ONLINE APPENDIX

Appendix B. The Covid Surveys

To assess risk perceptions during the Covid-19 pandemic, we conducted a survey of a diverse sample of over 1,500 Americans. The survey asked an array of questions related to beliefs, preferences and behavioral responses, as well as sociodemographic characteristics. We do not incentivize participants for accuracy given the large uncertainty surrounding the data on many of these issues. We first describe the structure and implementation of the first survey we ran, in May 2020, and then discuss the changes made in Waves 2 and 3. The survey instruments can be found at the conclusion of this section.

WAVE 1 SURVEY

To reach a diverse sample of Americans, we partnered with Qualtrics, who handled the recruitment and compensation of our participants. We specified a desired 1,500 respondents, who met the following quotas:

- Gender: Female (~50%); Male (~50%)
- Age: 18-34 (~25%); 35-49 (~25%); 50 - 69 (~30%); 70 and older (~20%)
- Household Income: <$50K (~35%); $50K-100K (~35%); >100K (~30%)
- Region: Midwest (~20%); Northeast (~20%); South (~40%); West (~20%)
- Race: White (~66%); Black (~12%); Latinx (~12%); Asian (~10%)

To guarantee representation in line with these quotas, the 5 demographic questions requesting this information were presented immediately following the consent form, allowing for screening out of participants as quotas were met. In addition, any participant who indicated they were younger than 18 years old or resided outside of the United States was screened out.

We also wanted to guarantee a minimum level of quality and thoughtfulness of participant responses. Immediately following the demographic screener questions, participants were told: “We care about the quality of our survey data and hope to receive the most accurate measures of your opinions. It is important to us that you provide thoughtful, careful answers to each question in the survey. Do you commit to providing your thoughtful and careful answers to the questions in this survey?” Participants had to select “I commit to providing thoughtful and careful answers” from 3 possible options in order to continue in the survey.

Finally, we wanted to familiarize participants with the question format they would see on much of the survey, while providing a further screen of their thoughtfulness and quality. Because objective likelihoods of suffering particular health consequences related to Covid-19 are in some cases quite small, it could be difficult for a typical participant to express their beliefs in a probability or percentage format. More generally, individuals often have difficulty interpreting probabilities, particularly in more abstract contexts. Gigerenzer and Hoffrage (1995) suggest that presenting or eliciting frequencies, rather than probabilities, improves participant understanding.

To address these concerns, we asked questions in terms of frequencies, but also began by familiarizing participants with the question format. We told respondents: “Many of the questions on this survey will ask you to make your best estimate as to how many out of 1,000 Americans will experience different events or have different features. To give you some practice and get you used to thinking in these terms, we have a few example questions for you to work through.”

For the first example, participants were told that, according to the United States Census, approximately 20 out of 1,000 Americans live in Massachusetts, and that this is equivalent to approximately 2% or 2 out of every 100. We then asked them, using this estimate, to tell us how many out of 5,000 Americans live in Massachusetts. Participants had to provide an answer of 100 (i.e. 2% of 5,000) in order to continue in the survey.

For the second example, participants were told that they would estimate the size of a group of Americans with a certain attribute. In particular, they were asked to provide their guess of how many Americans have red hair, both out of 1,000 and out of 10,000 (these two answer fields appeared in a random order). Only participants who estimated that fewer than 1,000 out of 1,000 Americans had red hair could continue in the survey. Participants also had to provide consistent answers: their answer to the “out of 10,000” question had to be 10 times their answer to the “out of 1,000” question in order to continue in the survey.

Following their successful completion of this question, we informed participants of what their red hair estimate implied both as a percentage and in terms of how many Americans out of 100, out of 1,000, and out of 100,000 would have red hair. We also provided an accurate estimate as a useful reference point: roughly 15 out of 1,000 Americans are estimated to have red hair, which we described to them as 1.5%, 1.5 out of 100, 15 out of 1,000, or 1,500 out of 100,000.

After completing these questions in line with our specified quality conditions, participants continued to our questions of interest. Qualtrics did not provide us with data on the participants who were screened out, nor did they inform us of the rate at which participants were screened out.

Participants completed several blocks of questions: Covid-19 Related Health Risks for People Like Self, Other Health Risks for People Like Self, Economic and Other Risks, Covid-19 Related Health Risks for Others, Demographics, and Preferences and Behavior. We asked about many sources of risk to assess whether the salience of Covid-19 health risks influences how other health and economic risks are judged.

A. *Covid-19 Related Health Risks for People Like Self*

In this block, we first ask participants to think about 1,000 people "very similar to you (i.e., in terms of age, gender, race socioeconomic status, zip code, health status, etc.)". We then ask "of these 1,000 people, how many do you believe will contract Covid-19 in the next 9 weeks?" We provide a time-frame to make the question more concrete, and we choose 9 weeks because we anticipate running multiple waves of this survey over time, approximately 9 weeks apart. We do not bound participants' answers.

Because this is the first risk elicitation question of this form, we contextualize this answer for all participants. In particular, after they provide their response, they are taken to a new survey page that informs them about the answer they just gave. Suppose they answered that they believe 300 of 1,000 people similar to them will contract Covid-19 in the next 9 weeks. The survey then repeats to them: "Just to clarify, by entering 300 for the question on the previous page, you are indicating that you believe 300 out of 1,000 people very similar to you will contract Covid-19 in the next 9 weeks. This is equivalent to 30%." Each participant is then asked if they would like to revise their answer, and if they indicate that they would, they have the opportunity to provide a new answer. In our analysis, we replace initial estimates with revised estimates for all participants who indicated they wished to revise their answer.

This block on Covid-19 related health risks for self includes two other risk assessment questions. Each asks people to consider 1,000 people very similar to them *who contract Covid-19 in the next 9 weeks*. They are then asked to estimate how many of these 1,000 people very similar to them who contract Covid-19 will require hospitalization. They are also asked to estimate how many of 1,000 people very similar to them who contract Covid-19 will die. The questions about hospitalization and death due to Covid-19 are both conditional on contracting Covid-19. These questions attempt to isolate beliefs about potential health consequences due to Covid-19 from beliefs about its prevalence or contagiousness.

B. *Other Health Risks for People Like Self*

We are interested in understanding how perceptions of Covid-19 related health risks compare to and interact with beliefs about other serious health risks faced by this same population. In this next block of questions, we adapt a similar question format to assessing other health risks. For each of the questions, participants are again prompted to consider 1,000 people "very similar to you (i.e., in terms of age, gender, race socioeconomic status, zip code, health status, etc.)". They are asked to estimate, out of those 1,000, how many will: (i) require hospitalization for a reason other than Covid-19 in the next 5 years, (ii) die for a reason other than Covid-19 in the next 5 years, (iii) have a heart attack in the next 5 years, and (iv) develop cancer in the next 5 years.

C. *Economic Risks and Other Threats*

We would also like to understand how participants perceive the economic risks surrounding the Covid-19 pandemic. Because these questions do not easily lend themselves to the "out of 1,000" format used for the health questions, we use the Likert-scale. For four different economic outcomes, we ask participants to assess the likelihood of this outcome on a 1 – 7 scale, where 1 indicates extremely unlikely and 7 indicates extremely likely.

We present two pairs of questions, the first related to the stock market and the second related to the unemployment rate. Within each pair, we present both a favourable and unfavourable outcome. For the stock market the two outcomes are: (i) the U.S. stock market drops by 10% or more in the next 9 weeks, (ii) the U.S. stock market grows by 10% or more in the next 9 weeks. For the unemployment rate the two outcomes are: (i) the U.S. unemployment rate reaches 20% or more in the next 9 weeks, and (ii) the U.S. unemployment rate falls below 5% in the next 9 weeks. By eliciting beliefs about good and bad outcomes we can assess not only general optimism or pessimism, but also perceived tail uncertainty.

D. *Covid-19 Related Health Risks for Others*

Participants' assessments of their own personal risk of dying from Covid-19 likely depend on their beliefs about the relative importance of different risk factors. We assess how participants believe the chances of *dying* from Covid-19 vary for different demographic groups. For the sake of simplicity, respondent time, and statistical power, we focus on three easy-to-describe demographic characteristics: age, race, and gender.

We craft the questions to parallel those from the first block of the survey, assessing Covid-19 death risks for people like the respondents themselves. This time, we ask participants to consider "1,000 people in each of the following [AGE/RACE/GENDER] categories who contract Covid-19 in the next 9 weeks." We ask them, within each category, to assess how many of the 1,000 Americans who contract Covid-19 in the next 9 weeks will pass away due to Covid-19. For the age category, participants make a forecast for 1,000 Americans under 40 years old, for 1,000 Americans between the ages of 40 – 69 years old, and for 1,000 Americans ages 70 and older. For the race category, participants make a forecast for 1,000 white Americans, for 1,000 Black Americans, for 1,000 Asian Americans, and for 1,000 Latinx Americans. For the gender category, participants make a forecast for 1,000 American men and for 1,000 American women.

E. *Sociodemographic Characteristics*

Recall that at the beginning of the survey, all participants are asked to report: year of birth, gender, race (White, Black, Asian, Latinx, check all that apply), approximate annual household income (choose from buckets of $25,000 increments), and region of the country (Northeast, South, Midwest, West). These questions appear as the very first five survey questions, so that Qualtrics can use them as screener questions in order to guarantee a stratified sample.

