and kurtosis, but the point is to pay attention to more than just the weighted average or mode of beliefs. One can only characterize bias and confidence if one elicits subjective belief distributions, which of course allow for the special case of degenerate beliefs held with certainty. Fully Bayesian epidemiological models of COVID-19 infections and deaths provide posterior predictive distributions of future levels, which can be used to also make determinations of whether subjective beliefs are "overconfident" or "insufficiently confident." Knowing bias and confidence are essential for normative educational and policy decisions.

We employ quadratic scoring rules to incentivize subjects to report beliefs over various outcomes. Matheson and Winkler (1976) demonstrated how to extend this scoring rule to elicit beliefs about continuous distributions, which is an appropriate characterization of the COVID-19 outcomes on which we focussed. Harrison et al. (2017) developed experimental tools to implement this elicitation approach, using discrete approximations to the underlying continuous outcomes. In effect, these approximations are akin to eliciting beliefs over histograms defined over the possible outcomes. Hence subjects could allocate 100 tokens over 10 histogram bins, where each bin represented an interval with an upper and lower outcome.

---

[7] We also considered beliefs about infections and deaths of the older segment of the population, although that is not a focus of attention here. Also not discussed here is a parallel, multi-wave experiment using the same elicitation methods in South Africa, albeit with appropriately different survey questions and a slightly different recruitment method.

The pandemic provides a unique setting over which to bracket the range of possible COVID-19 prevalence and mortality outcomes given the proliferation of estimates from epidemiological models. We rely on epidemiological models to bound prevalence and mortality outcomes for one-month, and December 1, 2020 time horizons. We review these models in Harrison et al. (2021), and develop a method to partition these bounds into intervals. We then ask subjects to place bets on these intervals, thereby revealing their beliefs. The intervals are constructed such that if beliefs are consistent with epidemiological models, subjects are best off betting approximately the same amount on every interval. The upshot is that we had four "frames" for each belief question, where each frame had slightly different bin labels to allow us to bracket a priori likely beliefs. Conditional on these frames, our elicitation methods were conventional (for us).

Figure 3 is a screen shot of the experimental software we developed, and documented in Harrison et al. (2020a), to elicit the beliefs of each subject about COVID-19 prevalence and mortality.[8] This subjective belief question was presented to subjects during wave 1 of our study, which took place on May 29, 2020. Figure 3 shows the actual bets, in the form of a token allocation, of subject #127, and the amount to be paid depending on the answer to the question. The answer was verified using the first public report provided by the Centers for Disease Control and Prevention (CDC) after the date in the question, a reference measure which was explained to subjects through audio-visual instructions before they completed the subjective beliefs task.

Figure 4 illustrates how these reports allow us to identify bias and confidence of beliefs. It compares the realized answer, as reported by the CDC, to the question from Fig. 3, and hypothetical bets that vary according to whether they are biased relative to the number of deaths by December 1, 2020, and the confidence with which these beliefs are held. Per the experimental protocol, the official reports from the CDC are treated as the correct answer that determine subject payments. The top left quadrant of Fig. 4 represents an unbiased, but relatively low confidence, set of bets, in the sense that the largest bet was placed on the correct answer, but bets were also made on other events. The bottom left quadrant also represents unbiased beliefs, but held with a degenerate level of confidence in the sense that all tokens were bet on the correct event. The two right quadrants represent biased beliefs because no tokens were allocated to the correct event, but clearly differ according to the strength with which beliefs were held.

We focus here on beliefs during 2020 about the cumulative level of deaths in the United States as of December 1, 2020. A key feature of this horizon, in relation to

---

[8] Harrison et al. (2020a) documents the online procedures used in these experiments, which have several novel features. It does not discuss results. Harrison et al. (2021) examines a specific hypothesis about the subjective beliefs we elicited: are they consistent with those from comparable epidemiological models, where both the subjective beliefs and the forecasts of the epidemiological models were evolving along with the pandemic? The present study examines the effect of the pathway of the pandemic on atemporal risk preferences, time preferences, intertemporal risk preferences, and subjective beliefs with respect to the mortality effects by the end of the year.
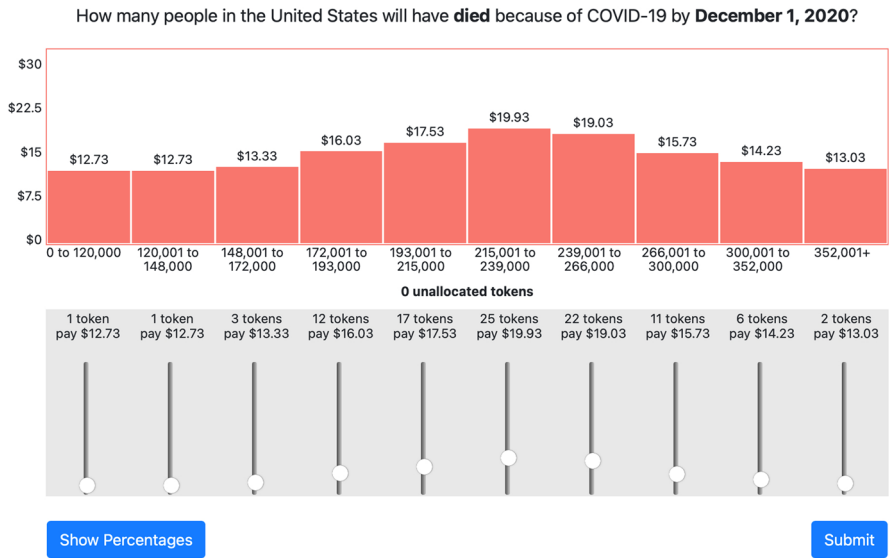
How many people in the United States will have **died** because of COVID-19 by **December 1, 2020**?

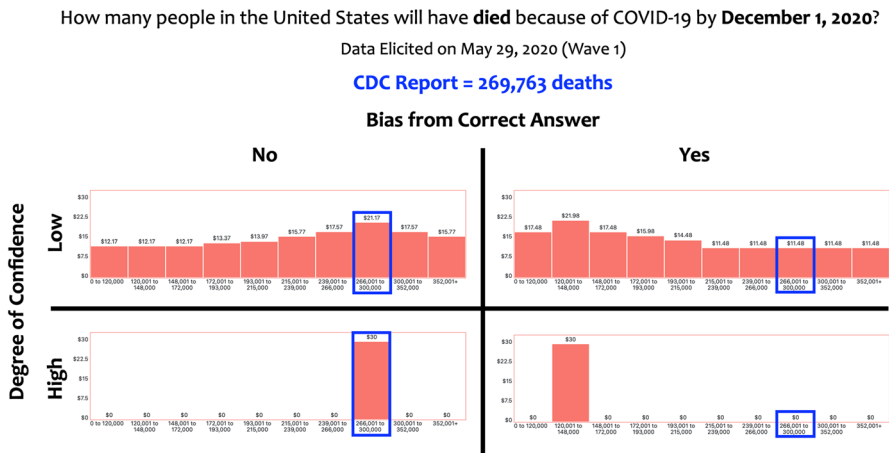**Fig. 3** Illustrative belief elicitation interface

Fig. 4 Bias and confidence in beliefs

the dates of elicitation of beliefs, is that in the life-cycle of the COVID-19 pandemic these are generally "long-run" beliefs for all but the last one or two waves.

This longer horizon for eliciting beliefs is of some significance because of the qualitative evolution of the major epidemiological models over 2020, documented well by Avery et al. (2020). This change is reflected in the densities of the bins over which our subjects allocated their tokens, as explained by Harrison et al. (2021).

4 Experimental procedures

For self-evident reasons we needed to conduct an online experiment. This led us to translate all of our existing tasks for eliciting preferences and beliefs to a software platform that facilitated this, and we selected *oTree*, developed by Chen et al. (2016). Further details of this software aspect of our procedures are given in Harrison et al. (2020a), which also documents the choice batteries employed.

4.1 Sampling

The waves of our experiment were run on May 29, June 30, July 31, August 31, September 29 and October 29. One of the benefits of an online experiment is that physical lab constraints do not limit the number of subjects on a specific day, which also means that we do not have the risk of confounds from day-to-day changes in the news surrounding the pandemic. Our overall sample of 598 consisted of distinct samples of 112, 130, 117, 99, 81 and 59 in each wave, respectively. The lower sample size in wave 6 is the direct result of massive power outages in the greater Atlanta area due to Tropical Storm Zeta. The long-term horizon for belief elicitation for all waves was December 1.

