Knowledge Acquisition in a High-Stakes Environment: Evidence from the Covid-19 Pandemic*

Charles Angelucci†, Amit K. Khandelwal‡, Andrea Prat‡, and Ashley Swanson‡

*MIT Sloan, †Columbia GSB

August 2021

Abstract

We measure Covid-19 knowledge at the onset of the pandemic. Information acquisition models predict that the precision of information increases in the benefits of improved decision-making and decreases in the costs of acquiring information. Indeed, we find that individuals with lower information costs are more informed, and that the benefits of information (proxied by mortality risk) predict information acquisition. High-risk individuals' knowledge is initially lower, but converges, albeit slowly, to that of low-risk individuals within 43 days. These findings highlight the importance of information cost heterogeneity and suggest a role for policies that increase information access for high-cost, high-benefit groups.

JEL codes: D83, I12

Keywords: information acquisition, health, Covid-19

*We thank Sylvain Chassang, Mark Dean, Andrey Simonov, Mirko Wiederholt, and seminar participants at MIT Sloan for useful suggestions and comments. We are grateful to Sergio Andres Florez Orrego, Jett Pettus, and Szymon Sacher for excellent research assistance. We thank the Jerome A. Chazen Institute for Global Business at Columbia GSB for funding. We obtained approval from the Columbia University IRB under protocol AAAS3426.

1 Introduction

Agents often behave suboptimally in complex and high-stakes environments, and lack of information is an important mechanism underlying this phenomenon.[1] Models of costly information acquisition predict that the precision of information is increasing in the benefits of improved decision-making and decreasing in the costs of acquiring and processing information. However, there is little empirical evidence on the processes underlying gaps in information, particularly when the stakes are high. This paper studies the onset of Covid-19 to assess whether canonical information models can rationalize how individuals acquire information in a sudden public health crisis that requires immediate behavior change.

We exploit data from surveys administered to representative samples of U.S. residents through the polling company YouGov. The first survey was launched one week after the U.S. declared Covid-19 a national emergency on March 13, 2020 and the last survey took place at the end of April 2020. Each survey round included a series of questions designed to elicit respondents' knowledge about Covid-19 and a series of questions about social distancing behavior.

Measuring knowledge about Covid-19 at the onset of the pandemic is important because the spread of an epidemic depends on how people behave, which depends in turn on their information.[2] Also, we can study knowledge about Covid-19 to learn about information acquisition processes more generally for several reasons. First, information about Covid-19 was novel. Few individuals in the U.S. were informed about Covid-19 even a few weeks prior to our first survey. Differences in knowledge thus reflect differences in the attention individuals recently paid to their environment, rather than pre-existing knowledge differences. Second, information about Covid-19 has instrumental value: it helps individuals adapt their behavior and reduce their risk of infection. In models of costly information acquisition, this implies that the precision of an individual's information should be increasing in her risk of harm from Covid-19.[3] Lastly, the information that we focus on was well-defined, uncontroversial, and publicly available in mainstream media and government publications.[4]

Our analysis is guided by a simple Rational Inattention framework, borrowed from Mackowiak et al. (2018). This approach requires us to measure individuals' costs and benefits of acquiring information. We proxy individuals' benefit with their age, gender, and race/ethnicity-specific mor-

[1]Such choice environments are diverse, encompassing the selection of insurance plans (Abaluck and Gruber 2011, Brown and Goolsbee 2002); the takeup of federal benefits (Bettinger et al. 2012, Bhargava and Manoli 2015); the choice of schools (Hastings and Weinstein 2008); and the purchase of automobiles and appliances (Davis and Metcalf 2016, Scott Morton et al. 2001).

[2]A recent literature formalizes this dependency with epidemiological–economic models of disease dynamics (Fenichel et al. 2011, Philipson 2000). These models have an explicit role for information regarding the risks individuals face. See Faia et al. (2021) regarding the type of information people acquired in the early stages of the pandemic. See Allcott et al. (2020) regarding partisan differences in social distancing behavior and beliefs.