We also ask non-required sociodemographic questions at the end of the survey: state of residence, whether their current place of residence is best described as urban, suburban, or rural, their educational attainment, whether they have been diagnosed with diabetes, heart disease, lung disease, hypertension, obesity, cancer, or another serious immunocompromising condition, whether they have been hospitalized for non-Covid-19 related reasons within the last year, whether a member of their family has been hospitalized for non-Covid-19 related reasons within the last year, and whether they have been unemployed anytime over the last 9 weeks.

F. *Preferences and Behavior*

Finally, we ask participants about their behavioral responses to the Covid-19 pandemic, and about their preferences regarding policy responses. We ask them how soon they believe "stay at home" measures should be lifted, and whether they would resume their normal activities if stay at home measures were lifted today. We ask about avoidance of medical care, specifically, how reluctant they would be to go to the emergency room today if they or someone in their family had an urgent medical issue, and whether they have avoided filling prescriptions, doctor's appointments, or other forms of medical care in the last few weeks. We then ask them approximately how many times per week over the last few weeks they have left their home to shop, do errands, socialize, etc. (specifically excluding work or exercise). Finally, we ask them, in their opinion, how likely is a significant resurgence of Covid-19 in the fall/winter of 2020.

G. Treatment Assignment and Order

We were also interested in assessing whether the salience of a certain demographic categorization (age, race, or gender) influenced individual perceptions of Covid-19 risks about oneself. For this reason we randomly assigned each participant to one of four treatments that tweaks the order of questions so that the subject is asked to assess Covid-19 risks for certain demographic groups before answering the Covid-19 Related Health Risks for People Like Self.

Specifically, in the control condition the order is exactly as described above, and we randomly assign, at the participant level, the age, race, and gender questions within the Covid-19 Related Health Risks for Others. In the other three treatments, we extract one of the three questions about others – either the age question, the race question, or the gender question – and move it to the front of the survey, immediately preceding the Covid-19 Related Health Risks for People Like Self block. The idea is to prime participants to think about risks in terms of age, race, or gender, before thinking about risks for people like themselves. For participants assigned to one of these three treatments, the remaining 2 questions about others are kept in their original place, in a random order, within the Covid-19 Related Health Risks for Others block later in the survey.

H. Implementation

Qualtrics obtained 1,526 responses to our survey between May 6 and May 13, 2020. Of those 1,526, we drop 4 observations: (i) two of these observations did not provide an answer to our first Covid-19 question asking for beliefs of contracting Covid-19 in the next 9 weeks, and (ii) two of these observations consistently provided answers greater than 1,000 to our questions asking for Covid-19 risk assessments out of 1,000 people.[26] The median time taken to complete our survey is approximately 10.5 minutes.

[26] As part of our IRB approval, respondents were permitted to skip questions. As a result, our number of observations for any particular question is often fewer than our total number of respondents, but typically close to the full sample.

WAVES 2 AND 3 SURVEYS

After analysing the data from our first wave, we conducted two additional waves of our survey. The most significant changes are the inclusion of additional questions, aimed at unpacking the surprising age result, an additional treatment related to question block order, and the addition of an information experiment (only in the Wave 3 survey). We describe these changes below.

Additional Questions

Waves 2 and 3 feature additional questions focused on personal experiences and activities. These questions are placed after the questions that appeared on the original survey, allowing for cleaner comparisons of answers to the original questions across survey waves.[27]

The first additional questions ask about interactions with individuals who might be perceived to be more vulnerable to Covid-19. In particular, we ask whether the individual has at least one young child at home (under 2), has at least one child under 18 at home, has elderly family members at home, or sees parents or other older family members on a regular basis.

We then turn our attention to three factors that we hypothesized might help to explain our age effect. We ask participants their extent of agreement (1 – 7 scale) with three statements: "at this stage in my life, it is possible/realistic to minimize risks," over the course of my life, I've experienced significant adversity," and "I was extremely surprised by the emergence of the Covid-19 pandemic." Following this, we ask specifically about experience with six particular forms of adversity: a serious, life-threatening illness, a serious life-threatening accident or injury, working a job that carries serious health or safety risks, serious illness, injury or untimely death of a loved one, military service, and poverty.

We also ask about personal experiences with Covid-19, asking participants whether they have been infected with Covid-19 (diagnosed by a medical professional), whether they personally know someone who has been infected by Covid-19, and separately, who has been hospitalized due to Covid-19, and separately, who has died due to Covid-19.

We close by asking about political orientation and news sources. Participants are asked to describe their political orientation, choosing from a list ranging from strongly democratic to strongly

[27] The one exception to this is that directly following the question asking how many times per week have you left your home, we add a follow-up questions that asks them specifically about different outside of the home activities (i.e. left home for work, went to a bar, ate indoors at a restaurant, etc.). The only "original" question that appears after this follow-up question is their beliefs about the likelihood of a resurgence.

republican. They are then asked about their frequency of consumption of Covid-19 related information from a variety of sources, as well as their degree of trust in those sources.

New Treatment Variation

In the first wave, we randomized the order in which certain survey blocks appeared. In particular, participants either answered questions about their own Covid-19 related health risks first, or saw one of the three blocks asking them to assess others (by age, race, or gender). In Waves 2 and 3, we introduce a new order variation. In particular, we randomize one-fourth of participants into seeing the block that asks about general health risks before they answer questions about their own Covid-19 related health risks. This allows us to ask how thinking about Covid-19 influences estimates of other health risks. We eliminate the treatment that asks participants to assess Covid-19 risks by gender as the first block, replacing it with this new treatment variation.

Information Experiment

In the third wave of the survey, we introduced an information experiment. This information experiment is placed right before the extended block of demographic and personal experience questions that previously closed the survey. In order to implement the experiment, we moved the question asking participants about their state of residence to the front of the survey (alongside our screening questions). Note that all respondents receive this information experiment.

In this experiment, we ask individuals for their best guess of how many people in their state died from Covid-19 between August 1, 2020 – October 1, 2020. Then, we provide them with truthful information about the number of Covid-19 deaths in their state during that time period (according to the Worldometer Covid-19 data tracker; this source is listed as the source for participants).

We then give participants an opportunity to provide a revised estimate of the Covid-19 hospitalization rate and death rate for Americans like themselves (as asked in the own Covid-19 health risks section of the survey). This allows us to consider reaction to information.

Implementation

Waves 2 and 3 were both implemented in partnership with Qualtrics under the same parameters as Wave 1. Qualtrics was instructed to exclude from participation any individual who had participated in a previous wave of our survey.

Wave 2 was conducted between July 15 – July 22, 2020. We were provided with a total of 1,557 responses. One response was dropped from analysis based upon providing multiple answers that exceeded 1,000 to questions that asked about rates out of 1,000; three responses were dropped from analysis because they skipped several consecutive questions.

Wave 3 was launched on October 30, 2020. Unfortunately, Qualtrics had difficulty fielding our targeted sample size of 1,500 respondents. Recruiting slowed significantly and we decided to close the survey with 1,453 responses on December 13, 2020. We dropped one response from analysis because they skipped several consecutive questions.

SUPPLEMENTARY SIMILARITY SURVEY

In May 2022, we ran a simple additional survey, aimed solely at assessing the subjective similarity of different experiences from our original surveys to a severe Covid outcome. We wanted to understand whether our intuitions about perceived similarity aligned with the views of a large, diverse sample, matched in terms of demographics to our original survey population.

Respondents were provided with a list of eight experiences, each of which was asked about in our original 2020 survey waves. The eight experiences were the two components of our "Health Adversities" index (if the respondent ever suffered a serious, life-threatening accident or injury;if the respondent ever suffered a serious, life-threatening illness), the four components of our "Non health adversities" index (if the respondent worked a job that carried serious health or safety risks; if the respondent experienced military service; if the respondent experienced poverty; if the respondent experienced serious injury, illness, or untimely death of a loved one), and two additional adverse experiences: having experienced a non-Covid hopsitalization and having experienced a family member hospitalization. The listed order of these experiences was randomized at the individual level.

We asked respondents to force rank the eight experiences according to how similar they perceived each to be to a serious Covid outcome in 2020, where 1 indicated most similar and 8 indicated least similar. We randomized respondents into one of three survey options. The first asked the respondent to rank the experiences according to how similar they were to a severe Covid case in 2020. The second asked the respondent to rank the experiences according to how similar they were to a Covid hospitalization in 2020. The third asked the respondent to rank the experiences according to how similar they were to a Covid death in 2020.

In order to enable Qualtrics to field a panel matched on demographics to our previous survey waves, respondents were asked to provide their sex, race/ethnicity, income, region, and age in the first block of the survey. In addition, participants had to indicate that they were willing to provide thoughtful answers in order to proceed.

Implementation

The similarity survey was implemented in partnership with Qualtrics under the same parameters as Waves 1 – 3 of our original survey. Data was collected from 1,046 respondents from May 24 – May 26, 2022. Median completion time for the survey was just over two minutes. We pre-registered the survey using AsPredicted; the pre-registration is available here: https://aspredicted.org/nu8xv.pdf. We pre-registered the plan to report the mean similarity ranks for each of the eight experiences, without updating our specifications for Table 2.

Results

In Table B1, we report the average rank assigned to each experience, alongside the 95% confidence interval, using each of the individual-level observations. The table is sorted according to perceived similarity. Recall that lower numbers indicate greater perceived similarity.

