A case study of an experiment during the COVID-19 pandemic: online elicitation of subjective beliefs and economic preferences

Glenn W. Harrison^{1,2,3} · Andre Hofmeyr^{3,4} · Harold Kincaid^{3,4} · Brian Monroe⁵ · Don Ross^{2,3,4,6} · Mark Schneider² · J. Todd Swarthout^{2,7}

Received: 15 December 2020 / Revised: 12 October 2021 / Accepted: 29 November 2021 /
Published online: 18 December 2021

© The Author(s), under exclusive licence to Economic Science Association 2021

Abstract

We convey our experiences developing and implementing an online experiment to elicit subjective beliefs and economic preferences. The COVID-19 pandemic and associated closures of our laboratories required us to conduct an online experiment in order to collect beliefs and preferences associated with the pandemic in a timely manner. Since we had not previously conducted a similar multi-wave online experiment, we faced design and implementation considerations that are not present when running a typical laboratory experiment. By discussing these details more fully, we hope to contribute to the online experiment methodology literature at a time when many other researchers may be considering conducting an online experiment for the first time. We focus primarily on methodology; in a complementary study we focus on initial research findings.

Keywords Experimental economic methodology · Online incentivized experiment · Risk preferences · Time preferences · Subjective beliefs

JEL Classification B41 · C88 · C91 · D81 · D90

1 Motivation

With the onset of the COVID-19 global pandemic during early 2020 and the associated social distancing mandates and guidelines that followed, many have adapted to new routines. Experimental economists who rely on traditional physical laboratories for conducting experiments have faced the reality of shuttered laboratories over intervals of uncertain duration. This situation has led many of us to consider conducting online experiments for the first time. We were keenly motivated to conduct a multi-wave

✉ J. Todd Swarthout
swarthout@gsu.edu

Extended author information available on the last page of the article

experiment to assess risk preferences, time preferences, and subjective beliefs related to the pandemic as it was unfolding, and the only way to do this was with an online experiment. We provide a detailed case study of our experiences, and review issues and solutions associated with designing and conducting online experiments.

Our online experiment was designed to explore preferences and beliefs during the evolution of the COVID-19 pandemic in South Africa (S.A.) and the United States (U.S.). While the pandemic unfolded, were risk and time preferences unconditionally stable? Did they vary with the progress of the pandemic in some conditional manner, or were they apparently disconnected from its course? 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 hypotheses we sought to evaluate with a multi-wave online experiment at monthly intervals between May and November 2020. Subjects were drawn at random from the same populations, with no subject asked to participate twice. Parallel experiments were undertaken in the United States ($N=598$) and South Africa ($N=544$).

The U.S. results from the experiment are presented in a complementary study by Harrison et al. (2022). Our objective here is to review the procedures and software we used to implement this online experiment. We provide our software source code for others to use if they find it useful: see <https://github.com/bamonroe/Covid19Experiment>.

In addition to incentivized elicitation of preferences and beliefs, we administered a complementary series of non-incentivized survey questions. We incorporated these survey questions in our overall experiment, even though the main focus was on incentivized responses. It was useful for us to have all of the data, experimental choices and survey responses, collected consistently and within the same framework.

We make no attempt here to undertake methodological evaluations of alternative elicitation methods, alternative recruitment methods, or alternative software for running the experiments online. Our objective was to apply four elicitation tasks that we have employed for many years, to address specific questions about the pandemic in a timely manner. We simply did not have time to evaluate alternatives. We used elicitation tools we have had a large hand in establishing in the field, recruitment from a known database of university students, and software that we were already familiar with. Therefore, this is a case study of *successful* methods, not an argument for *best* methods.