[3]By contrast, quantifying an individual's benefit from acquiring information would be difficult if we were studying information about, say, politics.

[4]Our knowledge questions originally included a question about whether hand-washing or mask wearing was recommended to prevent transmission of Covid-19. Given the CDC's reversal on this policy during our study, we dropped the question CDC (2020f,h).

tality risk, constructed using data from the Centers for Disease Control and Prevention (CDC). We exploit the fact that the variation in mortality risk across demographic groups was salient from the onset of the pandemic–as keywords from newspaper articles at that time confirm–to test whether individuals with higher mortality risk acquired more information.[5] To proxy individuals' cost of acquiring information, we leverage data from Angelucci and Prat (2021) who, starting in December 2018, have run regular surveys through YouGov to measure general knowledge of mainstream political news (the "AP surveys"). We use 2018-2019 AP surveys to train a machine learning model of news knowledge on a rich set of socioeconomic characteristics that overlap between the AP survey data and ours. We use the model to predict our survey respondents' knowledge of the news, whose inverse we use to proxy their costs of acquiring information.

Our main findings are as follows. Consistent with theory, individuals with higher costs of acquiring information exhibited significantly lower levels of knowledge: a standard deviation increase in our proxy for information costs was associated with a 0.22-0.24 standard deviation decrease in our index measuring knowledge about Covid-19. This first result suggests that steady-state knowledge of the news is a good predictor of individuals' short-run acquisition of suddenly relevant information. Next, we show that individuals who were at greater mortality risk from Covid-19 started the pandemic with lower levels of knowledge: a standard deviation increase in our mortality risk index was associated with a 0.05-0.14 standard deviation lower knowledge at our first survey round, fielded March 19-23.[6] However, individuals who were at greater risk largely caught up with the rest of the population by our fourth survey round in April 22-30; across specifications, the initial knowledge deficit among high-risk individuals was halved to fully eliminated by the end of our study. To investigate whether this catch-up was specific to knowledge about Covid-19, we exploit concurrent AP surveys on knowledge about political news *unrelated* to the pandemic. We indeed find that individuals with greater mortality risk did not exhibit increasing knowledge about non-Covid topics, suggesting that mortality risk drove the acquisition of information about Covid-19 specifically. We discuss the policy implications of our findings as well as caveats in Section 5.

This paper contributes to an empirical literature that studies information acquisition through the lens of the rational inattention framework–see, e.g., Cavallo et al. (2017) on information about inflation, Lacetera et al. (2012) on used-car purchases, Kacperczyk et al. (2016) on portfolio investments, Bartos et al. (2016) on house rental markets, Fuster et al. (forthcoming) on expectations regarding national home prices, and Bhattacharya and Howard (Forthcoming) about strategic choices in baseball–and more broadly as a function of the stakes involved in the decision at hand (Brown and Jeon 2020,Morrison and Taubinsky 2021).[7] This paper makes two contributions. First,

[5]Bundorf et al. (2021) find that Hispanic and Black Americans were aware of their greater risk of Covid-19 infection and severe health consequences, and older individuals were aware of the worse health outcomes conditional on infection.

[6]Alsan et al. (2020) measure knowledge about Covid-19 using a survey at the onset of the pandemic and find differences in knowledge along demographic dimensions, with Black Americans, men, and individuals aged 55 or less exhibiting lower levels of knowledge.

[7]This paper also relates to a literature on individuals' expectations about macroeconomic variables. See, for instance, Coibion and Gorodnichenko (2012), Coibion and Gorodnichenko (2015), and Gaglianone et al. (2020).

Covid-19 provides a unique opportunity to study information aquisition, as it was a completely new health crisis where individuals had no initial information and rapid behavior change was needed. Second, our study includes direct measurements of individual knowledge, as well as an individual-level proxy for the cost of aquiring information. As a result, the paper highlights the importance of information cost heterogeneity in the speed of acquisition of crucial information.