The spacing of the waves was designed to allow a month to pass between each wave, so that we get roughly even representation of behavior over the complete horizon of the experiment. The overall sample size was constrained by available budget at the time, and by widespread beliefs at the outset that the "next 6 months" would be critical for responses to the pandemic. In that belief we were broadly correct, but of course it would have been valuable to continue beyond the planned horizon given the continued spikes in infections and deaths through January 2021.

We recruited undergraduates from Georgia State University (GSU) for the experiment. We already maintained procedures to contact students, and possess a credible reputation amongst these students for paying for their participation. This reputation is especially important given the increased social distancing involved with purely online activities, and tasks relying upon future payments.

We had access to a recruitment database of current students who are interested in taking part in paid research through the Experimental Economics Center (ExCEN) at GSU. When registering in this system, students provide their name, campus ID, email address, and basic demographic information regarding age, gender, and ethnicity. As of May 11, 2020, there was a total of 2497 active subjects in the recruitment database, which is the pool of participants who were invited to take part. The average age of the subjects is 21 years, with a standard deviation of 3.5 years.

During our experiment virtually all in-person classes at GSU were conducted online. The formal announcement of the closing of all University System of Georgia campuses was made on March 16, 2020. Some residence halls at GSU remained open, with social distancing, to accommodate a limited number of domestic and foreign students who had logistical or financial problems with leaving campus. But the vast majority of students were off-campus for the rest of 2020. We have self-reported information on the country, U.S. state, and county in which they completed

the experiment, and 97% completed the experiment in Georgia. In turn the majority of those completed it within Greater Atlanta, a collection of 15 counties.

Figures A2 and A3 in the Online Appendix display infection and death rates for Georgia and Greater Atlanta, which might have been more salient to our subjects than national data. Given the nature of the media coverage at the time, and the focus on these numbers because of the political stakes during the year that included a Presidential election in early November, we believe that national news on infections and deaths dominated attention. However, it is useful to note how different the local patterns of deaths were compared to the national pattern. Death rates in Georgia and Atlanta were less dramatically spiked in structure than in the U.S. as as whole.

Stratified randomization was applied to the recruitment database. The demographic variables age, gender, and ethnicity were used to define the multiple strata of interest to create a set of balanced lists from which to recruit. The lists were defined by two across-subjects treatments: participation payments on offer ($5, $10, and $15); and 2 orders of presenting the health survey and the beliefs task. For each wave six lists were used for recruitment over the six treatments within a wave (3 participation payments × 2 task orders).[9] Because no subject took part in the experiment more than once, our data are a pseudo-panel, and our analyses of results control for basic demographics to make the results more comparable from wave to wave.[10]

We have pre-pandemic evidence for atemporal risk preferences and time preferences from subjects drawn from the same population as our estimates during the pandemic. In the case of atemporal risk preferences we have choices from Harrison et al. (2020d) for 232 subjects that actually participated in our COVID-19 experiment. These choices were made between May and October in 2019, and provide excellent comparisons. In the case of time preferences we have unpublished data on choices from 2013, with no overlap with the COVID-19 subjects. Although dated, this sample was drawn from the same general population, and faced comparable choices.[11]

---

[9] The participation fee treatments are to be used to examine the possible effects that sample selection on unobservables might have on inferences about preferences and beliefs. The order effect treatments are to be used for more detailed evaluation of the possible "priming effect" of health surveys on elicited beliefs about COVID-19.

[10] We are careful to only ever claim that we sample at random from the same population, and that we attempt, with controls of observable demographic characteristics, to mitigate differences across waves. The deeper issue is to control for sample selection effects conditional on *unobservables*, and evaluate the extent to which they are correlated with the preferences and beliefs we examine. Moreover, we would like to know if these potential correlations varied over the pandemic observed across our waves. The population of interest here consists of undergraduates at GSU. Our sample was recruited early in 2019 for general purposes: even that recruitment sample is subject to some potential selection effects, of course. But that recruitment sample was in place well before the pandemic, and we randomly sampled from it to obtain our samples in each wave. Hence there remains potential for pandemic-specific sample selection across waves, particularly given the significant progress of the pandemic over the course of our six waves. Variations in participation fees were implemented, following Harrison et al. (2009, 2020b), to allow for possible sample selection corrections with respect to unobservable risk preferences.

[11] One can also formally consider the effects on atemporal risk preferences and time preferences of sample *attrition* from our pre-pandemic sample from 2019 into the follow-up waves of 2020, using the methods of Harrison et al. (2020b, c).

4.2 Payments

When running an experimental session in a physical laboratory, an experimenter typically pays each subject in cash at the end of the session. Clearly this is not possible online, so alternative payment procedures are needed. Additionally, the temporal aspect of our experiment that involved some future-dated payments required careful consideration of how to remit payments after specific intervals. For example, given the study protocol, it was possible for a participant to be paid over 5 separate transactions that could span up to 7 months from their initial participation.[12] With 598 subjects in the sample and up to 5 potential payments per subject over time, the logistics of making nearly 3000 payments and recording the transactions, as required for filing and reimbursement purposes by GSU, called for an online payment platform to streamline these jobs.

Subjects were able to select their preferred method of payment upon entry into the study, with 54% of the subjects selecting *PayPal* and 46% *Venmo*. The mean, minimum and maximum of subject payments over all tasks were $121.59, $51.19, and $231.71, respectively, including non-salient participation fees. Hence total subject payments were $72,711. Funds of this scale, in the timeframe available, were provided by the Center for the Economic Analysis of Risk (CEAR) at GSU and personal chair funds.

5 Results

5.1 Atemporal risk preferences

Figure 5 displays estimates, with 95% confidence intervals, of atemporal risk premia across waves of our study, assuming alternatively EUT or RDU. The (atemporal) risk premium is just the difference between the expected value of a lottery and its certainty equivalent, where a positive risk premium indicates risk aversion. We use the equi-probable lottery L=($5, 1/3; $30, 1/3; $55, 1/3) as the reference lottery in the figure and the basis for the calculation of risk premia. This reference lottery has an expected value of $30 and is broadly representative of the range of prizes in our atemporal risk preference task. The blue and red bars below Fig. 5 are reproduced from Fig. 1 to show the evolution of infections (blue) and deaths (red) over the course of our study, where darker shades in each bar represent more infections and deaths, respectively.

To calculate risk premia we estimate EUT and RDU models pooling data across all subjects and waves of our study. We adopt the contextual utility error

---

[12] For example, subject payments for participation and the incentivized atemporal risk preference task are due within 24 hours of successfully completing the study. The payment dates for the other incentivized tasks vary, over known parameters, after a given session date. Future payments are made between 0 to 98 days after the session date for the time preference task; 7 to 49 days afterwards for the intertemporal risk aversion task; and the belief task paid out either one month after the session date or December 1, 2020.
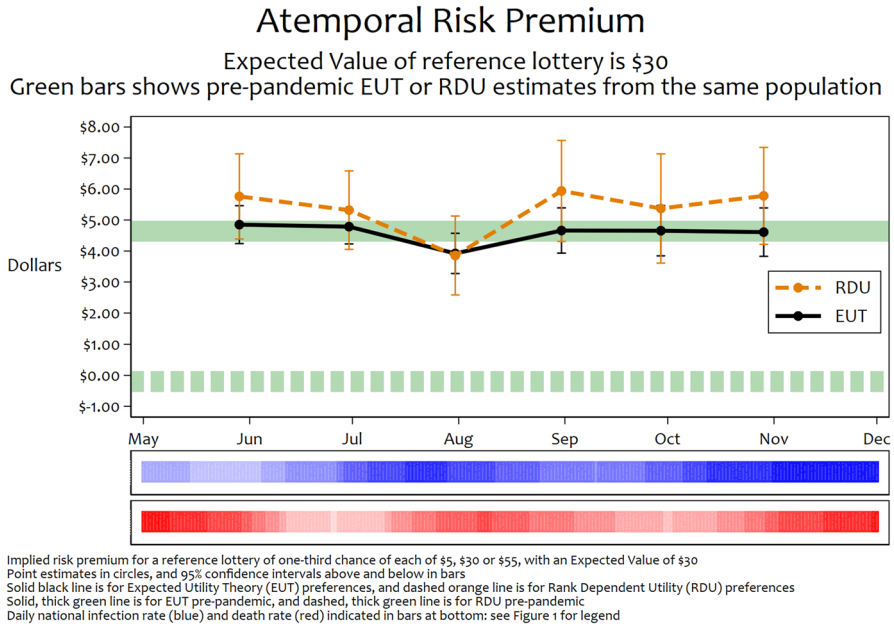
Atemporal Risk Premium