Table B1. Average Subjective Similarity Rank

	Average Rank at Individual Level	95% CI	
Serious Illness	3.26	3.13	3.39
Loss of Loved One	3.42	3.28	3.55
Accident or Injury	3.83	3.71	3.95
Family Hospitalization	4.29	4.17	4.41
Non-Covid Hospitalization	4.43	4.31	4.56
Dangerous Job	4.89	4.76	5.01
Poverty	5.54	5.41	5.67
Military Service	6.35	6.22	6.47

These results are quite similar when broken out separately according to similarity to a severe Covid case, similarity to a Covid hospitalization, or similarity to a Covid death. See Table B2 below.

Table B2. Average Subjective Similarity Rank, split by Type of Covid Experience

	Average Subjective Similarity Rank

	Serious Covid Case	Covid Hospitalization	Covid Death
Serious Illness	3.14	3.46	3.20
Loss of Loved One	3.47	3.50	3.27
Accident or Injury	3.96	3.86	3.67
Family Hospitalization	4.36	4.14	4.36
Non-Covid Hospitalization	4.42	4.37	4.50
Dangerous Job	5.01	4.70	4.95
Poverty	5.36	5.68	5.58
Military Service	6.27	6.29	6.47

Appendix C. Summary Statistics and Robustness on Covid Surveys

In this appendix we present:

1. Summary statistics, correlations, and description of the variables included in our analysis in Section 4;
2. The full version of tables 1 and 2. These include all the controls which were not shown in the main text, and regressions for beliefs on Covid infection and hospitalization.
3. A robustness exercise on interference.

Table C1

Summary statistics. The table describes if the variable was collected in all waves or just in waves 2 and 3 of the survey.

Variable	Waves	Min	Max	Mean	sd
Beliefs others death	All	0	1000	85.64	121.87
Beliefs own death	All	0	1000	53.12	114.78
Age	All	18	116	48.89	18.22
Red hair	All	0	1000	55.64	93.56
State Level	All	7	15669	4750.79	5086.03
Days since Peak	All	1	217	42.1	58
No. health conditions	All	0	7	0.88	0.83
Hospital self	All	0	1	0.1	0.3
Hospital family	All	0	1	0.18	0.38
Had Covid	2 & 3	0	1	0.04	0.2
Health adversities	2 & 3	0	2	0.37	0.56
Non health adversities	2 & 3	0	4	0.9	0.78
Subjective adversity	2 & 3	1	7	4.41	1.64

Table C1 presents summary statistics of our variables. Table C2 presents Pearson's correlation coefficients among them. We now give a fine-grained description of them:

- Beliefs others death is the belief on the number of deaths, out of 1000, conditional on contracting Covid in the next 9 weeks, averaging over estimates for gender groups (males/females), age groups (0-39; 40-69; 70+) and race groups (White; African-American; Asian-American; Latinx-American).
- Beliefs own death is the belief on the number of deaths, out of 1000, for "people like self" conditional on contracting Covid in the next 9 weeks.
- Age is the age of the respondent.
- Red hair is the belief of the respondent on the number of Americans, out of 1000, with red hair.
- Level is the cumulative number of deaths for Covid in the respondent's state, at the time of maximum weekly growth of deaths in the state. Maximum weekly growth is defined as the day with the highest

increase in 7 days rolling average of daily deaths increases, (death number on day t minus death number of day t-7).

- Peak is the number of days since the time of maximum weekly growth of cases in the State, where maximum weekly growth is defined in the same fashion as for deaths.
- Number of health conditions takes values from 0 to 7 and considers: diabetes; heart disease; lung disease; hypertension; obesity, cancer; other serious immunocompromising condition.
- Hospital self is a dummy equal to 1 if the respondent was hospitalized, not for Covid, in the last year.
- Hospital family is a dummy equal to 1 if a family member of the respondent was hospitalized, not for Covid, in the last year.
- Had Covid is a dummy equal to 1 if the respondent has been infected with Covid-19 (diagnosed by a medical professional).
- Health adversities takes values from 0 to 2 and considers if the respondent has personally experienced i) a serious, life-threatening accident or injury; ii) a serious, life-threatening illness.
- Non health adversities takes values from 0 to 4 and considers if the respondent has personally experienced any of the following: i) worked a job that carried serious health or safety risks; ii) serious illness, injury, or untimely death of a loved one; iii) military service; iv) poverty.
- Subjective adversity is the rate of agreement with the statement "Over the course of my life, I've experienced significant adversity". It takes values from 1 (not at all) to 7 (completely agree).

Table C2

Correlations among variables. Green correlation coefficient are significant at 5% level.

	Others death	Age	Red hair	Level	Days	Health cond	Hosp self	Hosp fam	Had Covid	Health adv	Non h adv	Subj adv
Beliefs others death	0.56	-0.28	0.18	0.09	0.02	-0.02	0.12	0.12	0.13	0.06	-0.05	0.11
Beliefs others death		-0.15	0.18	0.05	0.01	0.06	0.11	0.08	0.1	0.08	0	0.1
Age			-0.09	-0.2	-0.14	0.26	-0.14	-0.23	-0.11	0.06	0.09	-0.14
Red hair				0.05	0.05	0	0.05	0.03	0.03	0.02	-0.03	0.07
State Level					0.66	0	0.15	0.17	0.03	-0.02	-0.08	0.09
Days since Peak						0.03	0.14	0.15	0.03	0	-0.04	0.08
No. health conditions							0.11	0.06	0.06	0.28	0.19	0.13
Hosp self								0.39	0.13	0.17	0.01	0.13
Hosp fam									0.09	0.11	0.06	0.13
Had Covid										0.13	-0.02	0.09

Health adversities											0.07	0.21
Non health adversities												0.19

Table C3 presents the full list of selected regressors for beliefs about Covid fatality for others, for self, the risk of own infection and of own hospitalization. The risk of own infection is the belief on the number of Covid infections, out of 1000, for "people like self" in the next 9 weeks. Own hospitalization is the belief on the number of Covid hospitalizations, out of 1000, for "people like self" conditional on contracting Covid in the next 9 weeks. We can see that all the results regarding fatality also hold for infections and hospitalization.

Table C3

Own death is the belief on the number of deaths, out of 1000, for "people like self" conditional on contracting Covid in the next 9 weeks. Others death is the belief on the number of deaths, out of 1000, conditional on contracting Covid in the next 9 weeks, averaging over estimates for gender groups (males/females), age groups (0-39; 40-69; 70+) and race groups (White; African-American; Asian-American; Latinx-American). Own infection is the belief on the number of Covid infections, out of 1000, for "people like self" in the next 9 weeks. Own hosp is the belief on the number of Covid hospitalizations, out of 1000, for "people like self" conditional on contracting Covid in the next 9 weeks. All variables are standardized except for dummy variables (Hosp self; Hosp fam; Black; Asian; Rural). Red hair is the belief of the respondent on the percentage of red-haired Americans. Level is the cumulative number of deaths for Covid in the state, at the time of maximum weekly growth in the state. Days is the number of days since the peak of cases in the state. No. of health conditions takes values from 0 to 7 and considers: diabetes; heart disease; lung disease; hypertension; obesity, cancer; other serious immunocompromising condition. Hosp self (fam) is a dummy equal to 1 if the respondent (a family member) was hospitalized, not for Covid, in the last year. Income is the income of the respondent. Rural, Asian, and Black are dummies referring to the residential area or ethnicity of the respondent.

	Dependent variable:			
	Own death	Others death	Own infection	Own hosp
	(1)	(2)	(3)	(4)
Age	-0.131***	-0.236***	-0.183***	-0.112***
	(0.019)	(0.015)	(0.013)	(0.015)
Red hair	0.163***	0.155***	0.171***	0.130***
	(0.032)	(0.019)	(0.029)	(0.026)
State Level	0.037**	0.073***	0.071***	0.077***
	(0.015)	(0.014)	(0.014)	(0.017)
Days since Peak	-0.057***	-0.084***	-0.088***	-0.083***
	(0.013)	(0.015)	(0.012)	(0.018)
No. health conditions	0.090***	0.032***	0.027**	0.039***
	(0.015)	(0.011)	(0.013)	(0.013)
Hosp (self.)	0.245***	0.231***		0.319***
	(0.078)	(0.062)		(0.065)
Hosp (fam.)		0.093***	0.156***	0.099***
		(0.036)	(0.048)	(0.038)
Income	-0.036**	-0.044***	-0.083***	-0.043**

	(0.016)	(0.016)	(0.013)	(0.019)
Black	0.111**	0.164***		0.084**
	(0.053)	(0.048)		(0.042)
Asian		0.205***		
		(0.060)		
Rural	0.123***	0.068**		0.064*
	(0.033)	(0.030)		(0.035)
Constant	-0.084***	-0.103***	-0.027*	-0.086***
	(0.022)	(0.022)	(0.014)	(0.018)
Observations	4,514	4,477	4,506	4,511
R^2	0.073	0.122	0.081	0.063
Adjusted R^2	0.071	0.120	0.080	0.060
Note:				*p<0.1; **p<0.05; ***p<0.01 Clustered standard errors at state level

Table C4 presents the full output of Table 1 in column 1. Column 2 shows that our results, that higher non health adversities lead to lower pessimism, hold if we omit "serious injury, illness or untimely death of a loved one" from non-health adversities. Column 3 shows that political affiliation does not predict beliefs, nor does it affect the role of the other regressors.