2 Theory

Rational Inattention is a natural framework to analyze knowledge about Covid-19. We interpret the model presented in Section 2.1 of Mackowiak et al. (2018) in the context of our application. Consider an individual living through a pandemic: she decides how much information to gather about the disease and what precautionary measures to take. Specifically, individual i faces an unknown state of the world $x \sim N(0,1)$ and must choose her behavior y_i. Suppose her payoff is $U_i(x, y_i) = B_i x y_i - \frac{1}{2} C_i y_i^2$ and think of y_i as social distancing. The parameter x represents the benefit of that behavior in terms of reduced infection risk. Parameters B_i and C_i are instead individual-specific and they represent the benefit of reducing the infection risk and the intrinsic cost of the preventive behavior. We refer to $b_i \equiv \frac{B_i}{C_i}$ as the relative benefit of the behavior.

In a first stage, the individual acquires signal $s = x + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\epsilon^2)$, by choosing σ_ϵ^2. More precise signals are more costly and the information cost is given by $\lambda_i \kappa$, where λ_i is an individual-specific cost parameter and κ is the amount of information collected. The individual learns about the disease by consulting governmental sources, using the media, or talking to experts. The cost λ_i may be lower for certain individuals than for others. For instance, someone who regularly reads a newspaper is likely to receive information about Covid-19 with limited additional cost. Instead, someone who rarely follows the news may have a higher λ_i.

The amount of information κ is measured by the expected reduction in uncertainty, which is proportional to the conditional variance of the signal. An individual who acquires signal s pays an information cost proportional to:

$$\mathrm{Var}(x) - \mathbb{E}_s[\mathrm{Var}(x|s)] = 1 - \frac{\sigma_\epsilon^2}{1+\sigma_\epsilon^2} = \frac{1}{1+\sigma_\epsilon^2}.$$

In a second stage, the individual observes s and chooses the optimal behavior y_i. Mackowiak et al. (2018) show that the solution can be expressed in terms of the chosen level of attention:

$$\xi_i = 1 - \frac{\mathbb{E}_s[\mathrm{Var}(x|s)]}{\mathrm{Var}(x)} = \max\left(0, 1 - \frac{\sigma_\epsilon^2}{1+\sigma_\epsilon^2}\right),$$

a variable with values between 0 (no attention) and 1 (perfect knowledge). The optimal behavior is $y_i = b_i \mathbb{E}(x|s) = b_i \xi_i s_i$ and the expected behavior is an increasing function of the true state, where the strength of that relationship depends on ξ_i:

$$\mathbb{E}[y \mid x] = b_i \xi_i x. \tag{1}$$

If a behavior is beneficial ($x > 0$), better informed people do more of it. Anticipating her behavior, the optimal amount of information to collect is:

$$\xi_i = \max\left(0, 1 - \frac{\lambda_i}{2b_i^2}\right). \tag{2}$$

The agent's information precision depends positively on the relative benefit b_i and negatively on the information cost λ_i.

To summarize, the model makes two predictions: (i) the agent's information level ξ_i depends positively on her relative benefit b_i and negatively on her information cost λ_i and (ii) the agent's expected behavior $\mathbb{E}[y_i]$ is increasing in her level of information ξ_i and in her relative benefit b_i.

3 Surveys and Measurement

This section describes the survey methodology and measurement of key variables.

Surveys

We conducted four surveys through YouGov in spring 2020 (the "Covid surveys"). In addition to the information collected through the surveys, YouGov collects background information on survey respondents' socioeconomic status, education, political leanings, religiosity, etc.