Expected Value of reference lottery is $30
Green bars shows pre-pandemic EUT or RDU estimates from the same population

Implied risk premium for a reference lottery of one-third chance of each of $5, $30 or $55, with an Expected Value of $30
Point estimates in circles, and 95% confidence intervals above and below in bars
Solid black line is for Expected Utility Theory (EUT) preferences, and dashed orange line is for Rank Dependent Utility (RDU) preferences
Solid, thick green line is for EUT pre-pandemic, and dashed, thick green line is for RDU pre-pandemic
Daily national infection rate (blue) and death rate (red) indicated in bars at bottom: see Figure 1 for legend

Fig. 5 Atemporal risk premium

specification of Wilcox (2011) and assume a constant relative risk aversion (CRRA) utility function $u(x_t) = x_t^{1-r}/(1-r)$ for outcomes at time t, for both EUT and RDU. Under EUT $r > 0$ indicates risk aversion, $r = 0$ implies risk neutrality, and $r < 0$ indicates risk seeking behavior. For the RDU model we also employ the flexible two-parameter probability weighting function due to Prelec (1998): $\omega(p) = \exp(-\eta(-\ln p)^{\varphi})$, with $\varphi > 0$ and $\eta > 0$. EUT is nested within RDU when $\eta = \varphi = 1$, but if either or both of these conditions fail then r, φ, and η combine to determine risk preferences under RDU.

In each of the estimated equations[13] we include standard demographic characteristics (e.g., age, gender, and race or ethnicity), subject scores on the Generalized Anxiety Disorder 7-item screen developed by Spitzer et al. (2006), subject scores on the Patient Health Questionnaire measure of depression severity due to Kroenke et al. (2001), a categorical variable capturing the views of each subject on whether the threat of COVID-19 is exaggerated by the media, and dummy variables for each wave of the study.

We also interact the wave dummy variables to capture potential differences in risk preferences according to gender and ethnicity across waves. This "fixed effect" approach to capturing the effects of each wave has the advantage of being agnostic about what aspect of the progress of the pandemic, or just time itself, has on atemporal risk preferences. So these wave dummies are intended to help us statistically

---

[13] For r under EUT and for r, φ, and η under RDU.

capture potential changes over time. It follows that one cannot separate out effects from the pandemic from other things that might have occurred during the experiment. For example, the concurrent national election campaign generated considerable and salient "debate" over the facts and forecasts of the path of the pandemic. We refer to this as our baseline statistical analysis.

On the basis of the pooled preference model with dummies for each wave we directly estimate the atemporal risk premium of the reference lottery above by subtracting the certainty equivalent estimated assuming EUT or RDU from the expected value of the lottery ($30). This approach ensures that any uncertainty in the parameter estimates propagates into the calculation and visualization of risk premia.

Figure 5 shows that risk premia for the EUT model are remarkably stable over time. The risk premium assuming EUT over our elicitation period is shown in the solid black line, with 95% confidence bands. The only statistically significant differences in the CRRA parameter r, which uniquely defines the risk premium under EUT, are between wave 1 and wave 3 ($p = 0.043$) and wave 2 and wave 3 ($p = 0.049$). The average risk premium across all waves is approximately $4.60, and similar to the risk premium estimated for a sub-sample of the same subjects in 2019, and of course pre-pandemic, represented by the solid green bar in Fig. 5.

By contrast, risk premia for the RDU model vary more by wave, reaching a low of $3.86 in wave 3 and a high of $5.94 in wave 4, but the estimates are also more imprecise, at least relative to EUT. The risk premium assuming RDU over our elicitation period is shown in the dashed orange line, also with 95% confidence bands. The estimate of the risk premium for wave 3 is significantly lower than wave 1 ($p = 0.046$), wave 4 ($p = 0.049$), and wave 6 ($p = 0.062$), indicating a decrease in risk aversion in this wave. There are no other statistically significant differences in the atemporal risk premium across waves of our study.

To get some better sense of the effects of the pandemic, we also consider a complementary statistical analysis that replaces these wave dummies with covariates that more precisely reflect the course of the pandemic. We do this by including a covariate for national cumulative deaths, as displayed in Fig. 2. This analysis complements the analysis with wave dummies, and is intended to better track the effects of the pandemic. A more complete analysis of covariates of this kind is best done with estimates of preferences (and beliefs) at the individual level, and should account for lagged effects as well as a wider array of potential confounds. Nevertheless, this analysis indicates that there were no significant effects of the course of the pandemic, over the time period studied, in the atemporal risk preference parameters.[14]

We can comfortably reject the null hypothesis that EUT best characterizes the risk preferences of our sample as a whole, with p-values less than 0.001 for each wave when we test for the absence of probability weighting. However, this characterization of the representative agent hides the fact that we know from previous research that when risk preferences are estimated at the level of the individual, we

---

[14] Online Appendix C provides complete estimates, and formal hypothesis tests. Figure A4 in Online Appendix A displays point predictions of the effects for each parameter during the path of the pandemic that we studied.

see a predictable split of roughly 50:50 between EUT-consistent and RDU-consistent individuals using conventional significance levels to identify evidence of probability weighting.[15] Calculating risk premia at the individual level on the basis of the model of choice under risk that best characterizes that subject's choices will be undertaken in subsequent analyses.

The most striking result comes from comparing the pre-pandemic RDU estimates with the estimates over our waves during the COVID-19 pandemic. The thick dashed green line at the bottom of Fig. 5 shows that our pre-pandemic subjects were roughly *risk neutral overall*, with perhaps a slight hint of being risk-loving. This is completely different from the *overall risk aversion* we see during the pandemic, from the dashed orange line. Figure 6 shows what is happening here. In the top panels we show the pre-pandemic estimates of the probability weighting function and implied decisions weights. These decision weights assume equi-probable lotteries with 2, 3 or 4 prizes; RDU applies to non-uniform probabilities as well, of course, but this assumption makes it easy to see the pure effect of probability weighting. Pre-pandemic we have significant "probability optimism" leading to decision weights that put much greater weight on the better prizes. This is offset by a relatively concave utility function, with CRRA parameter $r = 0.71$, leading to the proximate risk neutrality overall that we see in the risk premium in Fig. 5.

Contrast the RDU risk preference estimates in the bottom two panels of Fig. 6, from our elicitations during the COVID-19 pandemic, recalling that we have a literal overlap of 232 subjects pre-pandemic and during the pandemic.[16] First, we see a decline in the concavity of the utility function, as the CRRA parameter drops from 0.71 to 0.40: *ceteris paribus* any probability weighting, this change would make the representative subject slightly less risk averse. But the *ceteris paribus* is violated, with probability weighting changing from global probability optimism to "inverse-S" probability weighting. In the bottom right panel of Fig. 6 we see the expected pattern of decision weights, with greater weights being placed on the extreme outcomes. The differential weights on the extremes are roughly symmetric, implying that overall risk aversion will be driven generally by the concave utility function. Hence the positive risk premium under RDU we find from our estimates during the pandemic.
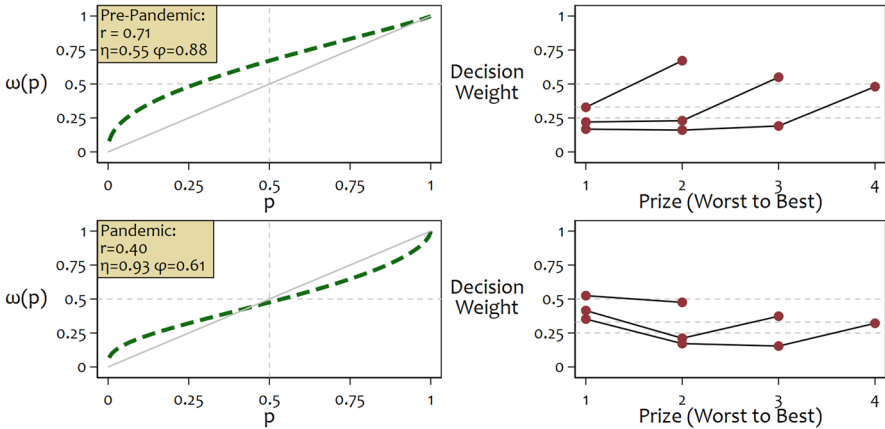
It is useful to stress that the comparison of atemporal risk preferences before the pandemic and during the part of the pandemic we studied refers to a *weighted average* of those preferences *over our 6 waves*. Our complementary analysis of the effect of the path of the pandemic did not show any significant changes *during* the pandemic beyond that weighted average change.