Table C4

Others death is the belief on the number of deaths, out of 1000, conditional on contracting Covid in the next 9 weeks, averaging over estimates for gender/age/race groups. More precisely, a first estimate is obtained averaging over beliefs for males and females; a second estimate is obtained averaging over beliefs for three age groups (0-39; 40-69; 70+); a third estimate is obtained averaging over beliefs for four race groups (White; African-American; Asian-American; Latinx-American). The final estimate is obtained averaging these three estimates. All variables, but dummies, are standardized. Health adversities is an index given by the sum of two dummies indicating 1) if the respondent ever suffered a serious, life-threatening accident or injury; 2) if the respondent ever suffered a serious, life-threatening illness. Non health adversities is an index given by the sum of four dummies: indicating 1) if the respondent worked a job that carried serious health or safety risks; 2) if the respondent experienced military service; 3) if the respondent experienced poverty; 4) if the respondent experienced serious injury, illness, or untimely death of a loved one. Non health adversities (small) does not consider the fourth one. Subjective adversity is the rate of agreement with the sentence "Over the course of my life, I've experienced significant adversity." Republican degree is a variable which measures political orientation of the respondent and it takes values from 1 (Strongly Democratic) to 7 (Strongly Republican).

	Dependent variable:		
	Others death		
	(1)	(2)	(3)
Had Covid	0.441***	0.446***	0.442***
	(0.167)	(0.167)	(0.167)
Health adversities	0.047**	0.046**	0.044**
	(0.019)	(0.019)	(0.019)
Non health adv.	-0.039***		-0.037**
	(0.015)		(0.015)

Non health adv. (small)		-0.031*	
		(0.016)	
Subj. adversity	0.043**	0.041**	0.043**
	(0.019)	(0.019)	(0.020)
No. health cond.	0.012	0.010	0.014
	(0.017)	(0.016)	(0.016)
Hosp (self.)	0.157**	0.160**	0.158**
	(0.073)	(0.073)	(0.073)
Hosp (fam.)	0.058	0.050	0.055
	(0.044)	(0.044)	(0.045)
State Level	0.059***	0.061***	0.059**
	(0.023)	(0.023)	(0.023)
Days since Peak	-0.097***	-0.097***	-0.096***
	(0.023)	(0.023)	(0.023)
Red hair	0.165***	0.166***	0.166***
	(0.033)	(0.033)	(0.032)
Age	-0.212***	-0.216***	-0.211***
	(0.021)	(0.020)	(0.023)
How Republican			-0.022
			(0.026)
Income	-0.043*	-0.042*	-0.039*
	(0.023)	(0.022)	(0.024)
Black	0.133**	0.136**	0.119**
	(0.054)	(0.054)	(0.053)
Asian	0.249***	0.252***	0.249***
	(0.092)	(0.091)	(0.089)
Rural	0.113**	0.116***	0.101**
	(0.044)	(0.044)	(0.044)
Constant	-0.128***	-0.129***	-0.123***
	(0.030)	(0.029)	(0.027)
Observations	2,953	2,953	2,944
Adjusted R^2	0.133	0.132	0.133
Note:	*p<0.01;**p<0.05;***p<0.01		
	Clustered standard errors at state level		

Beliefs, political affiliation, and behavior. Existing work on the pandemic has stressed the importance of political beliefs in shaping behaviour (e.g. wearing a mask) and policy views. Do memory-based beliefs about the lethality of Covid, which are only modestly influenced by politics, affect behaviour?

Our survey measured behaviour and attitudes, including how often respondents leave home for reasons other than work or exercise, whether they have recently forfeited medical care to avoid leaving home, and whether they favour lifting the lockdown measures in place at the time of the survey. Of course, past experiences may affect behaviour through a variety of channels. For instance, respondents with past health adversities may refrain from going out because it is harder for them to do so, not necessarily because they are more pessimistic about Covid. To address this issue, we use the "red hair" proxy as an instrument for beliefs. The idea is that "red hair" captures respondent's general tendency to overestimate unlikely events, regardless of whether they concern risk or not. As a result, if "red hair" helps explain behaviour, it arguably does so via beliefs.[28]

Table C5 reports our regressions. In columns (1), (3) and (5) we show the role of beliefs in OLS specifications in which we control for the best predictors of behaviour selected by our method. In columns (2), (4) and (6) we instrument beliefs using the red hair proxy. Relative to Table 1, we add political affiliation ('how republican') which, while not selected as a predictor of beliefs, is a commonly cited predictor of attitudes towards the pandemic (Bursztyn et al 2020).

Respondents who estimate higher "red hair", and hence have more pessimistic beliefs about Covid, behave more cautiously. Interference in retrieval affects beliefs and, through this channel, memory affects behaviour. This only occurs, however, for individual decisions, not for a policy preference such as whether to lift the lockdown. Political affiliation instead emerges as a key predictor for policy preferences, consistent with existing work.

Table C5

Going out is the answer to the question "Over the last few weeks, approximately how many times per week have you left your home to shop, do errands, socialize, etc.?". It takes values 1 (never), 2 (once a week), 3 (twice a week), 4 (three or more times a week). Med avoid is the answer to the question "Have you avoided filling prescriptions at the pharmacy, doctor's appointments, or other forms of medical care in the last few weeks?". It takes values 1 (Yes, completely), 2 (Somewhat), 3 (Not at all). Lift lockdown is the answer to the question "Would you resume your normal activities if lockdown or "stay-at-home" measures were lifted today?". It takes value from 1 (Definitely yes) to 5 (Definitely not). Death others is the belief on Covid death for others, as described in tables 1 and 2. It is obtained as the average of the estimated risk of death for separate age, ethnicity and gender classes. This is instrumented with the estimated number of red-haired Americans (F >> 10 in all cases). Republican degree is a variable which measures political orientation of the respondent and it takes values from 1 (Strongly Democratic) to 7 (Strongly Republican). All variables are standardized and controls include variable which were selected by performing a dependent variable specific model selection algorithm. Max weekly growth death is the maximum weekly growth of Covid deaths in the state. Days since weekly death peak is the number of days since Covid deaths peak in the state. Current level death is the current cumulative level of Covid

[28] Red hair also has a low correlation with the other predictors of beliefs. It has a -0.09 correlation with "Age". The next variable in the survey whose correlation with red hair is highest in magnitude is "Subjective Adversities" which has a 0.07 correlation with red hair.

deaths in the state. Unemployment is a dummy equal to 1 if the respondent experienced unemployment in the last nine weeks.

	Dependent variable:					
	Going out	Going out	Med avoid	Med avoid	Lift Lockdown	Lift Lockdown
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Death others	-0.071***	-0.228**	-0.057**	-0.278**	-0.002	-0.119
	(0.023)	(0.112)	(0.023)	(0.114)	(0.019)	(0.098)
Max weekly growth death	-0.057***	-0.055***				
	(0.014)	(0.014)				
Days since wk death peak	0.044*	0.036				
	(0.023)	(0.023)				
Current level death			-0.019	-0.028		
			(0.023)	(0.020)		
Age	0.065***	0.023	0.227***	0.169***		
	(0.018)	(0.031)	(0.016)	(0.031)		
Age squared			0.065***	0.076***		
			(0.016)	(0.015)		
Female	-0.051***	-0.049**			0.113***	0.115***
	(0.019)	(0.020)			(0.020)	(0.021)
Black					0.026	0.034*
					(0.018)	(0.019)
Asian	-0.071***	-0.062***			0.056***	0.066***
	(0.019)	(0.017)			(0.014)	(0.018)
Rural			-0.102***	-0.089***		
			(0.020)	(0.019)		
Education			-0.092***	-0.093***		
			(0.017)	(0.019)		
West					0.025	0.022
					(0.023)	(0.024)
Suburban					0.083***	0.072***
					(0.016)	(0.017)

Income					-0.092***	-0.091***
					(0.017)	(0.018)
No. health conditions	-0.083***	-0.076***	-0.084***	-0.076***	0.056**	0.056**
	(0.020)	(0.018)	(0.020)	(0.022)	(0.022)	(0.022)
Hosp (fam)	0.056***	0.064***				
	(0.016)	(0.017)				
Hosp (self)			-0.082***	-0.067***		
			(0.020)	(0.025)		
Unemployment			-0.032*	-0.028		
			(0.019)	(0.018)		
State population	-0.038**	-0.035**	-0.034**	-0.026*	-0.079**	-0.064
	(0.017)	(0.015)	(0.014)	(0.016)	(0.038)	(0.042)
Republican degree	0.090***	0.083***	0.012	0.003	-0.261***	-0.267***
	(0.022)	(0.023)	(0.018)	(0.025)	(0.043)	(0.047)
Constant	0.115***		-0.042*		0.082***	
	(0.018)		(0.024)		(0.021)	
Controls	YES	YES	YES	YES	YES	YES
Observations	2,962		2,960		2,963	
R^2	0.043		0.141		0.122	
Adjusted R^2	0.039		0.138		0.119	
Note:						*p<0.1;**p<0.05;***p<0.01
						Clustered standard errors at state level

Table C6 presents a more complete analysis of interference. It reports the coefficient of the interaction among all Covid and non-Covid adversities. We also report the coefficient of the interaction of a variable with itself, obtained by adding the square of that variable to the corresponding regression. For the sake of clarity and brevity, health adversities include serious injury, serious illness, and hospital self. Hence, it is defined from 0 to 3, differently from Table 2. Green indicates agreement with our theory, yellow disagreement. A darker color corresponds to a lower p-value. We can see that, consistent with Figure 4, interference is present across the board, with the strongest ones being among i) Level and family hospital; ii) health conditions and family hospital. The square of the number of health conditions has a strong and negative coefficient, meaning that numerous health conditions interfere one with the other in shaping pessimism.