2 Software

The primary technical hurdle when running an online experiment is delivering the experiment software over the internet. We used version 3.3.6 of the oTree framework developed by Chen et al. (2016).¹ oTree experiments run natively within any

¹ Alternatives exist, of course. For example, Duch et al. (2020) describe zTree Unleashed, a method to run zTree online. This relies on Windows emulation to run a zTree server and all clients on a single Linux server, allowing stateful client-server connections via localhost loopback, and then uses the X Window System to stream bitmaps of each client GUI over HTTPS. The result is that each subject views a zTree client (i.e., a zleaf) within a browser, but the client is actually running on a remote server. Other alternatives are considered by Zhao et al. (2020) and Arechar et al. (2018). To allow oTree to generate survey questions, we used version 0.9.2 of oTreeutils due to Konrad (2019).

modern browser and on any modern device, valuable features for online experiments, although one always should check the implementation on different browsers and devices.

When conducting an online experiment, consideration should be given to the scalability of the chosen software. Coming from laboratory experiments, one is accustomed to thinking in terms of the number of sessions that must be conducted in support of a given research project, which is often due to the constraint of the number of physical seats within the laboratory. This usually results in running sessions over many days and weeks. An online experiment does not have this constraint, allowing many more subjects to participate simultaneously. This can greatly reduce the amount of time spent collecting data, but only if one has a software approach that effectively scales, and the Django framework on which oTree 3.3.6 was built has a proven track record of extreme scalability.

We deployed our oTree software over the internet by renting an Ubuntu Linux virtual machine at little cost compared to the overall budget of the study,² and took advantage of well-established open-source software to handle the tasks required to connect computers over the internet. Version 3 of oTree itself is built on the Django web framework,³ and uses the PostgreSQL database software to manage the data needed to deploy oTree and record subject responses. We connected our oTree instances to the internet at-large by passing HTTPS requests through an Apache web server,⁴ which allowed us to deploy multiple instances of oTree on the same virtual machine with different subdomain names. The collective software stack we used, including Linux, Apache, PostgreSQL, and Django, consists entirely of open-source software used globally by millions of users daily, and supported by generous corporate sponsorship to ensure reliability, all free of cost to us as the end users.

All components of this software stack are designed for extreme scalability, which allowed us to conduct experimental sessions with more than 200 simultaneous subjects, and left us with little doubt that we could support several hundred more subjects. By utilizing widely used and widely supported open-source software to manage connectivity over the internet, store and retrieve data from databases efficiently, and serve content to our subjects through web browsers, we were able to focus our efforts primarily on the tasks for which we have the greatest comparative expertise: the design of economic experiments.

² The monthly cost of renting the machine was roughly equivalent to half the average amount paid to a single subject in our experiment. We used a commercially available machine, based in the United States. Time constraints did not allow us to request access to such machines within a university environment, particularly since our universities were under COVID lockdown. Issues of data privacy also arise in the selection of location and institutional setting of the host machine and researchers must abide by any IRB requirements.

³ Django is a popular open-source, Python-based web framework used by many well-known, large-scale websites, such as Instagram and Pinterest, to deploy content. For more information see <https://www.djangoproject.com/>. We used version 3.7 of Python. oTree version 5 has now abandoned the Django framework and instead relies upon the Starlette framework, which appears to be just as scalable. For example, we ran an experiment in June 2021 where we collected data from over 250 people in a single day.

⁴ As of December 2020, approximately 35% of all websites on the internet are deployed using an Apache web server; see https://w3techs.com/technologies/history_overview/web_server.

3 Design considerations

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, so we must have elicitation methods that allow us to jointly estimate time preferences with atemporal risk preferences, and then in turn to estimate intertemporal risk preferences with time preferences and atemporal risk preferences. There are no elicitation methods that allow one to jointly infer all three of these types of preferences with one task.

Our experimental elicitation method for atemporal risk preferences follows the unordered binary lottery choices popularized by Hey and Orme (1994). To elicit time preferences, we employ the approach of Andersen et al. (2014). To elicit intertemporal risk preferences, which really involve the conceptual interaction of risk and time preferences, we follow Andersen et al. (2018). Each of these references provide discussion of previous literature, and evaluations of alternative approaches. Again, our intent here was not to invent any new elicitation mouse-traps, or test the ones we used: this is a policy application, under time pressure as the pandemic evolved.