---

[15] The same general conclusion of heterogeneity with respect to the *type* of atemporal risk preference that best characterizes individuals is drawn if one uses Bayesian Hierarchical models or finite mixture models.

[16] Strictly speaking we are comparing the RDU risk preferences of all 598 pandemic subjects with the 232 of them that made risky choices pre-pandemic. The pandemic RDU risk preferences of the sub-sample of 232 that participated in experiments during 2019 and 2020 are virtually the same as the pandemic RDU risk preferences of the full sample of 598. Indeed, the estimates of r, φ, and η are not statistically significantly different for the sub-sample compared to the full sample.

Fig. 6 Probability weighting and Decision weights pre-pandemic and during the pandemic

5.2 Time preferences

In a similar vein as for atemporal risk preferences, Fig. 7 displays estimates, with 95% confidence intervals, of the present value of a $50 reward available in 14 days across waves of our study. We assume either Exponential or Quasi-Hyperbolic discounting. Using each of these characterizations of time preferences, we calculate the present value of $50 in 14 days for two reasons. The amount $50 is representative of the larger, later (LL) rewards in our time preference task for the $40 principal, since the mean of these LL rewards is $48. The horizon of 14 days, which was one of the time horizons in our task, shows the effect of the present-bias parameter β on estimated present values in the Quasi-Hyperbolic model.[17]

To calculate present values we estimate Exponential and Quasi-Hyperbolic discounting models jointly with the RDU model of risk preferences.[18] We pool data across subjects and waves, and admit heterogeneity in our estimates by incorporating

---

[17] The discount rate of the Quasi-Hyperbolic model asymptotes to the discount rate of the Exponential model, represented by the parameter δ, at longer time horizons. This is why the parameter δ in the Quasi-Hyperbolic or β-δ discounting model is often referred to as the long-term discount rate.

[18] Andersen et al. (2008) show how to estimate discounted *utility*, as opposed to discounted *value*, by incorporating the curvature of the utility function identified by a risk preference task when estimating time preferences. Harrison et al. (2018) discuss the importance of accurately characterizing choice under risk when drawing inferences about discounting behavior.
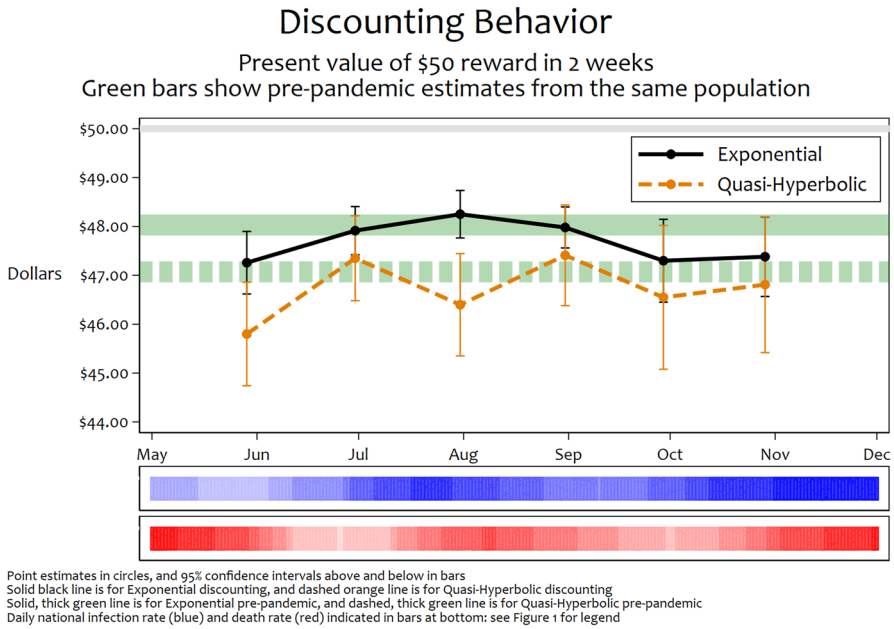
Discounting Behavior
Present value of $50 reward in 2 weeks
Green bars show pre-pandemic estimates from the same population

Point estimates in circles, and 95% confidence intervals above and below in bars
Solid black line is for Exponential discounting, and dashed orange line is for Quasi-Hyperbolic discounting
Solid, thick green line is for Exponential pre-pandemic, and dashed, thick green line is for Quasi-Hyperbolic pre-pandemic
Daily national infection rate (blue) and death rate (red) indicated in bars at bottom: see Figure 1 for legend

Fig. 7 Discounting behavior

the same set of covariates used in our models of risk preferences. Using the pooled, heterogenous preference Exponential and Quasi-Hyperbolic models we then directly estimate the present value of $50 in 14 days, so that any uncertainty in the estimates propagates into inferences we draw from the data.

Figure 7 shows that present values computed using the Exponential model vary slightly over time, but within a very narrow band. Present value estimates for this model range from a low of $47.26 in wave 1 to a high of $48.25 in wave 3. The solid green bar in Fig. 7 represents comparable Exponential discounting estimates obtained in 2013 from a sample drawn from the same population.

Despite varying within a narrow band, there are a number of statistically significant differences in the estimated present value of $50 in 14 days. Specifically, the estimated present value in wave 1 is significantly *lower* than wave 2 ($p=0.068$), wave 3 ($p=0.003$), and wave 4 ($p=0.030$), indicating greater discounting of delayed rewards in wave 1. By contrast, the estimated present value in wave 3 is significantly *higher* than wave 5 ($p=0.023$) and wave 6 ($p=0.041$), indicating less discounting in wave 3. Thus, present value estimates suggest similar levels of discounting in waves 2, 3, and 4, and similar, but higher, levels of discounting in waves 1, 5, and 6. While these present value differences are indeed statistically significant, the only statistically significant difference in estimates of δ is between wave 1 and wave 3 ($p=0.050$), leading us to conclude that time preferences, under the assumption of Exponential discounting, were relatively stable over the course of the pandemic.

Present values computed assuming Quasi-Hyperbolic discounting exhibit more variability relative to the Exponential model, but also more variance: compare the

95% confidence intervals of the Quasi-Hyperbolic model to the confidence intervals for the Exponential model. Quasi-Hyperbolic present value estimates, which combine the present-bias parameter β and the long-term discount rate δ, range from \$45.80 in wave 1 to \$47.41 in wave 4. The dashed green bar in Fig. 6 represents comparable pre-pandemic estimates with the Quasi-Hyperbolic model. The present value estimate in wave 1 is significantly lower than waves 2 and 4 ($p < 0.05$ in both comparisons), indicating higher discounting in wave 1 relative to waves 2 and 4. There are no statistically significant differences in any of the other present value comparisons, demonstrating again the stability of time preferences over the course of our study.

The differences in present values calculated assuming the Exponential or Quasi-Hyperbolic models suggest that present-bias may influence discounting behavior in our sample. Indeed, we can comfortably reject the null hypothesis ($β = 1$) that Exponential discounting best characterizes time preferences for the sample as a whole, with p-values less than 0.001 for each wave.[19] Again, as with the atemporal risk preference results, there is undoubtedly heterogeneity in the model that best characterizes each subject's discounting behavior, and we plan to explore this, and the impact it has on estimated present values, using Bayesian Hierarchical models.

The complementary analysis of the effect of the course of the pandemic over the 6 waves of our study, as captured by the cumulative level of deaths, indicates that the time preference parameter δ *did* change significantly for the Quasi-Hyperbolic model as the pandemic evolved.[20] However, the change in δ has a parabolic U-form, implying estimates in wave 1 that are similar to estimates in wave 6.

5.3 Intertemporal risk preferences

Given the popularity of the CRRA function in the microeconomic and macroeconomic literatures, Andersen et al. (2018) propose this non-additive structural specification of the intertemporal utility function: $U(x_t, x_{t+\tau}) = E[(D_t u(x_t) + D_{t+\tau} u(x_{t+\tau}))^{(1-\rho)}/(1-\rho)]$, where ρ is the intertemporal relative risk aversion parameter ($\rho \neq 1$) and $D_t = \beta/(1+\delta)^t$ is the discount factor of the Quasi-Hyperbolic specification. This intertemporal utility function is separable but not additive when $\rho \neq 0$, and collapses to being additively separable and reflecting intertemporal risk neutrality at $\rho = 0$. Hence we can focus our characterization of intertemporal risk preferences over the pandemic in terms of the evolution of estimates of ρ.