Table C6

Each cell reports the interaction estimated between the row and the column, together with their p values in parentheses. A green cell indicates that the sign of the coefficient directionally matches the prediction of the theory, a yellow cell indicates that it does not. Darker colors indicate lower p value. Interactions were estimated adding them to the model presented in table 1 column 2, if the two variables were available in all waves. They were estimated adding them to the model presented in table 2 column 2, if at least one of the two variables was available only in waves 2 and 3. The interaction of a variable with itself represents the coefficient of the square of the variable. Health adversities takes values from 0 to 3 and it includes serious injury, serious illness, and own hospital.

Others Death	Level	Health cond	Family hosp	Health adv	Had Covid
Level	-0.009 (0.399)	-0.007 (0.572)	-0.072 (0.000)	-0.032 (0.061)	-0.153 (0.052)
Health conditions		-0.011 (0.006)	-0.112 (0.000)	-0.015 (0.298)	-0.077 (0.459)
Family hospital				-0.013 (0.762)	-0.132 (0.714)
Health adversities				-0.007 (0.660)	0.022 (0.875)

Figure C1 extends the analysis of Figure 3 by examining interference between non-Covid health adversities and the experience of having had Covid. As in Figure 3, having had Covid reduces the marginal impact of non-Covid health experiences, and vice versa, except for the index of health adversities (serious illness or injury, or own hospitalization).

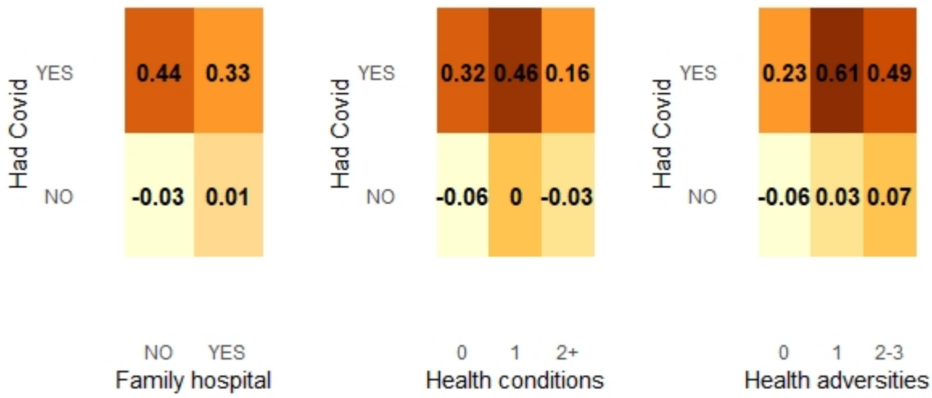

Figure C1.

The Figure reports the residuals of the standardized beliefs of FATALITY (for others), estimated by removing from the model of Table 2's Column 2 the variables "Had Covid" and i) "Family Hospitalization", ii) "Number of Health Conditions", and iii) "Health Adversities". Health adversities refer to the sum of serious injury, serious illness, and self hospitalization dummies. Reported values are average residuals in each cell. Different colours indicate different average residuals up to the third decimal.

Appendix D. Model Selection

The regressions presented in the main text show output models obtained from best subset selection. In our survey, we collect several demographics and ask several behavioral questions, along with beliefs about Covid. This is a typical case where we might want to remove irrelevant predictors. There are two compelling reasons to do that: i) when the number of predictors is high, prediction accuracy of the OLS model will be good but there might be a lot of variability in the least squares fit; ii) interpretability of models which include a lot of predictors is difficult. It is often the case that some or many of the variables used in a multiple regression model are in fact not associated with the response. Including such irrelevant variables leads to unnecessary complexity in the resulting model. By removing these variables—that is, by setting the corresponding coefficient estimates to zero—we can obtain a model that is more easily interpreted. Although in our case the number of observations is much higher than the number of potential covariates (hence variability should not be an issue), we still aim at keeping only the most relevant predictors. To do so, we employ a machine learning algorithm called best subset selection (Guyon and Elisseeff, 2003; James et al., 2013). Other applications of best subset selections in economics include Alabrese and Fetzer (2018) and Becker et al. (2017). The method works as follows: we fit a separate least squares regression for each possible combination of the *p* predictors. That is, we fit all *p* models that contain exactly one predictor, all $\binom{p}{2}$ models that contain exactly two predictors, and so forth. We then look at all of the resulting models, with the goal of identifying the one that is best, according to some information criteria. More formally, the algorithm entails the following steps:

1) We denote $\mathcal{M}_0$ the *null model*, containing no covariates;
2) For k $\in \{1,2, \dots, \mathrm{p}\}$ we:
   a) Fit all $\binom{p}{k}$ models containing k covariates;
   b) Pick the best of these $\binom{p}{k}$ models and denote it $\mathcal{M}_{k}$. The best model is the one with the highest R^2. In every set of models with k covariates, we can compare them by using the R^2, since the number of covariates is fixed within the set;
3) Select the best model, among $\mathcal{M}_0, \dots, \mathcal{M}_p$ using cross-validation or an information criterion (Mallow's C_p, BIC, adjusted R^2).

We can express the best subset selection problem as a nonconvex and combinatorial optimization problem. The objective is to find the optimal s for:

$$\min_{\beta} \sum_{i}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\,\beta_j \right)^2 \text{ subject to } \sum_{j=1}^{p} I\big(\beta_j \neq 0\big) \leq s$$

This requires that the optimal solution involves finding a vector β such that the residual sum of squares is minimized and no more than s coefficients are different from 0. The algorithm presented above (points 1-3) solves this optimization problem for every value of s and then picks among the optimal models for the different values of s. Best subset selection can thus be expressed as a regularized regression with penalization term equal to $\sum_{j=1}^{p} I\big(\beta_j \neq 0\big)$.

In point 3 of our description of the algorithm, we refer to the selection of the best model, among $\mathcal{M}_0, \dots, \mathcal{M}_p$. We will discuss three information criteria: Mallow's C_p, Bayesian information criterion (BIC), and adjusted R^2. Mallow's C_p is defined as $C_p = \frac{1}{n}(RSS + 2d\hat{\sigma}^2)$, with RSS being the residual sum of squares, d the total number of parameters used and $\hat{\sigma}^2$ is an estimate of the variance of the error ϵ associated with each response measurement. In the case of the linear model with Gaussian errors, C_p is equivalent to the Akaike information criterion (AIC). BIC is defined as $BIC = \frac{1}{n}(RSS + log(n)\, d\hat{\sigma}^2)$. The BIC replaces $2d\hat{\sigma}^2$ with $log(n)\, d\hat{\sigma}^2$. Since, $log(n) > 2$ if $n > 7$, the BIC places a heavier penalty on models with many variables and it usually selects smaller models than the C_p. As can be easily guessed, to identify the best model we aim at minimizing either the Mallow's C_p or the BIC. The adjusted R^2 is defined as adj$R^2 = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)}$ where TSS is the total sum of squares. The best model is the one which maximizes the adjusted R^2. Finally, we can use m-fold cross-validation. This proceeds as follows: i) divide the sample of n observation in into m non-overlapping groups (folds), each containing around $\frac{n}{m}$ observations; ii) for each $z \in \{1,2, \dots, m\}$ treat fold z as a validation set, fit the model on the remaining folds and compute the mean squared error, MSE_z pertaining to the withheld validation set z; iii) compute $CV_m = \frac{1}{m}\sum_{z=1}^{m} MSE_z$. We will then choose the model with the lowest cross-validation error. What is the best criterion to use is an issue which goes beyond the scope of this discussion. We can refer the reader to Ding et al. (2018). To give a sense of this discussion, in figure A1 we show a comparison of the four decision criteria, applied to the choice of the best model to predict the number of times the respondent had gone out in the period before the survey (table 3 column 1).

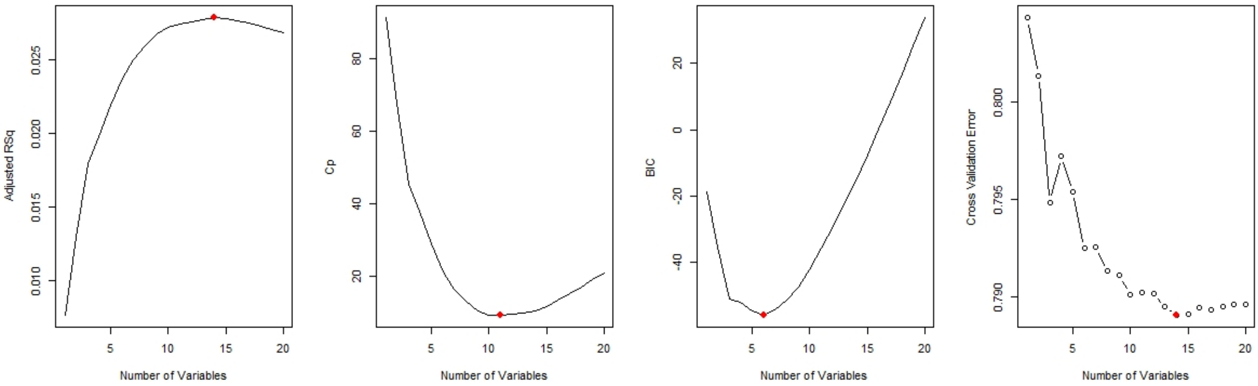

Figure A1

Adjusted R^2, Mallow's C_p, BIC and cross-validation error to select the best model to describe the propensity to go out. The best model, according to each criterion, is highlighted in red.