We are also interested in eliciting *subjective belief distributions* for individuals with respect to the short-term and longer-term progress of the COVID-19 pandemic. We are specifically interested in beliefs about the levels of infection (prevalence) as well as about the levels of deaths (mortality) of the populations of the United States or South Africa. The short-term horizon is always one month from the day of elicitation. The longer-term horizon was December 1, 2020, implying a varying-length horizon over the waves of the experiment.

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. We employ a Quadratic scoring rule (QSR) to incentivize subjects to report beliefs over various outcomes, as implemented by Harrison et al. (2017). In effect, we elicit beliefs about histogram bins defined over the possible pandemic outcomes. Subjects could allocate 100 tokens over 10 histogram bins, where each bin was defined by an upper and lower outcome. Details of the procedures for selecting bin intervals are in Harrison et al. (2021). The upshot is that we had four “frames” for each belief question, randomly assigned to a subject, where each frame had slightly different bin labels to allow us to bracket a priori likely beliefs.

None of these elicitation tasks required any subject interaction with other subjects. This allowed us freedom to provide a specific 24-h window for subjects to respond, facilitating flexible participation around different lock-down schedules and availability. As it happens, oTree allows for synchronous, real-time subject interaction, but this would have required more coordination of subject schedules and raised (potential) issues with intra-session attrition.⁵

⁵ Alternatively, we could have implemented asynchronous interaction using the “strategy method” and resolved outcomes at the end of the calendar day.

4 Experiment procedures

4.1 Recruitment, samples, and retention

The procedures to *recruit and contact* student participants vary slightly between our institutions due to the software and institutional infrastructure available. We recruited students from Georgia State University (GSU) and the University of Cape Town (UCT) for the experiment. We already maintained procedures to contact students, and possess a credible reputation amongst them for paying for their participation. This reputation is especially important given the increased social distance involved with purely online activities, and tasks relying upon future payments.

At GSU researchers have access to a recruitment database of current undergraduate students who are interested in taking part in paid research through the Experimental Economics Center (ExCEN). 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 recruiter database, which is the pool of participants that were invited to take part in the U.S. portion of the study.

Our researchers at the University of Cape Town (UCT) approach this slightly differently and build a recruitment database from scratch for each project undertaken. This is accomplished by submitting a request to the UCT central administration along with a copy of the study announcement. Once approved, an email with the study announcement is sent to all undergraduate and postgraduate students. The announcement outlines the study, the incentivized nature of the research, and then invites the student to fill out an online questionnaire if they are interested in being considered for participation. The questionnaire captures their name, email address, student number, and similar basic demographic information regarding age, gender, and ethnicity. As of May 11, 2020, there were over 1700 students at UCT who had confirmed their interest in participating in the study. This is the pool from which we selected participants in the South African portion of the study.

For both locations *stratified randomization* was performed on the recruitment databases. 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 three across-subject treatments: three initial waves of data collection; three participant payments on offer (\$5, \$10, and \$15 in the U.S. and R40, R60, and R80 in South Africa)⁶; and two orders of presenting the health survey and the beliefs task. Thus a total of 18 balanced sample lists (3 waves $\times$ 3 participation payments $\times$ 2 task orders) were generated for each location and used to recruit participants. These 18 lists were later adopted to serve 6 waves of data collection in total. For each wave, six lists were used for recruitment over the six treatments

⁶ The purchasing power parity (PPP) for the U.S. dollar and South African rand was approximately \$1 = R6.8 in December, 2020. See <https://www.imf.org/external/datamapper/PPPEX@WEO/OEMDC/ADVEC/WEOWORLD/ZAF>.

within the wave (3 participation payments $\times$ 2 task orders). Additional details of randomization procedures are in Online Appendix G.

Using these recruitment lists, initial *emails inviting* students to sign up to participate in the study were sent about 5 days ahead of the launch date of each wave. The email informs potential participants that they are invited to take part in a study regarding population health risks and health outcomes. It lists the date and time of the study, the anticipated time required for their participation, the amount of the fixed participant payment if they complete all tasks, and explains that they will have four additional opportunities to earn extra compensation depending on their decisions.⁷ The initial recruitment email concludes with a request for the student to log into their account and confirm participation.⁸ Participation was capped at 150 participants per wave (25 participants per treatment) in the U.S. and 120 per wave (20 participants per treatment) in South Africa to accommodate budgets. Individuals were told that sign-ups to participate were handled on a first-come, first-served basis, and that once quotas were met no further sign-ups were considered. If the quotas were not met and additional capacity remained, then the same recruitment email was sent a second time 2 days ahead of the wave launch date as a reminder to all non-confirmed individuals to consider participating.