Inferences about ρ depend on the use of EUT or RDU models for atemporal risk preferences, and the use of Exponential or Quasi-Hyperbolic models for time

---

[19] To the extent that other non-Exponential discounting models characterize the data better than the Quasi-Hyperbolic model, the present values generated by the Quasi-Hyperbolic model can be viewed as just a descriptive way of documenting varying discount rates by horizon. There is some evidence from Andersen et al. (2014) that the Quasi-Hyperbolic model does indeed perform poorly for the Danish population when characterizing time preferences compared to the many alternative specifications they explore.

[20] See Fig. A4 in Online Appendix A, and estimates and hypothesis tests in Online Appendix C.

preferences, but allow any combination of these and other possible variants. The notational extensions can be cumbersome, but are not conceptually difficult: see Andersen et al. (2018; §5.4 and §5.5).

If this intertemporal utility function is additively separable, then the inverse of the intertemporal elasticity of substitution is equal to the coefficient of atemporal risk aversion. This assumption is solely one of parametric convenience and is popular in models of intertemporal choice. The linear specification of intertemporal utility is then equal to a weighted sum of atemporal utility flows, where the weights are determined by discount factors. However, when the intertemporal utility function is not additively separable, it is necessary to estimate the extent of intertemporal risk aversion (or intertemporal risk seeking) to accurately characterize preferences over serially correlated intertemporal lotteries.

Figure 8 shows estimates of the intertemporal risk preference parameter ρ across the waves of our study. Figure 8 is derived from an intertemporal risk preference model estimated jointly with the RDU model of choice under atemporal risk and the Quasi-Hyperbolic discounting specification.[21] Again, we pool data across subjects and waves, and incorporate heterogeneity in every equation of interest (r, φ, η, β, δ,

Intertemporal Risk Preferences
Estimated intertemporal risk preference parameter ρ across waves
Intertemporal risk neutrality if ρ=0, and intetemporal risk aversion if ρ>o

Point estimates in orange circle, and 95% confidence intervals above and below in bars
Estimates of parameter ρ determine intertemporal subtitutability of present values of time-dated, risky lotteries
Underlying discounting model is Quasi-Hyperbolic, and underlying atemporal risk preference model is Rank Dependent Utility
Daily national infection rate (blue) and death rate (red) indicated in bars at bottom: see Figure 1 for legend
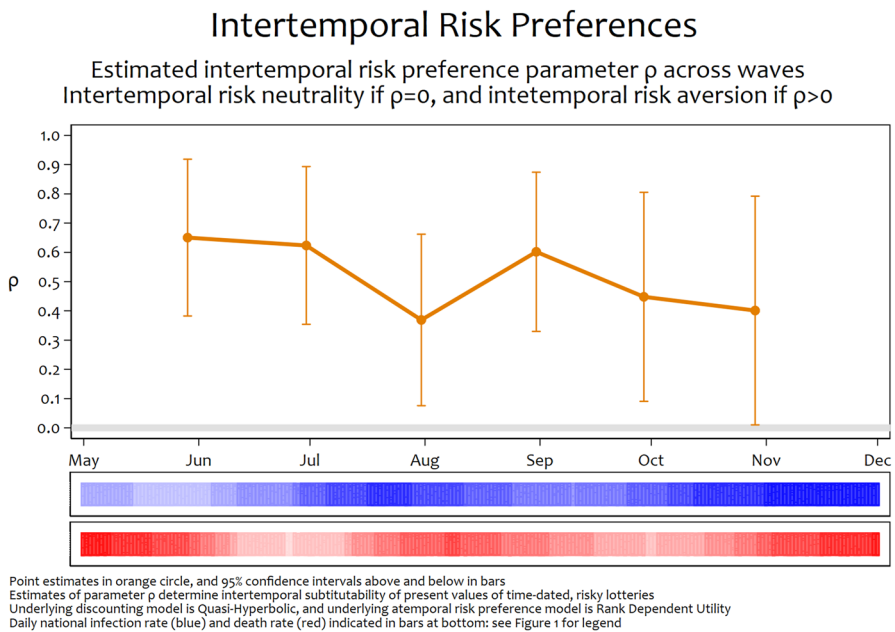
Fig. 8 Intertemporal risk preferences

---

[21] Given that we find statistically significant evidence of RDU and Quasi-Hyperbolic discounting in our data, and that they nest the EUT and Exponential alternatives, we use these models in the joint estimation framework for intertemporal risk preferences.

and ρ) using the same set of covariates and interactions.[22] Fig. 8 clearly shows that
our sample is intertemporally risk averse, with none of the 95% confidence intervals
spanning zero, indicated by the gray line in the figure. Thus, the standard assump-
tion of an additively separable intertemporal utility function does not hold in our
sample. Although there appears to be a downward trend in intertemporal risk aver-
sion over time, none of the estimates for ρ are significantly different across waves
given the large 95% confidence intervals of the estimates. In other words, intertem-
poral risk aversion is stable over time for the period of our elicitations.

The complementary analysis of the effect of the course of the pandemic, over the
6 waves of our study, indicates that the intertertemporal risk preference parameter ρ
did not change significantly as the pandemic evolved.[23]

5.4 Subjective beliefs

The elicited subjective beliefs tell a major story about the manner in which the
implications of the COVID-19 pandemic were perceived by the public, as repre-
sented by our sample. We understand well the sample selection processes that led to
our sample, and the complications of rigorously untangling them to make inferences
about the population. But for the moment accept that this sample tells us how "The
Street" perceived the risks of COVID-19.

Figures 9, 10 and 11 display the recovered beliefs[24] for cumulative deaths by
December 1, 2020. Figure 9 shows in detail the elicitation undertaken on May 29,
2020, wave 1. Figure 10 displays elicitations over all six waves, ensuring compa-
rability across waves by using the same axes for each wave. Figure 11 repeats the
display from Fig. 9, but for the elicitation undertaken on October 29, 2020, our final
wave 6. The range of these *elicitations* was between 0 and 328 million, the U.S.
population at the time. However, the *displays* of recovered beliefs focus on the range
between 0 and 550,000 for Figs. 9 and 10, and then on the smaller range between
210,000 and 310,000 for wave 6 in Fig. 11. The reason for having a smaller range for
wave 6 is apparent from inspection of the relatively tight belief distribution for wave
6 in Fig. 10 in comparison with the belief distributions for waves 1 through 5.

The solid, gray, vertical bar in Figs. 9, 10 and 11 shows the realized outcome for
deaths up to the day prior to the elicitation. If subjects were paying attention to this

[22] We also estimate behavioral error specifications for atemporal risk preferences, time preferences, and intertemporal risk preferences, but these are not equations of interest for present purposes.

[23] See Figure A4 in Online Appendix A, and estimates and hypothesis tests in Online Appendix C.