The set of potential predictors is the set of demographics and we can see that the BIC selects the regression with 6 covariates, namely age, dummy for female, dummy for Asian, Number of health conditions, family member been hospitalized (not for covid), and population of the state, which we included as controls in table 3.[29] Figure A1 offers the perfect insight to reflect on the different information criteria. BIC suggests that the best model is the one with 6 covariates. We have already explained why the BIC tends to select more parsimonious models. In this case both the adj. R^2 and cross-validation suggest to use a 14 covariates model and Mallow's C_p suggests to include 11 covariates. However, we can see that the 6 variable model is very close to the best model for each of the four criteria. This was the principle which guided us in our work. We usually selected the best model, according to the BIC criterion, and verified if this was close to be optimal for the other three.

We now give some more details on how we selected the best model for each of our dependent variables. Tables 1 and 2 report the output of the models we selected to describe beliefs about Covid death. A similar procedure is employed to describe beliefs about Covid infection and hospitalization. We split the variables in 3 sets:

1) Set A: state level Covid dynamics. For all the three waves it contains the following variables (for Covid cases or deaths): current level; maximum weekly growth; days since growth peak; current weekly growth; level at the time of maximum growth;

[29] Table A4 reports also variables on Covid dynamics, which were the object of a separate variable selection and politics, which was added for theoretical reasons.

2) Set B: personal characteristics and Covid experiences. For all the three waves it contains the following variables: age, gender, ethnicity, region, income, urbanization, employment, a lot of health info on the self and family, state population, the estimated number of red haired Americans;
3) Set B': these are additional variables in waves 2 and 3: interactions with family members, several measures of adversities in life, several measures of direct and indirect exposure to Covid; political preferences; several opinions on Covid.

One caveat with best subset selection is that certain variables may be dropped in case they are highly correlated with each other. This is why, in some cases we perform some minimal form of supervision, like for example retaining some predictors which are very relevant according to our memory model, but were not selected by the machine learning algorithm.[30]

Our model selection consists of the following stages:

1) We perform model selection, for each of the 4 dependent variables (Covid infection, hospitalization, and death for self, Covid death for others), in set A of state level Covid dynamics (10 predictors);
2) We perform some minimal supervision on model selection. We select the model that contains the most robust predictors across the four types of beliefs. This leads to the inclusion of the days since the weekly cases growth peak, and the level of cases in the state of the respondent at the time of maximum weekly growth of cases;[31]
3) We perform model selection, for each of the 4 dependent variables, in set B and B' of demographics (23 predictors for all waves; 35 predictors for waves 2 and 3);
4) We show the resulting models which contain the variables selected in stages 1-3 in table 1;
5) Table 2 column 2 contains the best model obtained when performing model selection in waves 2 and 3, plus all the covariates which were selected on all waves (table 1 column 2), even if they were excluded by performing model selection in the last two waves.

A similar procedure is employed to select the best subset of predictors from set B to predict the number of times the respondent had gone out, the tendency to avoid medical appointments, and the

---

[30] For example, health adversities and non health adversities. Each of them had been considered separate potential predictors and serious injury only had been selected. We decided to include them jointly as indices.
[31] To give a sense of how our mild supervision worked, best subset selection suggested those two predictors for all but one dependent variable. For beliefs about infection, the best model would have included the maximum weekly growth of cases in the state, instead of the level. The model we picked had negligible differences with the “optimal” one, in terms of prediction accuracy.

preference for lifting lockdown. These are included in table 3. We included political orientation as a control in table 3, since this is believed to be a relevant factor in orienting behavior and policy preference regarding "stay-at-home" measures.

References for Appendix D

Alabrese, E. and Fetzer, T. Who is Not Voting for Brexit Anymore? (2018). CESifo Working Paper No. 7389, Available at SSRN: https://ssrn.com/abstract=3338743

Becker, S. O., Fetzer, T., and Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. Economic Policy, 32(92): 601-650

Ding, J., Tarokh, V., and Yang, Y. (2018). Model selection techniques: An overview. *IEEE Signal Processing Magazine*, *35*(6), 16-34.

Guyon, I. and A. Elisseeff (2003). An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3, 1157–1182

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An introduction to statistical learning. New York: Springer

Appendix E. The Primed-Recall Experiment

The following pages present the primed-recall experimental materials.

What is the purpose of this research?
This study aims to understand how past experiences inform beliefs about the future.

What can I expect if I take part in this research?
If you take part in this study, you will complete a 15-minute survey. You will be asked questions about your past experiences and about your expectations about the likelihood of different events in the future. You will be asked to provide brief written answers about these events. You will also be asked some basic questions about yourself. If you complete the survey, you will be paid $3.25. This payment will be made via the Prolific platform within 48 hours of your participation.

Please be aware that there are comprehension checks that you must answer correctly in order to complete the study. Failure to answer these comprehension checks correctly could result in dismissal from the study without pay.

While we believe the risks of this study are minimal, you should be aware that some of the experiences we ask you about may be unpleasant to think about. You may choose to skip questions about past experiences that you are uncomfortable answering.

You will not interact with any other participants in this study.

Your answers will be linked to your Prolific ID at the time of data collection. The research team will remove your Prolific ID from the dataset before analyzing the data or sharing aggregate results.

Please do not provide any other information in your answers to this survey that would allow for others to identify you – this includes names, zip codes, contact information, and other personal details. The research team will attempt to remove any individually identifiable information from your responses before analyzing or sharing your data. This de-identified data will be stored for future research use and may be shared with other researchers or research participants.

What should I know about a research study?
- Whether or not you take part is up to you.
- Your participation is completely voluntary.
- You can choose not to take part.
- You can agree to take part and later change your mind.
- Your decision will not be held against you.
- Your refusal to participate will not result in any consequences or any loss of benefits that you are otherwise entitled to receive.
- You can ask all the questions you want before you decide.

You may not be told everything
As part of this research design, you may not be told everything about the purpose or procedures of this research. There are different versions of this study. While you will be fully informed about the version of this study that you have been assigned to, you will not be informed about different versions of this study that other participants are in.

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the

research team: Katherine Coffman, kcoffman@hbs.edu, 617 495 6538.

Thank you for participating today.
We care about the quality of our survey data and hope to receive the most accurate measures of your beliefs. It is important to us that you provide thoughtful, careful answers to this survey.
Do you commit to providing your thoughtful and careful answers to the questions in this survey?

- I commit to providing thoughtful and careful answers
- I do not commit to providing thoughtful and careful answers
- I can't promise either way

Some of the questions in this survey will ask you to make your best estimate as to the likelihood of different things. To give you some practice and get you used to thinking in these terms, we have a few example questions for you to work through. Note that you are required to answer the following questions in order to proceed.

Example 1: According to the United States Census, **approximately 2 out of every 100 Americans** live in Massachusetts. This is equivalent to approximately 2%.

Using this estimate, tell us **how many out of 5,000 randomly-chosen Americans** live in Massachusetts?

Example 2: Think about **a group of 100 randomly-chosen American cities**. How many of them do you believe receive more than 1ft of snow in a typical winter?

When asked, "Out of 100 randomly-chosen American cities, how many receive more than 1ft of snow in a typical winter?", you entered that you believe ${q://QID49/ChoiceTextEntryValue/1} out of 100 randomly-chosen American cities receive more than 1ft of snow in a typical winter.

This is equivalent to X%.

Now, using the same estimate you provided above, please tell us, out of **1,000** randomly-chosen American cities, how many receive more than 1ft of snow in a typical winter?

Great! From this point forward, you may choose to skip questions that you feel uncomfortable answering. Please click next to get started on the survey.

[Note: Individuals are randomly-assigned to see one of the three 'experience prompts' below, or randomly-assigned to a control treatment that proceeds directly to the next section.]

EXPERIENCE PROMPTS
In this part of the survey, we will ask you about an experience you may have had. We appreciate you taking the time to provide thoughtful, honest answers to these questions. An important reminder before you start:
Please do NOT provide any information in your answers to these questions that would allow others to identify you – this includes, for example, names, zip codes, or contact information.

[IDTHEFT]
Have you, a loved one, or your employer dealt with identity theft, a data breach, stolen credit card information, or a compromised password in the past?
Yes
No → skip to next section

Thinking back on this experience, please respond to the questions below.

- What were the intentions of the perpetrators? If you do not know, please give us your best guess.
- Please briefly describe any damages, losses, or inconvenience you, your loved one, or your employer suffered as a result of this experience.
- What emotions do you associate with thinking back on that experience?

[FINANCE]
Have you or a loved one ever struggled with finances?
Yes
No → skip to next section

- What was the main cause of that struggle?
- What losses, difficulties, or sacrifices do you remember from that time?
- What emotions do you associate with thinking back on that experience?

[LOSS]
Have you lost a loved one to illness?
Yes
No → skip to next section

- Thinking back on this experience, please respond to the questions below.
- Briefly - using no names or other personally identifiable information - describe who the loved one was and what the illness was.
- What losses, difficulties, or sacrifices do you remember from that time?

[For all Experience Prompts]

- Please take the next few minutes to type 3 – 4 sentences about this personal experience. Feel free to write whatever comes to mind as you think about this experience.

- On a scale of 1 – 7, where 1 is not at all and 7 is extremely vividly, how vividly do you remember this personal experience?