The first survey round took place March 19-23, roughly a week after the U.S. declared Covid-19 a national emergency on March 13. The remaining three rounds took place between March 26-31, April 8-13, and April 22-30.[8]

The first survey was administered to 2,000 individuals. The next two surveys were each administered to 1,000 individuals who did not participate in a previous survey (the cross-section sample) and to a subset of individuals who participated in a prior survey (the panel sample). The final survey was administered to 1,500 cross-section respondents and to 500 panel respondents.[9] Our survey took roughly 16-18 minutes to complete. Participants received $1.5 on average (paid via gift cards). Payments included a $1 participation fee and a $1 bonus if they accurately completed a knowledge quiz (discussed below). To supplement YouGov's attribute data, we collected additional respondent characteristics, including recent as well as pre-pandemic news consumption habits, flags for the presence of children and elderly individuals in the household, and flu vaccine takeup.

We also take advantage of the AP surveys, which were conducted on a different sample of YouGov respondents during 2018-2020. As described below, Angelucci and Prat (2021) developed a methodology to measure individuals' information about political news that we adapted for our sample. We use their 2018-19 surveys to predict Covid survey respondents' information costs (λ_i), and their 2020 surveys are used in placebo regressions to assess how general political news knowledge

[8] No subsequent survey was administered due to funding constraints.

[9] For details on how YouGov constructs each sample of new respondents, see https://smpa.gwu.edu/sites/g/files/zaxdzs2046/f/downloads/YG_Matching_and_weighting_basic_description.pdf.

(unrelated to Covid-19) evolved in spring 2020. Figure A.1 reports the timeline of the Covid and AP surveys against the spread of Covid-19 in the U.S.

Knowledge Index

In each survey round, we asked a series of questions to assess respondents' knowledge of Covid-19. The questions were designed to assess knowledge about the properties of the disease, and their general level of attention to the pandemic, covering topics such as symptoms, the consequences of contracting the virus, the means through which it spreads, where and when it had spread, etc. We also quizzed respondents on political news related to Covid-19. Finally, we asked respondents to guess the number of confirmed cases and deaths in their states.

We aimed to capture respondents' knowledge of publicly-available information regarding Covid-19. An important feature of this environment is that it was truly new, hence the commonly-used term "novel coronavirus." This is advantageous for a test of costly information acquisition since the information at issue would not have been widely disseminated even six weeks prior to our first survey.[10] Relatedly, in many cases, the answers to our knowledge questions could not easily have been guessed based solely on knowledge of earlier viral outbreaks such as the 1918 influenza epidemic (which was particularly deadly for young adults, unlike Covid-19), the first SARS outbreak of 2003 (which did not originate in Wuhan, China), the 2009 H1N1 pandemic (which had a greater impact on children than adults), or the most recent outbreak of Ebola in 2014-2016 (Ebola does not involve respiratory symptoms) (Lovelace 2020, da Costa et al. 2020). This suggests that respondents who correctly answered our questions were actively following the Covid-19 pandemic. Of course, accurate information about Covid-19 properties has been far from perfect throughout the pandemic. Our guiding principle was to formulate questions based on publicly posted/reported information, most often available concurrently at the CDC and the World Health Organization (WHO).

The first set of knowledge questions measured respondents' factual knowledge on Covid-19. The *same* factual knowledge questions were asked to the cross-section sample of (new) respondents in each survey round in order to track changes in a fixed measure of knowledge over time. The questions are stated in Part A of Table B.1 (rows 1-6). The panel sample respondents were asked new questions in each survey round. These questions are stated in Part C of Table B.1 (rows 9-30).

The second set of knowledge questions asked respondents to guess the total number of confirmed Covid-19 cases and deaths in their state, as of the day before the survey date. For these questions, we code the response as correct if the numeric response was within one standard deviation on either side of the truth.[11] The wording of these questions is reported in Part B of Table B.1 (rows 7-8).

[10] See Figure A.2, which reports the weekly count of unique news articles mentioning Covid-19 between October 2019 and May 31, 2020. To give a sense of scale, the analogous frequency is shown for a similarly high-profile topic (President Donald Trump). There were few articles about Covid-19 until mid-January, around the time of the first case in the U.S. In mid-February, Covid-19 was as frequently covered as President Trump. After mid-February, articles about Covid-19 dwarfed articles about him.