On the evening prior to a wave going live, a list of names and email addresses for all confirmed participants was downloaded, and the session closed from accepting any new registrations. The list of confirmed participants was then matched to a list of unique URLs that contained encrypted wave and treatment information, allowing oTree to correctly display a given subject's participation payment amount, the treatment to which the subject was assigned, and the wave that the subject was in. Since these access URLs are unique to every individual, we used mail merge from Microsoft *Word*, *Excel* and *Outlook* and sent the invitation emails directly to the participant from a university email address about 4 h before the links became live and provided entry to the study.

In addition to the unique URL, the invitation email contained other pertinent information related to the study: the exact 24-h window when the study was available; the expected time to complete; instructions on how to log back onto the study and pick back up where they left off, if necessary; internet browser requirements; an overview of the components of the study; and the payment methods that would be utilized. We provide templates for the emails sent to U.S. subjects in Online Appendix B.

The experimenters *monitored* the number of participants who completed a particular wave throughout the day it took place. In every wave, a reminder was sent at around 3 p.m. to encourage those participants who had yet to complete the study. The encouragement email was another mail merge that provided the unique web

⁷ Participants knew the amount of the fixed participation payment but did not know to which task order they were assigned.

⁸ For GSU students, confirmation of participation was done through the ExCEN recruitment system. For UCT students, participation was confirmed through a special instance of UCT's online collaboration and learning environment, *Vula*, which is used to support UCT courses as well as other UCT-related groups and communities.

link along with brief text reminding the participant that the study had to be finished within the 24-h window for them to be eligible for payment.

4.2 Payment

When running an experiment session in a physical laboratory, an experimenter typically pays each subject in cash at the end of the session. This is not possible online and *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 five separate transactions that could span up to 7 months from their initial participation.⁹ With 598 (544) subjects in the U.S. (S.A.) sample alone and up to 5 potential payments per subject over time, the logistics of making nearly 3000 (2700) payments and recording the transactions, as required for filing and reimbursement purposes by our universities, called for an online payment platform to streamline these jobs. Further, online payment processors differ across countries.

We began this project not knowing if our U.S. subjects would be best served with payments via PayPal, Venmo, or a mailed check originating from an online bank, so we provided all three as options in a test run¹⁰ of our protocol ahead of our first wave. While handling these pilot payments, we quickly learned that our subjects do not select a mailed check as an option. Thus we eliminated that option and focused on PayPal and Venmo payments in the U.S. 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.¹¹ In contrast, payments to all S.A. subjects were made using the Standard Bank Instant Money service.¹²

Subjects were told in the invitation letter and through the informed consent process that, after completing the study, any payments owed and marked payable on the same day as a session would be verified by authorized research staff and sent within 24 h of completion; that any task payments marked payable on a future date would

⁹ For example, subject payments for participation and the incentivized atemporal risk task are due within 24 h 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 and 98 days after the session date for the time preference task; 7–49 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.

¹⁰ We invited graduate research assistants and other students already known by the researchers to participate in a trial run of our online protocol, including payments. No test data were retained.

¹¹ In 2013, PayPal purchased Venmo and their platforms have been increasingly integrated over the years. This is especially true when using a PayPal Business account with the Payouts module enabled. The module allows one to upload a file containing all the payment information for a particular day and will send payments to multiple subjects at once, directly to their PayPal or Venmo account.

¹² Each of the “big 5” banks in S.A. provide instant money payment options. We chose Standard Bank for the following reasons: it is the largest bank in S.A. (and Africa); instant money can be redeemed at Standard Bank ATMs, numerous stores, and Spar supermarkets, one of the largest and most widespread supermarkets in the country; and it provides a “bulk issuing” facility to make multiple payments at once, which is typically only available to companies, but we could apply to use because one of us is a Standard Bank customer.

be sent on that future date; and if a task payment is dependent on the outcome of a future event, and we only know the true outcome after a specified date in the future, research staff will verify the true outcome and make payment within 14 days of that date passing. Subjects were told that they could revisit the custom web link sent in their invitation to view their payment receipt sheet for each incentivized task decision selected for payment, and the dates associated with them.