[24] Harrison et al. (2017) established that EUT decision-makers, with risk preferences in the range widely observed in experiments, will have reports (of token allocations) that provide good approximations to recovered beliefs. However, our pooled data indicate RDU decision-makers, and that can generate first-order changes between reports and recovered beliefs. Harrison and Ulm (2016) provide theoretical results that allow exact calibration of recovered beliefs under EUT and RDU if one knows the (point) estimates of the relevant structural parameters. Our approach is to use those theoretical results on calibration to jointly estimate atemporal risk preferences and beliefs, assuming that beliefs are bounded between 0 and the total U.S. population. We then employ a Beta distribution to constrain beliefs to that interval, allowing the mean and standard deviation of the estimated distribution to vary with the same covariates used to estimate risk preferences.
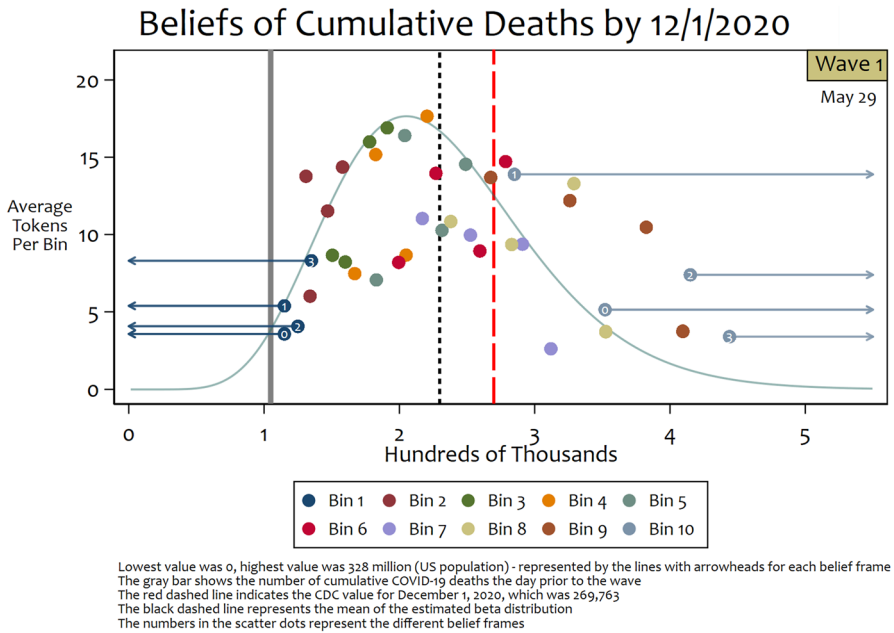
Fig. 9 Beliefs of cumulative deaths by 12/1/2020
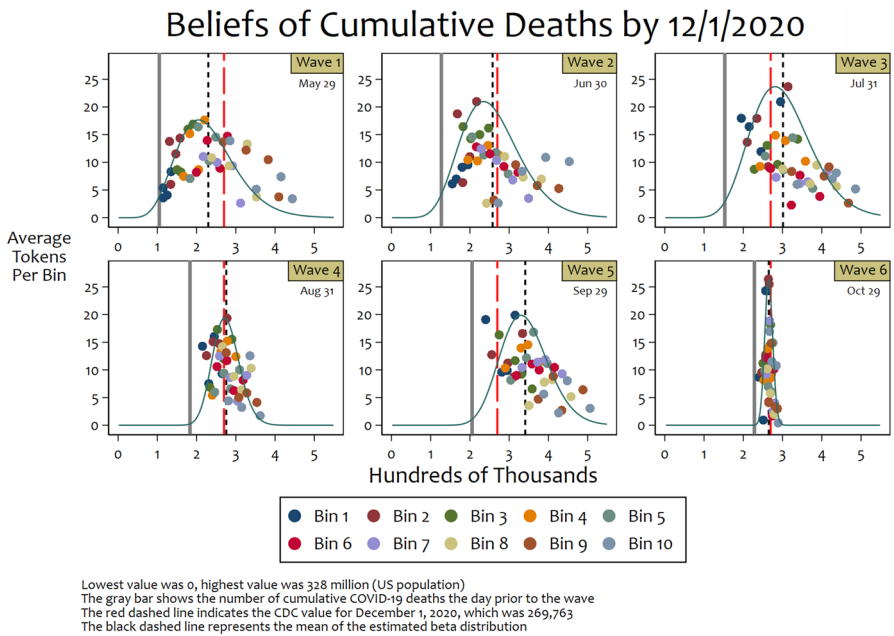

Fig. 10 Beliefs of cumulative deaths by 12/1/2020

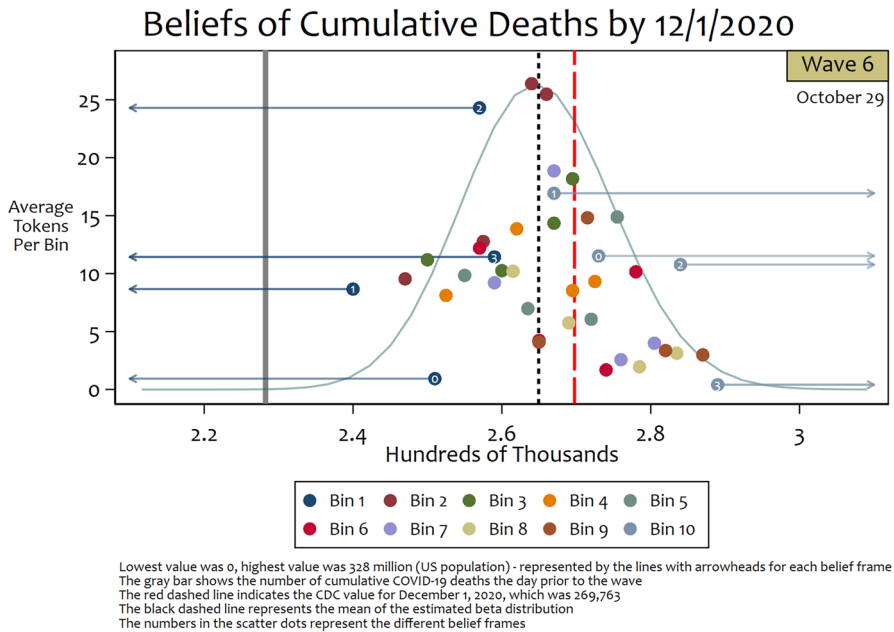## Beliefs of Cumulative Deaths by 12/1/2020

Fig. 11 Beliefs of cumulative deaths by 12/1/2020

outcome, this realized value would then anchor their beliefs on the lower end. The red, long-dashed vertical line shows the realized value for cumulative deaths as of December 1, 2020 reported by the CDC, which was 269,763, so this was in fact the observed outcome over which beliefs were being elicited. The black, short-dashed, vertical line is the mean of the estimated Beta distribution of beliefs in that wave. Close inspection of the *color* of the scatter dots allows one to identify the belief elicitation bin intervals listed in the legend, to facilitate understanding the translation from the belief elicitation interface in Fig. 3 to these inferred beliefs. The *numbers* inside the scatter dots in Figs. 9 and 11 refer to the elicitation frames 0, 1, 2 or 3, as documented in Harrison et al. (2021).

Figures 9 and 11 additionally show arrows on the left and right, to visualize the lowest elicitation bounds and highest elicitation bounds, respectively. The arrows on the left of the display refer to an elicitation bound for bin 1 in Fig. 3 that stretches from the numbered scatter dot on the right of the arrow, down to zero on the left of the arrow. So we can see in Fig. 9 (11) how these lower bounds varied from each other across elicitation *frames* within *wave* 1 (wave 6). Similarly, the arrows on the right of the displays in Figs. 9 and 11 refer to an elicitation bound for bin 10 in Fig. 3 that stretches from the numbered scatter dot on the left of the arrow, up to 328 million on the right of the arrow. Figure 9 shows that in wave 1 we had similar lower bounds for elicitation bin 1 across frames, but much more variability in the upper bounds for elicitation bin 10 across frames. Figure 11 shows that in wave 6 we have considerable variation in both the lower and upper bounds across frames.

Figure 10 allows us to track the evolution of beliefs over the six waves of our experiment. In terms of confidence, subjects held relatively diffuse beliefs during waves 1, 2, and 3: the standard deviations are 75,089, 76,610, and 78,246 deaths, respectively. But something striking then happened between waves 3 and 4. The standard deviation of the belief distribution in wave 4 is only 34,242, implying far greater confidence than in the prior 3 months. Confidence relaxed again in wave 5 (standard deviation = 60,985) compared to wave 4, but then tightened up significantly in wave 6 (standard deviation = 10,180). These variations over time lead us to measure the extent of the difference in the quality of predictions. Although we are limited here to looking at pooled beliefs, albeit adjusted for differences in demographic characteristics across waves, we infer that there were marked changes in confidence over the period of our study.

Inspection of Fig. 1 provides clues as to what led to this revision of beliefs between wave 3 and wave 4. Between waves 2 and 3 a spike of infections (on a daily basis) was underway in the Southern U.S., where most of our subjects were located, and the steady rise in deaths (on a daily basis) was prominent in the news. In a related vein, there was considerable media discussion and political debate about the implications for continued increases in infections and deaths later in the year, as well as recognition that spikes in deaths lagged spikes in infections but were positively correlated with infections. One natural hypothesis is that the elevated infection level made both infections and deaths more salient, leading to the tightening of beliefs evident from wave 4 on.

With regard to bias, Fig. 10 shows that estimates of the mean of the belief distributions in wave 1 and wave 2 (229,756 and 257,648, respectively) are less than the CDC report of 269,763. Although the estimate for wave 1 is significantly less than the CDC report ($p < 0.001$), the estimate for wave 2 is only barely statistically different to the CDC report at $p = 0.088$. By contrast, the mean of the belief distributions in wave 3 and wave 5 are significantly larger than the CDC report ($p < 0.001$), and the mean for wave 4 is higher than the CDC report, but again only barely statistically different at $p = 0.082$. Finally, Fig. 11 shows that the mean of the belief distribution in wave 6 of 264,944 is significantly lower than the CDC report ($p < 0.001$). Comparing the extent of bias across waves, Fig. 10 clearly shows that bias was greatest in wave 1 and wave 5, with a mean difference of 40,007 *less* deaths in wave 1, and 71,157 *more* deaths in wave 5, compared to the CDC report of 269,763. At the other extreme, bias was clearly smallest in waves 4 and 6, with a mean difference of only 5,823 *more* deaths in wave 4, and 4,819 *less* deaths in wave 6, again compared to the CDC report.