[11] The standard deviation was taken across the true cumulative case or death count across all respondents in each round, excluding outliers (above the 95th percentile).

The third set of knowledge questions covered recent political developments related to the pandemic. The methodology we employed to construct the quiz follows Angelucci and Prat (2021). Respondents were given 60 seconds to take a quiz which included three true and three fake political news statements about Covid-19. Respondents were told that exactly three statements were true and they received a $1 bonus if they selected all three true statements. We relied on a panel of three journalists to select the three political news statements they felt were the most important in the week before the start of every survey round. We relied on the same journalists to write three false but plausible news stories that could have happened during the same period.[12] The same news knowledge quizzes were administered to cross-section and panel respondents within each survey round. New quizzes were used in each survey round. These news quiz statements are reported in Table B.2.

Our overall knowledge index, k_{it}, is constructed as simply the fraction of answers that are correct.

In one set of analyses, we explore how non-Covid-related knowledge evolved in the early weeks of the pandemic, as a placebo exercise. We analyze general news quiz questions from three AP surveys that took place in February, April, and May 2020. The general knowledge index is the fraction of statements selected correctly, k_{it}^{AP}. These news quizzes are reported in Part B of Table B.3. For the sake of comparability, we create a "Covid-19 knowledge quiz" index k_{it}^{q}, which is the fraction of Covid news quiz answers that are correct.

Mortality Risk Index

To proxy for each respondent's private benefit of reducing infection risk, we construct a measure of relative Covid-19 mortality risk as a function of age, gender, and race/ethnicity. While we find it unrealistic to suppose that sample respondents went to the CDC website to analyze Covid-19 mortality statistics, the patterns of mortality across age, gender, and race/ethnicity were widely discussed in the media in the early weeks of the pandemic.[13] Figure 1 reports the percent of news articles about Covid-19 mortality risk that mentioned age, gender, race/ethnicity, and pre-existing conditions in each week of 2020, beginning with the date of the first confirmed case of Covid-19 in the U.S. In the week following January 19, 29% of articles on Covid-19 mortality risk mentioned age, 30% mentioned gender, and 10% mentioned race/ethnicity. After January, age was uniformly the most common risk factor mentioned. Gender was initially more prominent than race/ethnicity, while race/ethnicity was more frequently mentioned than gender after early April.

We define mortality risk using the rate of Covid-19 confirmed deaths per population as of May 5 (CDC 2020i). A high index of deaths per population may reflect a high infection risk, a high

[12] The panel of journalists selected the most important news stories related to Covid-19 from Reuters wire stories dedicated to U.S. politics (www.reuters.com/news/archive/politicsNews); see Angelucci and Prat (2021) for details.

[13] On the relationship between socioeconomic factors and infection as well as death rates, see Brown and Ravallion (2020) and Knittel and Ozaltun (2020). Fan et al. (2020) look at the relationship between socioeconomic factors and partisanship and individuals' beliefs about health outcomes. Finally, using survey data from eight countries, Galasso et al. (2020) document differences between men and women in beliefs and attitudes.