When conducting a multi-site study involving multiple currencies, researchers must make decisions regarding exchange rates. The official exchange rate¹³ between the two currencies, while variable over the waves of our study, was \$1=R15.14 in December, 2020, and the PPP exchange rate was approximately \$1=R6.80. We opted to take a rough average of the PPP and official exchange rate, and scale our prizes across countries with \$1=R10.

The minimum, mean, and maximum of subject payments over all tasks are \$51.19, \$121.59, and \$231.71 in the U.S., and R460, R1,174, and R2,100 in South Africa, respectively. The median time spent completing all tasks within a session is 147 min.¹⁴ Total subject payments across all waves were \$72,711 in the U.S. and R638,656 in S.A., resulting in a grand total across countries of \$114,900 for subject payments.

4.3 Experimental design

We conducted six online sessions, which we refer to as *waves*, at roughly one month intervals.¹⁵ Each session was active for 24 h on the given session date, which meant that a subject completed all experiment tasks during the time span from midnight to midnight local time. Restricting responses to specific dates in

¹³ See <https://www.resbank.co.za/webindicators/ExchangeRateDetail.aspx?DataItem=EXCX135D>.

¹⁴ Pearson's χ^2 test suggests no statistically significant difference in total time on task across our two study locations (p -value=0.486).

¹⁵ Sessions in both countries were conducted during 2020 on May 29, June 30, July 31, August 31, September 29 and October 29, and referred to as waves 1 through 6. The total number of participants by wave for South Africa are 92, 93, 100, 100, 80, and 79 and for the United States are 112, 130, 117, 99, 81, and 59 for waves 1 through 6, respectively. We had originally planned to have up to 120 in South Africa and 150 in each session in the United States. The recruitment levels for the first 3 waves in the United States are consistent with our experience with in-person recruitment. The lower sample size in wave 6 in the United States is the direct result of massive power outages in the greater Atlanta area due to Tropical Storm Zeta. The small declines in waves 4 and 5 in the United States are probably due to us recruiting from one overall database of potential subjects, and attracting those relatively keen to participate in earlier waves. See Harrison et al. (2022), §4.1 for additional information on the sampling procedure used for the United States waves. The recruitment levels for the first four waves in South Africa are consistent with our experience there with in-person recruitment, and the slight drops in the final two waves are probably due to the end of the academic year. Our experiments were not longitudinal, so these declines are not due to attrition in the usual sense. Instead we view these slight declines as due to a familiar sample selection effect: unobserved willingness to participate in experiments leads to higher recruitment rates in earlier sessions, leaving only those who are less willing to fill later sessions. This sample selection effect is not unique to online samples. Formal econometric methods for addressing sample selection, and attrition for longitudinal experiments, are demonstrated by Harrison et al. (2009, 2020). These studies also demonstrate that these sampling biases can affect inferences about atemporal risk preferences unless corrected. We followed their procedures for identifying sample selection by varying the non-stochastic participation fee offered to subjects, and one could therefore apply formal econometric corrections to the data we collect. Again, none of these general corrections are specific to our experiments being online.

this manner allows our *ex post* analysis to better control for effects of emergent news regarding the pandemic. The initial page provided informed consent information and subjects could not progress past this page if they arrived before the experiment start time.

Since this is an individual choice experiment, each subject was free to arrive, progress, and complete the experiment at her desired pace during the 24-h active period, but was paid for participation only if all tasks were completed by midnight of the session date. Every page displayed a countdown timer showing the time remaining in the 24-h time span of a session. Figure 1 summarizes the hour in which subjects completed the study distributed over the 24 h of the active participation window. The spread of responses between 10 a.m. and midnight suggests that there is some value in allowing online subjects a wider window for a session than is conventional in physical laboratories. This is particularly true for such a long experiment, spanning many tasks.¹⁶ Of course, the decision to widen the participation window of a session will be influenced by whether subjects participate singly or in groups: the larger the group size, the greater the benefit of coordinating on common start times.