  ______

[LIKELIHOOD ASSESSMENTS – Everyone sees these]

Please take a moment to read the following information about cyberattacks. This information is adapted from IBM’s web resources.
Please take a moment to read the following information about cyberattacks. This information is adapted from IBM’s web resources.

Cyberattacks are unwelcome attempts to steal, expose, alter, disable or destroy information through unauthorized access to computer systems. They can be launched by individual actors, like hackers, or by criminal organizations or state actors. While not always the case, cyberattacks can be associated with cyber warfare or cyberterrorism.

Cyber attacks target things like financial data, personal or sensitive data, intellectual property, infrastructure, or government departments and agencies. Not all cyberattacks are successful. Many organizations and governments have implemented plans to limit the possibility of disruption or harm. But, if successful, cyberattacks can cause significant damage - data loss or manipulation, major service disruptions, and financial losses.

We are interested in your beliefs about the chances of a significant cyberattack in the future. In particular, please consider the questions below.

What do you believe is the likelihood that **you will be significantly impacted** by a cyberattack over the next 5 years? This impact could come in the form of financial loss or through damage to critical infrastructure, such as power lines, hospitals, banking systems, communication satellites, or manufacturing.

Please indicate on a scale of 0 - 100, where **0 indicates that there is no chance** at all that you will be significantly impacted by a cyberattack and **100 indicates that you will definitely be significantly impacted**.

______

Now, think of **1,000 people just like you** in the United States.

Out of those 1,000 people, how many do you believe will be significantly impacted by a cyberattack over the next 5 years? Again, this impact could come in the form of financial loss or through damage to critical infrastructure, such as power lines, hospitals, banking systems, communication satellites, or manufacturing.

When answering the previous question, how vividly did you imagine what an infrastructure-disrupting cyber attack would be like?

Please answer on a scale of 1 – 7 where 1 is not vividly at all and 7 is extremely vividly.

Please take a few minutes now to write down 3 – 4 sentences about any thoughts that came to mind when you were answering the previous questions on cyberattacks.

For each of the experiences below, please indicate whether you have had this past experience.

	Yes (1)	**No** (0)
Have you, a loved one, or your employer dealt with identity theft, a data breach, stolen credit card information, or a compromised password in the past? (Had_ID)	○	○
Do you have memories of the Sept. 11th, 2001 terrorist attacks in the United States? (Had_Sept11)	○	○
Have you lost a loved one to serious illness?	○	○
Have you or a loved one experienced a recent extreme weather event in your area, such as a blizzard, hurricane, tornado, or flood?	○	○
Have you or a loved one ever struggled with finances?	○	○
Have you or a loved one been hospitalized recently?	○	○
	○	○
Have you or a loved one dealt with addiction?	○	○
Have you or a loved one had a serious accident or injury recently?	○	○

Page Break

For each of the experiences below, please indicate **how similar you believe that experience is to a significant cyberattack -** in particular, a cyberattack that significantly disrupts critical civilian infrastructure, such as power lines, hospitals, banking systems, communication satellites, or manufacturing.

Please answer on a scale of 1 – 7 where **1 is Not at all similar** and **7 is Extremely similar**.

	1 2 3 4 5 6 7
Identity theft, a data breach, or having had credit card information or a password stolen ()	
The Sept. 11th terrorist attacks ()	
Losing a loved one to serious illness ()	
An extreme weather event such as a blizzard, hurricane, tornado, or flood ()	
Struggling with finances ()	
Having been hospitalized recently ()	
Dealing with addiction ()	
A serious accident or injury ()	

[DEMOGRAPHICS]

What is your year of birth?
What is your sex?

- Male
- Female
- Non-binary
- Prefer not to say

Choose one or more races that you consider yourself to be:

- White
- Latino or Latina
- Black or African American
- Asian
- Other
- Prefer not to say

Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that includes your entire household income in (previous year) before taxes.

- Less than $25,000
- $25,000 to $49,999
- $50,000 to $74,999
- $75,000 to $99,999
- $100,000 to $124,999
- $125,000 to $150,000
- $150,000 or more

What region best describes where you live?

- Northeast
- South
- Midwest
- West

- Other

What is the highest level of school you have completed or the highest degree you have received?

- ○ Less than high school degree
- ○ High school graduate (high school diploma or equivalent including GED)
- ○ Some college but no degree
- ○ Associate degree in college (2-year)
- ○ Bachelor's degree in college (4-year)
- ○ Master's degree
- ○ Doctoral or Professional degree (PhD, JD, MD)

In a typical year, approximately how much snow falls in the area you live in? We encourage you to use the internet to look up the answer for your area.

Appendix F. Supplemental Results for Primed Recall Experiment

Table F1.
The Impact of Primed and Lived Experiences on Cyberattack Estimates, ITT

	OLS Predicting Index of Cyberattack Estimates High Quality Only		
	1	2	3
ID theft prime	0.15**	0.15**	0.12*
	(0.061)	(0.060)	(0.072)
Finance prime	0.14**	0.15**	0.19**
	(0.062)	(0.060)	(0.077)
Loss prime	0.057	0.054	0.14*
	(0.061)	(0.060)	(0.078)
Below Median Similarity x *ID theft* prime			0.058
			(0.085)
Below Median Similarity x *Finance* prime			-0.085
			(0.086)
Below Median Similarity x *Loss* prime			-0.15*
			(0.086)
Had *ID theft*		0.46***	0.46***
		(0.050)	(0.050)
Had *finance*		0.22***	0.21***
		(0.064)	(0.064)
Had *loss*		0.092*	0.086*
		(0.047)	(0.047)
Snow	0.16***	0.17***	0.17***
	(0.023)	(0.022)	(0.022)
Controls	Y	Y	Y
Observations	2090	2083	2083
Adjusted R-squared	0.049	0.102	0.103

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year. The *prime* indicators take 1 if the individual was randomly-assigned to that treatment. The *had* indicators take 1 if the individual reported having had that personal experience. *Below median similarity* is an indicator that takes 1 if the individual reported a below-median similarity assessment of the primed experience compared to others in the same treatment.

Table F2.
The Impact of Primed and Lived Experiences on Cyberattack Estimates,
No High Quality Restriction

	OLS Predicting Index of Cyberattack Estimates		
	ITT	TOT	
	1	2	3
ID theft prime	0.18***	0.26***	0.24***
	(0.051)	(0.058)	(0.072)
Finance prime	0.14***	0.14***	0.26***
	(0.050)	(0.053)	(0.066)
Loss prime	0.12**	0.17***	0.33***
	(0.050)	(0.058)	(0.073)
Below Median Similarity x *ID theft* prime			0.044
			(0.087)
Below Median Similarity x *Finance* prime			-0.23***
			(0.076)
Below Median Similarity x *Loss* prime			-0.29***
			(0.087)
Had *ID theft*		0.36***	0.36***
		(0.050)	(0.050)
Had *finance*		0.17**	0.17**
		(0.070)	(0.069)
Had *loss*		0.071	0.068
		(0.049)	(0.049)
Snow	0.16***	0.18***	0.18***
	(0.018)	(0.020)	(0.020)
Controls	Y	Y	Y
Observations	2976	2424	2424
Adjusted R-squared	0.052	0.090	0.096

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year. The *prime* indicators take 1 if the individual was randomly-assigned to that treatment. The *had* indicators take 1 if the individual reported having had that personal experience. *Below median similarity* is an indicator that takes 1 if the individual reported a below-median similarity assessment of the primed experience compared to others in the same treatment.

Table F3.
The Impact of Primed and Lived Experiences, Individual Cyberattack Estimates

	OLS Predicting Cyberattack Estimate for Self (z-score)			OLS Predicting Cyberattack Estimate for People Like You (z-score)		
	High Quality Only					
	ITT	TOT		ITT	TOT	
	1	2	3	4	5	6
ID theft prime	0.22***	0.32***	0.31***	0.058	0.092	0.054
	(0.061)	(0.068)	(0.085)	(0.060)	(0.068)	(0.085)
Finance prime	0.19***	0.18***	0.31***	0.079	0.051	0.092
	(0.061)	(0.065)	(0.082)	(0.061)	(0.065)	(0.082)
Loss prime	0.11*	0.15**	0.29***	0.00048	0.038	0.11
	(0.061)	(0.070)	(0.091)	(0.060)	(0.070)	(0.091)
Below Median Similarity x			0.027			0.077
ID theft prime			(0.10)			(0.10)
Below Median Similarity x			-0.23**			-0.074
Finance prime			(0.093)			(0.093)
Below Median Similarity x			-0.24**			-0.14
Loss prime			(0.11)			(0.11)
Had *ID theft*		0.34***	0.34***		0.38***	0.38***
		(0.063)	(0.063)		(0.063)	(0.063)
Had *finance*		0.25***	0.25***		0.28***	0.28***
		(0.083)	(0.083)		(0.083)	(0.083)
Had *loss*		0.050	0.045		0.0052	0.0035
		(0.059)	(0.059)		(0.059)	(0.059)
Snow	0.11***	0.13***	0.13***	0.19***	0.21***	0.21***
	(0.022)	(0.025)	(0.025)	(0.022)	(0.025)	(0.025)
Controls	Y	Y	Y	Y	Y	Y
Observations	2097	1709	1709	2091	1704	1704
Adjusted R-squared	0.047	0.083	0.088	0.045	0.079	0.079

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year. The *prime* indicators take 1 if the individual was randomly-assigned to that treatment. The *had* indicators take 1 if the individual reported having had that personal experience. *Below median similarity* is an indicator that takes 1 if the individual reported a below-median similarity assessment of the primed experience compared to others in the same treatment.