The preceding discussion highlights an obvious problem with relying on the standard, statistical definition of bias: with a large enough sample, and a small standard error of the point estimate of the mean of a belief distribution, one will always find statistically significant evidence of bias. The estimate for wave 6 exemplifies this point: the difference between the mean and the CDC report is the smallest of all waves, and yet the difference is statistically significant ($p < 0.001$). Nevertheless, the standard, statistical definition of bias has some value when comparing degrees of bias *across waves*, by focusing solely on the difference between the CDC report and the mean of the estimated belief distribution.

Predicted Subjective Beliefs Over the Path of the Pandemic Studied

Effect on Mean of Belief Distribution

Effect on Standard Deviation of Belief Distribution
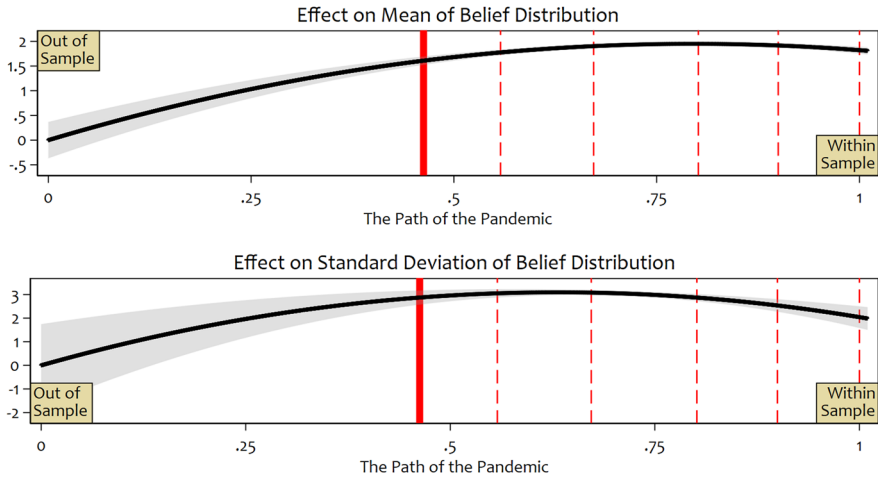
The Path of the Pandemic is captured by the normalized cumulative death level in the United States up to the last wave on October 29, 2020
The model for atemporal risk preferences and subjective beliefs includes this variable in a quadratic form, affecting each parameter
The first wave of elicitation on May 29, 2020 is indicated by a thick, solid red line, and later waves by dashed red lines
The solid black lines show the predicted effect of the Path of the Pandemic, and the shaded regions around the predictions are 95% confidence intervals

Fig. 12 Predicted subjective beliefs over the path of the pandemic studied

From economic and policy perspectives, a more substantive issue than statistical bias is whether estimated belief distributions have "sufficient" density around the CDC report, because this will influence risk mitigation efforts of individuals and public policy approaches to COVID-19 containment. To draw meaningful inferences of this kind, one needs to focus on the standard *deviation* of the belief distribution as opposed to the standard *error* of the point estimate of the mean of the distribution. Harrison et al. (2021) discuss a Bayesian approach that can be used to construct a "region of practical equivalence," or ROPE, around the CDC report and then compare that ROPE to highest density intervals of the estimated belief distributions. We will adopt these methods in subsequent work, which will allow us to say more than whether beliefs were statistically biased: we can determine if those biases have "policy significance."

The complementary statistical analysis of the effect of the course of the pandemic, over the 6 waves of our study, indicates that subjective beliefs did change significantly as the pandemic evolved.[25] Fig. 12 displays the estimated effects. The horizontal axis displays the course of the pandemic, measured by the normalized value of cumulative deaths from COVID-19 during the period of our experiments. So a value of 1 on the horizontal axis denotes the level of cumulative deaths at the time of the final wave, on October 29, 2020. The solid, vertical red line shows the first wave, and the dashed, vertical red lines show later waves. For completeness,

---

[25] See Figure A4 in Online Appendix A, and estimates and hypothesis tests in Online Appendix C.

we project prior to our first wave, back to the origins of the pandemic when there were zero recorded deaths. As one would expect from "out of sample" predictions, the confidence intervals widen when forecasting to before our first wave (the same is true, of course, for forecasts to later pandemic outcomes). The results in Fig. 12 accord with intuition: as cumulative deaths grew, forecast averages for that level grew. And the rationale behind the decline in the standard deviation of beliefs around waves 3 and 4 has been just discussed.[26]

6 Discussion

The data from our experiment span many preferences and beliefs that are relevant to descriptively understanding and normatively evaluating observed responses to the pandemic, and to designing appropriate policy. We stress that such analyses should exploit the fact that our experimental design provides data for rich, structural inferences at the level of the individual. The present analysis focuses on inferences for pooled data, although conditioned on a large set of covariates to mitigate wave-to-wave sampling variability. We focus here on two clear results from that pooled analysis, each of which suggest important hypotheses for future evaluation.

6.1 The effect of background risk on foreground risk aversion

The striking result for atemporal risk preferences appears to be exactly what Quiggin (2003) "anticipated" would happen under RDU with additional background risk being present.[27] The literature on background risk established conditions under which a background risk with zero average effect on payouts would be associated with a change in foreground risk aversion.[28] Intuition allows for any effect, and effects that are sensitive to the size of the background risk. Modest increases in background risk would presumably lead to *more* risk aversion over foreground choices, termed "risk vulnerability." On the other hand, the psychological notion of "diminished sensitivity" suggests that if an individual is already at a point of sufficiently high (background) risk, the addition of a small amount of (foreground) risk will not be particularly salient. And severe, life-threatening background risk can be a factor leading to risk-management decisions that *appear* to be risk-loving behavior, such as the migration patterns observed during times of localized famine (Sen (1981)).

---

[26] Figure 12 indicates a smooth reduction in variance, but Fig. 10 suggests that it was a sharper reduction. These differences are directly due to the use of "fixed effect dummies" for waves in the analyses underlying Fig. 10, and the quadratic (or "smooth polynomial") specification of the path of the pandemic used in the analyses underlying Fig. 12.

[27] The formal results of Quiggin (2003) were derived assuming the special case of RDU referred to as "dual theory," but readily generalize.

[28] Nobody suggests that the pandemic posed zero-mean effects on outcomes for our subjects. This is just assumed to identify theoretically the pure mechanisms at work from additional background risk.

Gollier and Pratt (1996) established conditions, generalized by Eeckhoudt et al. (1996), to show that under EUT the risk vulnerability result is associated with all weakly decreasing absolute risk averse utility functions, such as the CRRA specification we employed. Assuming EUT, the results in Fig. 5 would suggest that the pandemic did *not* generate risk vulnerability, since the implied risk premia are the same as for our pre-pandemic sample. Quiggin (2003) established more general results when one allows for RDU, showing that the effect of background risk depends on the nature of probability weighting, with potential contrasts to the predictions under EUT. Our results confirm the critical importance of the probability weighting "pathway" to the risk premium.

Further investigation of this finding will employ a Bayesian Hierarchical model, following Gao et al. (2020), to allow inferences about individual RDU risk preferences, conditioning on pre-pandemic data as well as the evidence from data collected during the pandemic. It will also consider the effect of imposing a dogmatic prior that individuals view compound lotteries consistently with the Reduction of Compound Lotteries axiom, since the relationship between background risk and foreground risk defines a compound lottery.[29] These analyses will complement an emerging literature on risk preferences in the context of natural disasters and violence, as reviewed by Drichoutis and Nayga (2021).

6.2 Evolving subjective beliefs

Our second striking result is the evolution of subjective beliefs about cumulative deaths due to the pandemic. We observed, again at a pooled level, considerable changes in the average belief over our six waves, as well as considerable changes in the confidence of those beliefs. The mere fact that the mean of beliefs is greater or smaller than the realized value is not the same as whether these differences have "policy significance," and by itself says nothing about overconfidence or insufficient confidence.[30]

The next step, again, is to examine these evolving beliefs in a Bayesian Hierarchical model that allows inferences at an individual level, and to consider the effect of our rich set of covariates about COVID-19 information sources and credibility. This will allow tests of hypotheses about the drivers of any changes in bias and confidence. We can also compare the evolution of these beliefs on "The Street" with the evolution of predictions from epidemiological models, as summarized by Harrison et al. (2021) using our data.