Subjects completed the following *task sequence* after progressing past the initial informed consent page: a welcome page which collected payment details and reiterated basic information previously provided during recruitment; a general demographics questionnaire; an incentivized belief elicitation task focusing on COVID-19 infections and deaths; a health questionnaire and short, validated screens for anxiety and depression¹⁷; an incentivized atemporal risk preference task; an incentivized time preference task; an incentivized intertemporal risk preference task; and a final earnings summary. Subjects received a non-salient participation payment and also received earnings from a randomly determined choice from *each* of the four incentivized tasks. We presented a “roadmap” page before each task, which kept each subject apprised of her current progress through the experiment. We provide screen shots of all components of the experiment in Online Appendix C.

We presented *video instructions* to subjects before each task. We had already begun using video instructions for experiments run within the physical laboratory before the pandemic,¹⁸ while also providing printed instructions. Our experience from the laboratory was that most subjects focused on watching and listening to the video instructions instead of reading the printed instructions, so we opted for only video instructions in the present study. We produced the videos with *Camtasia* and *ScreenFlow*, and hosted the videos on Vimeo. We then presented the videos by

¹⁶ In the physical laboratory, we often split tasks over two distinct sessions on separate days, to allow time for all tasks and avoid an unduly long single session.

¹⁷ We swapped the order of the belief elicitation task and the health questionnaire for approximately half of the subjects since these both ask about COVID-19, and it is possible that initial exposure to the questionnaire may affect subsequent responses to the task, or vice versa.

¹⁸ Video instructions provide greater control over the delivery of information to subjects than the common practice of an experimenter standing at the front of the laboratory and presenting or discussing the instructions with the subjects. “Live” presentation of instructions, even if following a script, increases the possibility of extemporaneous speech and other ad hoc verbal and nonverbal communication which could bias subject behavior and confound inferences.

Fig. 1 Local time of day session completed

embedding the Vimeo streaming video player on each instructions page before the task. Additionally, a subject could re-watch instructions during the task as needed. The scripts we used to create the U.S. video instructions are given in Online Appendix D and the videos are all available online.¹⁹

The *survey questions* were split into two questionnaires presented to the subjects. One set consisted of standard demographic questions while the other set included COVID-19 questions along with anxiety and depression screening questions. There were some country-specific variations in questions. The survey questions are provided in Online Appendix E.

The *beliefs task* elicited responses to eight questions pertaining to COVID-19 infection and mortality counts 1 month after the session date and on December 1, 2020. The questions refer to the country in which the session was conducted, and separate questions were asked for the overall population, and over-65 population in the U.S. and over-60 population in S.A. For each question, the response space was partitioned into 10 ordered bins and a subject reported an allocation of 100 tokens over these bins. Each response was incentivized via a QSR that rewarded a minimum of \$0 and maximum of \$30 depending on response accuracy. Figure 2 shows an example of a beliefs task decision page. Online Appendix F provides the specific parameters for each question.

The *atemporal risk preference task* consisted of 90 pair-wise lottery choices, primarily designed to allow econometric estimation of Expected Utility Theory and Rank-Dependent Utility models *at the individual level*. Figure 3 shows an atemporal

¹⁹ Our project page at <https://cear.gsu.edu/gwh/covid19/> provides URLs to the video instructions hosted on Vimeo.

Task 1

Decision: **1** of 8

Show instructions

This example shows a beliefs choice screen with a provisional token allocation. Each bin reports the payoff that will be realized *conditional on the future true value being within the given bin*.

Fig. 2 Interface for belief elicitation

risk preference choice with a “Double or Nothing” option, used to test the Reduction of Compound Lotteries axiom. Each lottery pair was drawn randomly, without replacement, from this battery and presented to subjects one at a time. The task used prize magnitudes between \$0 and \$70, and probabilities which varied in increments of 0.05 between 0 and 1. Parameters for the 90 lottery pairs are provided in Online Appendix F.