Table F4.
Similarity and Vividness

	OLS Predicting **Index of Cyberattack Estimates**		OLS Predicting **Reported Vividness** of Cyberattack	
	High Quality Only, Treatment-on-Treated, Omitting Control Group			
	1	2	3	4
Perceived Similarity of Primed Experience	0.056***	-0.034	0.060***	-0.015
	(0.013)	(0.051)	(0.012)	(0.046)
Reported Vividness of Primed Experience	0.062***	0.00061	0.19***	0.14***
	(0.021)	(0.039)	(0.019)	(0.036)
Similarity x Vividness		0.016*		0.014*
		(0.0087)		(0.0080)
Snow	0.17***	0.17***	0.034	0.031
	(0.031)	(0.031)	(0.029)	(0.029)
Controls	Y	Y	Y	Y
Observations	1172	1172	1161	1161
Adjusted R-squared	0.060	0.062	0.114	0.115

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. The reported vividness of cyberattack measure in Columns 3 and 4 is the z-score of the individual's self-reported vividness with which they imagined a cyberattack. The reported vividness of primed experience measure is the individual's self-reported vividness with which they imagined the primed experience on a 1 – 7 scale. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table F5.
Simulation and Vividness
OLS Predicting **Index of Cyberattack Estimates**
High Quality Only

Reported Vividness of Cyberattack	0.28***
	(0.022)
Snow	0.15***
	(0.022)
Controls	Y
Observations	2014
Adjusted R-squared	0.118

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. The reported vividness of cyberattack measure is the z-score of the individual's self-reported vividness with which they imagined a cyberattack. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table F6.
Similarity and Responsiveness to Experiences, ITT

OLS Predicting Index of Cyberattack Estimates
High Quality Only, Intent-to-Treat

	Pooled		Below Median *Snow*		Above Median *Snow*	
	1	2	3	4	5	6
$\bar{S}_i(E_i)$, Total Similarity of Lived Experiences	0.19***	0.19***	0.15***	0.15***	0.23***	0.24***
	(0.025)	(0.025)	(0.035)	(0.035)	(0.037)	(0.037)
$S_i(e_p)$, Similarity of Primed Experience	0.14***	0.20***	0.16***	0.19***	0.13***	0.21***
	(0.025)	(0.029)	(0.034)	(0.040)	(0.036)	(0.042)
$\bar{S}_i(E_i)\ x\ S_i(e_p)$		-0.061***		-0.039		-0.084***
		(0.018)		(0.026)		(0.024)
Snow	0.16***	0.16***	0.36***	0.36***	0.087**	0.099**
	(0.022)	(0.022)	(0.080)	(0.080)	(0.044)	(0.044)
Controls	Y	Y	Y	Y	Y	Y
Observations	2088	2088	1061	1061	1027	1027
Adjusted R-squared	0.114	0.119	0.086	0.087	0.117	0.127

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. $\bar{S}_i(E_i)$ is the z-score of the average perceived similarity of all lived, non-primed experiences; an individual who reports 0 lived experiences has a pre-standardized $\bar{S}_i(E_i)$ of 0. $S_i(e_p)$ is the z-score of the perceived similarity of the primed experience; an unprimed individual has a pre-standardized $S_i(e_p)$ of 0. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table F7.
Similarity and Responsiveness to Experiences, No Restriction to High Quality

	OLS Predicting Index of Cyberattack Estimates					
	Treatment-on-Treated					
	Pooled		Below Median *Snow*		Above Median *Snow*	
	1	2	3	4	5	6
$\bar{S}_i(E_i)$, Total Similarity of Lived Experiences	0.17***	0.18***	0.13***	0.14***	0.22***	0.23***
	(0.023)	(0.023)	(0.031)	(0.031)	(0.033)	(0.033)
$S_i(e_p)$, Similarity of Primed Experience	0.13***	0.21***	0.16***	0.22***	0.11***	0.20***
	(0.022)	(0.027)	(0.031)	(0.038)	(0.033)	(0.040)
$\bar{S}_i(E_i)\ x\ S_i(e_p)$		-0.069***		-0.056**		-0.082***
		(0.016)		(0.023)		(0.022)
Snow	0.18***	0.18***	0.29***	0.30***	0.12***	0.12***
	(0.020)	(0.020)	(0.073)	(0.073)	(0.040)	(0.040)
Controls	Y	Y	Y	Y	Y	Y
Observations	2427	2427	1267	1267	1160	1160
Adjusted R-squared	0.112	0.119	0.085	0.088	0.100	0.111

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. The index of the cyberattack estimates is constructed by taking the z-score of each cyberattack estimate for the individual, averaging them, and then computing the z-score of the averaged measure. $\bar{S}_i(E_i)$ is the z-score of the average perceived similarity of all lived, non-primed experiences; an individual who reports 0 lived experiences has a pre-standardized $\bar{S}_i(E_i)$ of 0. $S_i(e_p)$ is the z-score of the perceived similarity of the primed experience; an unprimed individual has a pre-standardized $S_i(e_p)$ of 0. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table F8.
Similarity and Responsiveness to Experiences, Individual Outcome Measures

	OLS Predicting Z-Score of Cyberattack Estimate for Self High Quality Only, Treatment-on-Treated					
	Pooled		Below Median *Snow*		Above Median *Snow*	
	1	2	3	4	5	6
$\bar{S}_i(E_i)$, Total Similarity of Lived Experiences	0.18***	0.18***	0.14***	0.14***	0.21***	0.23***
	(0.028)	(0.028)	(0.039)	(0.039)	(0.039)	(0.040)
$S_i(e_p)$, Similarity of Primed Experience	0.17***	0.22***	0.21***	0.24***	0.14***	0.21***
	(0.027)	(0.033)	(0.038)	(0.047)	(0.039)	(0.048)
$\bar{S}_i(E_i)\ x\ S_i(e_p)$		-0.051***		-0.029		-0.071***
		(0.020)		(0.030)		(0.027)
Snow	0.13***	0.13***	0.22**	0.22**	0.045	0.050
	(0.025)	(0.025)	(0.089)	(0.089)	(0.049)	(0.049)
Controls	Y	Y	Y	Y	Y	Y
Observations	1712	1712	873	873	839	839
Adjusted R-squared	0.111	0.114	0.093	0.093	0.108	0.115

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. $\bar{S}_i(E_i)$ is the z-score of the average perceived similarity of all lived, non-primed experiences; an individual who reports 0 lived experiences has a pre-standardized $\bar{S}_i(E_i)$ of 0. $S_i(e_p)$ is the z-score of the perceived similarity of the primed experience; an unprimed individual has a pre-standardized $S_i(e_p)$ of 0. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table F8b. Similarity and Responsiveness to Experiences, Individual Outcome Measures

	OLS Predicting Z-Score of Cyberattack Estimate for People Like You High Quality Only, Treatment-on-Treated					
	Pooled		Below Median *Snow*		Above Median *Snow*	
	1	2	3	4	5	6
$\bar{S}_i(E_i)$, Total Similarity of Lived Experiences	0.17***	0.18***	0.11***	0.12***	0.22***	0.23***
	(0.028)	(0.028)	(0.039)	(0.039)	(0.040)	(0.040)
$S_i(e_p)$, Similarity of Primed Experience	0.070**	0.12***	0.093**	0.13***	0.052	0.13***
	(0.028)	(0.034)	(0.038)	(0.047)	(0.040)	(0.049)
$\bar{S}_i(E_i)\ x\ S_i(e_p)$		-0.057***		-0.043		-0.076***
		(0.020)		(0.030)		(0.027)
Snow	0.20***	0.20***	0.43***	0.43***	0.12**	0.12**
	(0.025)	(0.025)	(0.089)	(0.089)	(0.050)	(0.050)
Controls	Y	Y	Y	Y	Y	Y
Observations	1707	1707	869	869	838	838
Adjusted R-squared	0.083	0.087	0.050	0.052	0.076	0.084

Notes: * denotes p<0.1, ** p<0.05, *** p<0.01. Controls include age, sex, race/ethnicity, education, income, region. $\bar{S}_i(E_i)$ is the z-score of the average perceived similarity of all lived, non-primed experiences; an individual who reports 0 lived experiences has a pre-standardized $\bar{S}_i(E_i)$ of 0. $S_i(e_p)$ is the z-score of the perceived similarity of the primed experience; an unprimed individual has a pre-standardized $S_i(e_p)$ of 0. *Snow* is the z-score of the individual's estimated share of U.S. cities receiving more than 1ft of snow in a typical year.

Table 9. The Role of Similarity in Experience Effects

	I	II
ID Theft Prime	0.15**	0.13**
	(0.060)	(0.059)
Financial Prime	0.15**	0.14**
	(0.060)	(0.060)
Lost Loved One Prime	0.054	0.035
	(0.060)	(0.059)
Had ID Theft	0.46***	0.13
	(0.050)	(0.12)
Had Financial Struggles	0.22***	0.064
	(0.064)	(0.077)
Had Loss of Loved One	0.092*	-0.063
	(0.047)	(0.058)
Weighted ID Theft		0.053***

		(0.017)
Weighted Financial Struggles		0.048***
		(0.014)
Weighted Loss of Loved One		0.072***
		(0.017)
Snow	0.17***	0.16***
	(0.022)	(0.022)
Controls	Y	Y
Observations	2083	2080
Adjusted R-squared	0.102	0.124