---

[29] There exist formal structural models, referred to as Recursive RDU models, that relax this axiom but maintain the Compound Independence axiom: see Harrison and Ng (2018) for an exposition and application.

[30] There is a literature that refers to overestimation or overplacement as overconfidence, but that is simply poor use of statistical terminology. These are measures of bias, and may or may not reflect statistically significant bias depending on the variance of beliefs. The sole measure of overconfidence one should use is overprecision: see Harrison and Swarthout (2020; §5) for further discussion of this literature.

7 Conclusions

The COVID-19 pandemic presents a remarkable opportunity to put to work all of the research that has been undertaken in past decades on the elicitation and structural estimation of subjective belief distributions as well as preferences over atemporal risk, patience, and intertemporal risk aversion. As contributors to elements of that research in laboratories and the field, we drew together those methods and applied them to a series of multi-wave, online, incentivized experiments in the United States. The resulting data will provide the basis for investigation of several hypotheses emerging from our initial evaluation of pooled data, particularly with respect to atemporal risk preferences and subjective beliefs.

References

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583–618.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting behavior: A reconsideration. *European Economic Review, 71*, 15–33.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2018). Multiattribute utility theory, intertemporal utility, and correlation aversion. *International Economic Review, 59*(2), 537–555.

Angrisani, M., Cipriani, M., Guarino, A., Kendal, R., & Pina, J. (2020). Risk preferences at the time of COVID-19: An experiment with professional traders and students, *Staff Reports #927*, Federal Reserve Bank of New York.

Avery, C., Bossert, W., Clark, A., Ellison, G., & Ellison, S. F. (2020). An economist's guide to epidemiology models of infectious disease. *Journal of Economic Perspectives, 34*(4), 79–104.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88–97.

Coller, M., & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics, 2*, 107–127.

Drichoutis, A. C., & Nayga, R. (2021). On the stability of risk and time preferences amid the COVID-19 pandemic. *Experimental Economics*. https://doi.org/10.1007/s10683-021-09727-6

Eeckhoudt, L. C., Gollier, C., & Schlesinger, H. (1996). Changes in Background Risk and Risk-Taking Behavior. *Econometrica, 64*, 683–689.

Epstein, L. G., & Tanny, S. M. (1980). Increasing generalized correlation: A definition and some economic consequences. *Canadian Journal of Economics, 13*, 16–34.

Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica, 57*(4), 937–969.

Gao, X. S., Harrison, G. W., & Tchernis, R. (2020). Behavioral welfare economics and risk preferences: A Bayesian approach, *NBER Working Paper 27685*, National Bureau of Economic Research.

Gassmann, X., Malézieux, A., Spiegelmanc, E., & Tisserand, J. (2020). Preferences After Pandemics: Time and Risk in the Shadow of Covid-19, *Working Paper*, Burgundy School of Business, Université Bourgogne FrancheComté, France.

Gollier, C., & Pratt, J.W. (1996). Risk vulnerability and the tempering effects of background risk, *Econometrica*, *64*, 683–689.

Guenther, B., Galizzi, M. M., & Sanders, J. G. (2021). Heterogeneity in risk-taking during the COVID-19 pandemic: Evidence from the UK lockdown. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.643653

Harrison, G. W., Hofmeyr, A., Kincaid, H., Monroe, B., Ross, D., Schneider, M., & Swarthout, J. T. (2020a). A case study of an experiment during the COVID-19 pandemic: Online elicitation of subjective beliefs and economic preferences, *Working Paper 2020–21*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University; forthcoming, *Journal of the Economic Science Association*.

Harrison, G. W., Hofmeyr, A., Kincaid, H., Monroe, B., Ross, D., Schneider, M., & Swarthout, J. T. (2021). Eliciting beliefs about COVID-19 prevalence and mortality: Epidemiological models compared with the street. *Methods*, *195*, 103–112.

Harrison, G. W., Hofmeyr, A., Ross, D., & Swarthout, J. T. (2018). Risk preferences, time preferences, and smoking behavior. *Southern Economic Journal, 85*(2), 313–348.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2009). Risk attitudes, randomization to treatment, and self-selection into experiments. *Journal of Economic Behavior and Organization, 70*(3), 498–507.

Harrison, G. W., Lau, M. I., & Williams, M. B. (2002). Estimating individual discount rates for Denmark: A field experiment. *American Economic Review, 92*(5), 1606–1617.

Harrison, G. W., Lau, M. I., & Yoo, H. I. (2020b). Risk attitudes, sample selection and attrition in a longitudinal field experiment. *Review of Economics and Statistics, 102*(3), 552–568.

Harrison, G. W., Lau, M. I., & Yoo, H. I. (2020c). Constant Discounting, Temporal Instability and Dynamic Inconsistency in Denmark: A Longitudinal Field Experiment, *CEAR Working Paper 2020–09*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., & Ulm, E. (2017). Scoring Rules for Subjective Probability Distributions. *Journal of Economic Behavior and Organization, 134*, 430–448.

Harrison, G. W., Morsink, K., & Schneider, M. (2020d). Do No Harm? The Welfare Consequences of Behavioural Interventions, *CEAR Working Paper 2020–12*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Harrison, G. W., & Ng, J. M. (2018). Welfare effects of insurance contract non-performance. *Geneva Risk and Insurance Review, 43*(1), 39–76.

Harrison, G. W., & Rutström, E. E. (2008). Risk aversion in the laboratory. In J. C. Cox & G. W. Harrison (Eds.), *Risk aversion in experiments.* Emerald.

Harrison, G. W., & Swarthout, J. T. (2020). Belief distributions, Bayes Rule and Bayesian overconfidence, *Working Paper 2020–11*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Harrison, G. W., & Ulm, E. R. (2016). Recovering subjective probability distributions, *Working Paper 2016–01*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*(6), 1291–1326.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613.

Li, Z., Lin, P.; Kong, S., Wang, D., & Duffy, J. (2020). Conducting large, repeated, multi-game economic experiments using mobile platforms, *Working Paper*, Department of Economics, University of California, Irvine.

Matheson, J. E., & Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management Science, 22*(10), 1087–1096.

Prelec, D. (1998). The probability weighting function. *Econometrica, 66*, 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*(4), 323–343.

Quiggin, J. (2003). Background risk in generalized expected utility. *Economic Theory, 22*, 607–611.

Sen, A. (1981). *Poverty and famines*. Oxford University Press.

Shachat, J., Walker, M. J., & Wei, L. (2021a). The Impact of an Epidemic: Experimental Evidence on Preference Stability from Wujan, *American Economic Revuew (Papers & Proceedings)*, 111: 302–306.

Shachat, J., Walker, M. J., & Wei, L. (2021b). How the onset of the COVID-19 pandemic impacted prosocial behavior and individual preferences: Experimental evidence from China. *Journal of Economic Behavior and Organization, 190*, 480–494.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097.

Wilcox, N. T. (2011). 'Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk". *Journal of Econometrics, 162*(1), 89–104.

Authors and Affiliations

Glenn W. Harrison[1,2,3] · **Andre Hofmeyr[3,4]** · **Harold Kincaid[3,4]** · **Brian Monroe[5]** · **Don Ross[2,3,4,6]** · **Mark Schneider[2]** · **J. Todd Swarthout[2,7]**

Andre Hofmeyr
andre.hofmeyr@uct.ac.za

Harold Kincaid
harold.kincaid@uct.ac.za

Brian Monroe
brian.monroe@ucd.ie

Don Ross
don.ross931@gmail.com

Mark Schneider
mschneider@gsu.edu

J. Todd Swarthout
swarthout@gsu.edu

[1]   Department of Risk Management and Insurance, Robinson College of Business, Georgia State University, Atlanta, USA

[2]   Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, Atlanta, USA

[3]   School of Economics, University of Cape Town, Cape Town, South Africa

[4]   Research Unit in Behavioural Economics and Neuroeconomics, University of Cape Town, Cape Town, South Africa

[5]   School of Philosophy and School of Economics, University College Dublin, Dublin, Ireland

[6]   School of Society, Politics and Ethics, University College Cork, Cork, Ireland

[7]   Department of Economics and Experimental Economics Center (ExCEN), Andrew Young School of Policy Studies, Georgia State University, Atlanta, USA