The *time preference task* consisted of 20 choice pages with 5 pair-wise choices on each page. Each choice consisted of a smaller, sooner (SS) amount to be paid on an earlier date, and a larger, later (LL) amount to be paid on a later date. On a given choice page the SS and LL delays, along with the SS amount, were held constant, leaving only the LL amount to vary across choices on the page. In turn, all of these values varied across pages. A calendar was presented on each choice page highlighting the current date, and the dates on which the SS and LL rewards would be received. We used three front-end delays (FEDs) to the SS reward: 0, 7, and 14 days. We also used two principals (\$25 and \$40), four time horizons (7, 14, 42, and 84 days), and LL rewards that increased in increments of \$1, up to a maximum of \$64, in the task. Figure 4 presents a screenshot of the time preference

Fig. 3 Interface for atemporal risk preference task

task. Parameters for this task are provided in Online Appendix F, with each subject receiving a random sample from the full set of parameters.

The *intertemporal risk preference task* consisted of 40 pair-wise choices over intertemporal lotteries that varied payouts over time. Each lottery had two possible outcomes, with each outcome paying an amount at *two* future dates. A calendar was included in the task showing the current date and the two future dates when payments would be received depending on the realization of the chosen intertemporal lottery. To construct our battery of intertemporal lotteries, we used a 7-day FED to the sooner reward, probabilities that increased in increments of 0.1 between 0.1 and 0.9, two time horizons of 14 days and 42 days between the rewards, and the following two sets of larger and smaller amounts: (\$45, \$2) and (\$26, \$1). Each intertemporal lottery pair was drawn at random, without replacement, from this battery and presented to subjects sequentially. Figure 5 shows an example of this task and Online Appendix F provides the parameters used for all intertemporal risk preference questions.

After completing all tasks, a subject arrived at the *final earnings summary*. This summary page recapped all earnings owed to the subject for participating in the experiment, which included the fixed participation payment along with the previously realized earnings from each of the four incentivized tasks. Additional tabs on this page allowed the subject to view the specifics of each task realization. Generally speaking, a subject received payment on more than one date given the intertemporal

Task 3

Decision: 1 of 20

Show instructions

June							July							August							September							October								
Su	Mo	Tu	We	Th	Fr	Sa	Su	Mo	Tu	We	Th	Fr	Sa	Su	Mo	Tu	We	Th	Fr	Sa	Su	Mo	Tu	We	Th	Fr	Sa	Su	Mo	Tu	We	Th	Fr	Sa		
1	2	3	4	5	6					1	2	3	4						1					1	2	3	4	5						1	2	3
7	8	9	10	11	12	13	5	6	7	8	9	10	11	2	3	4	5	6	7	8	6	7	8	9	10	11	12	4	5	6	7	8	9	10		
14	15	16	17	18	19	20	12	13	14	15	16	17	18	9	10	11	12	13	14	15	13	14	15	16	17	18	19	11	12	13	14	15	16	17		
21	22	23	24	25	26	27	19	20	21	22	23	24	25	16	17	18	19	20	21	22	20	21	22	23	24	25	26	18	19	20	21	22	23	24		
28	29	30					26	27	28	29	30	31	23	24	25	26	27	28	29	27	28	29	30				25	26	27	28	29	30	31			

14 July 2020 (14 days from today)		OR	25 August 2020 (56 days from today)	
\$40 in 14 days			\$44 in 56 days	
\$40 in 14 days		OR	\$48 in 56 days	
\$40 in 14 days		OR	\$50 in 56 days	
\$40 in 14 days		OR	\$53 in 56 days	
\$40 in 14 days		OR	\$54 in 56 days	

Submit

This example shows a time preferences choice screen with SS delay of 14 days, SS amount of \$40, and LL delay of 56 days. The LL amount varies across choice rows. Provisional selections of the SS option in the first three rows and LL option in the last two rows have been made, and the subject may change these selections or submit.