infection fatality rate, or both; we normalize mortality risk by population rather than by confirmed cases in order to avoid bias driven by disparities in testing or access to health care. We measure how mortality risk varies along key demographic dimensions: age-by-gender (ag) and race/ethnicity (r). Age is specified in ten-year bins, with separate bins for under five years and over 74 years of age. For each demographic grouping $j \in \{ag, r\}$ and each cell within that grouping, we construct a relative mortality index $RMI_c^j = \left(\frac{deaths_c^j}{population_c^j}\right) / \left(\frac{deaths_{total}}{population_{total}}\right)$.[14] The full details are shown in Table A.1. For example, $RMI_{>75,m}^{ag} = 10.992$ and $RMI_{Black}^{r} = 1.752$ indicate that males aged 75 and over and Black individuals had mortality rates (relative to population) approximately 11 and 1.8 times that of the average person in the U.S., respectively. The index is steeply increasing in age, is uniformly higher for males than females, and is lower for White people than for all other race/ethnicity groups except American Indian/Alaska Native. To obtain an overall relative mortality index for each survey respondent, we then take the product of the mortality risk across the age-gender and race/ethnicity dimensions: $RMI_i = RMI_{c(i,ag)}^{ag} * RMI_{c(i,r)}^{r}$. This index is unable to account for, for instance, different age distributions or different age-mortality profiles by race/ethnicity, but reflects the best available public-facing evidence on mortality risk variation across demographic groups as of early May 2020.[15]

Information Costs

A key determinant of knowledge in the model is an individual's information cost, λ_i. While we did not survey our respondents before March 2020, Angelucci and Prat (2021) conducted regular surveys on different samples of respondents in 2018-20 to measure knowledge of U.S. national politics. As described above, the main instrument through which knowledge is measured are news quizzes designed by journalists in which respondents are incentivized to select three true news stories from a list containing three true and three false stories. We use their 2018-19 surveys (reported in Part A of Table B.3) to train a random forest model of knowledge of news about political events (as measured by the number of correct answers to a quiz) on the rich set of demographic and other characteristics, including religion, partisanship, and media diet, that overlap between the AP surveys and the Covid surveys.[16] In a regression of the actual number of correct answers on the random-forest-predicted number of correct answers, the out-of-sample $R^2 = 0.085$.[17] Using the

[14] Bundorf et al. (2021) document that while overall individuals tended to overestimate their risk of Covid-19 infection and the health implications of infection, they tend to correctly assess their *relative* risks.

[15] Subsequent data that accounts for different age distributions across racial/ethnic groups has found that, relative to White people, excess all-cause mortality in spring 2020 was 77% higher for Asian people, 154% higher for American Indian people, 177% higher for Hispanic people, and 340% higher for Black people (Polyakova et al. 2021). The independent contributions of the underlying causes of disparities in the mortality impact of Covid-19 are not yet fully known. Thus, it is unknown how survey respondents' own circumstances would affect their interpretation of the early Covid-19 mortality data.

[16] The number of trees in the random forest model was chosen to maximize out-of-sample fit in a cross-validation exercise, where the model was trained on 90 percent of the data and then R^2 was calculated in the remaining 10 percent.

[17] The seemingly low R^2 is explained by the nature of the prediction being made as well as patterns in the raw data. First, the number of correct answers is a variable taking four values (i.e., $\{0, 1, 2, 3\}$). Second, the average respondent in Angelucci and Prat (2021) selects about 2.3 correct answers, with very few respondents selecting 0 or

model structure and estimated parameters, we can construct $\widetilde{\lambda}_i$ as a proxy for information costs, for each respondent in our sample. Specifically, $\tilde{\lambda}_i$ equals -1 times the predicted number of correct answers from the random forest model, so that a higher index indicates greater predicted difficulty acquiring information.

Behavior Index

We also surveyed respondents on their social distancing practices. Specifically, we asked a series of 17 questions regarding recent behavior relevant for an individual's exposure to the virus (see Table B.4). These questions elicited respondents' number of outings, gatherings, trips to stores to make essential and non-essential purchases, etc. This information was collected on both a "past 24 hours" and "past 7 days" basis. We form the behavior index beh_{it} by constructing an individual's deviation from perfect compliance with social distancing guidelines issued by the CDC. For behaviors classified by the CDC as desirable in spring 2020 (e.g., hand-washing), we leave the answers coded as is for binary answers, and winsorize and standardize continuous answers. For behaviors classified as undesirable (e.g., participating in gatherings with 50 or more people), we winsorize and standardize continuous answers, and multiply all answers by -1 so that an increase in our behavior index reflects greater compliance. Winsorization is applied at the 99^{th} percentile for desirable behaviors, and at the 1^{st} percentile for undesirable behaviors.