Fig. 4 Interface for time preference task

structure of several tasks. The final earnings summary clearly conveyed the date on which each payment would be made. The server hosting this experiment was online throughout the timespan of sessions and subsequent payoffs, and subjects were reminded that they could always return to their final earnings summary page, which included no personal identifying information, via their unique participation URL in order to access their payment details.

5 Conclusions

Conducting an online experiment for the first time is challenging, but manageable. As is often characteristic of the dynamics of innovation, we soon discovered some respects in which our proficiency improved. The areas which required the most attention and offered the most scope for new efficiencies were software design, mechanics of recruitment, flowing subjects into the sessions, and payments.

Although short-term start-up costs of adapting to online experiments are inevitable, there are definite benefits to consider. Aside from the obvious benefit of simply collecting data during a time of laboratory closures, we found it attractive to be able to conduct a session with many more subjects simultaneously participating than is possible in a traditional laboratory. A case in point is our experiment: 1142 subjects participated in the 6 sessions, yielding an average of 190 subjects per day. This throughput is likely not obtainable in most physical laboratories, and where, in extremis, it is, it is surely not sustainable. Online experiments have the potential to be highly efficient from a data collection perspective.

Task 4

Decision: **1** of 40

Show instructions

Fig. 5 Interface for intertemporal risk preference task

Another benefit of online experiments is the relative ease of conducting a multi-site, multi-country study. Although any replication of a study protocol will result in some degree of procedural variation and issues of data security, even across locations within a single study, an online protocol using a common software codebase allows researchers to explicitly determine a priori which aspects of the protocol will remain constant across locations and which will vary. This is due primarily to the necessity of structuring most procedures of an online experiment in advance, in contrast to procedures in physical laboratories that naturally afford a higher degree of human interaction.

As advocates of laboratory experiments we would be remiss not to acknowledge the greater experimental control provided by the physical laboratory. The extent to which experimental control is lost in online experiments remains an open question for us, and once more laboratories are open again after the pandemic we will likely see an influx of such assessments. The equilibrium solution, we predict, will feature judicious mixtures of face-to-face and online experimentation, but with a higher prevalence of the latter than was the norm before COVID-19 emerged as a mother of invention.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40881-021-00115-7>.

Funding We are grateful to the Center for the Economic Analysis of Risk and the University of Cape Town for funding this research; in all other respects the funders had no involvement in the research project. We are grateful to the editor and two referees for comments.

References

- 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.
- Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, 21, 99–131.
- 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.
- Duch, M.L., Grossmann, M.R.P., Lauer, T. (2020). z-Tree unleashed: A novel client-integrating architecture for conducting z-tree experiments over the internet. *Working Paper 99*, Department of Economics, University of Cologne.
- 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. <https://doi.org/10.1016/j.ymeth.2021.04.003>
- Harrison, G. W., Hofmeyr, A., Kincaid, H., Monroe, B., Ross, D., Schneider, M., & Swarthout, J. T. (2022). Subjective beliefs and economic preferences during the COVID-19 pandemic. *Experimental Economics*, forthcoming. <https://doi.org/10.1007/s10683-021-09738-3>
- 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., & Yoo, H. I. (2020). Risk attitudes, sample selection and attrition in a longitudinal field experiment. *Review of Economics and Statistics*, 102(3), 552–568.
- Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., & Ulm, E. (2017). Scoring rules for subjective probability distributions. *Journal of Economic Behavior & Organization*, 134, 430–448.
- Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 62(6), 1291–1326.
- Konrad, M. (2019). oTree: Implementing experiments with dynamically determined data quantity. *Journal of Behavioral and Experimental Finance*, 21, 58–60.
- Zhao, S., López Vargas, K., Friedman, D., and Gutierrez, M. (2020). *UCSC LEEPS lab protocol for online economics experiments*. available on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3594027.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Glenn W. Harrison^{1,2,3} · Andre Hofmeyr^{3,4} · Harold Kincaid^{3,4} ·
Brian Monroe⁵ · Don Ross^{2,3,4,6} · Mark Schneider² · J. Todd Swarthout^{2,7}

Glenn W. Harrison
gharrison@gsu.edu

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

- 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, Andrew Young School of Policy Studies, Georgia State University, Atlanta, USA