4 Results

As described above, our survey took place over four rounds between March 19 and April 30, 2020. Our regressions below include both the "panel sample" (p) of survey respondents who participated in multiple rounds and answered different questions, and also the "cross-section sample" (xs) of respondents per round, who each participated in a single round and all of whom answered the exact same questions on the Covid-19 pandemic.

Building on the simple model of costly information acquisition outlined above, we estimate the following specification:

$$k_{it} = \alpha + \beta_1 RMI_i + \beta_2 Days_{it} + \beta_3 RMI_i * Days_{it} + \widetilde{\lambda}_i + X_{it}\beta + \varepsilon_{it} \tag{3}$$

where i indexes an individual and t indexes the day on which she responded to the survey; k_{it} is our knowledge index, proxying here for information ξ_{it}; RMI_i is our relative mortality index, proxying here for relative benefit b_i; and $\widetilde{\lambda}_i$ is our random forest-predicted proxy for information costs, based on the news quiz in Angelucci and Prat (2021). As described above, $Days_{it}$ is the days since March 19, normalized by the total time horizon of our surveys (43 days). Thus, our trend varies from zero (beginning of survey) to one (end of survey). It also varies by i based on the date of her survey response. This approach allows us to flexibly capture the differential knowledge

1 correct statements.

available to late versus early respondents within each survey round. Finally, X_{it} includes controls intended to capture variation across individuals that may be correlated with both their levels of and trends in Covid-19 mortality. Summary statistics for all control variables used in our analysis are presented in Table A.2. Also included in X_{it} are dummy variables for each combination of sample and round, to account for variation in difficulty across question sets, and a set of fixed effects as specified in the notes to each table and figure below. To aid in interpretation, the knowledge index k_{it} is standardized within each survey round-by-sample (xs/p), to account for the different set of questions asked to each group over time. The information acquisition cost proxy $\tilde{\lambda}_i$, and the behavior index beh_{it} and media consumption variables in alternative specifications below, are standardized within the full ($xs + p$) sample.

Table 1 presents our baseline results. In columns 1-2, we proxy for information costs using the Angelucci and Prat (2021) index $\tilde{\lambda}_i$ defined above. In columns 3-4, we include all controls in our dataset, which is more flexible as they "absorb" the Angelucci and Prat (2021) variable. Columns 1 and 3 control for state/employment/occupation fixed effects, columns 2 and 4 also include individual fixed effects for the panel sample.

The results tell a consistent story. There is a negative association between mortality risk and knowledge: a standard deviation increase in mortality risk is associated with a roughly 0.05-0.14 standard deviation reduction in knowledge. The interaction between mortality risk and the linear trend substantially counteracts this association. As we move from the beginning to the end of the survey time horizon, high-mortality respondents' knowledge disadvantage is reduced. For example, consider the regression with full controls (column 4). The results indicate that on the date of the first survey ($t = 0$), a high-risk individual (with RMI_i one standard deviation above the mean) had a 0.138 (s.e. 0.036) standard deviation lower Covid-19 knowledge than a respondent with mean mortality risk. By April 30 ($t = 1$), that same high-risk individual's knowledge disadvantage was reduced to 0.138-0.082=0.056 (s.e. 0.03) standard deviations. A similar pattern is shown in the specification with $\tilde{\lambda}_i$ and a reduced set of controls (column 2), though the estimates are less precisely estimated. The same analyses are available for the xs and p subsamples in Table A.3. This pattern is nearly identical in the cross-section sample (columns 1 and 4), and somewhat muted in the panel sample (columns 2-3 and 5-6). In other words, we find that high-mortality-risk respondents, who stand the most to gain from acquiring accurate information about risk and mitigation behaviors, "caught up" with their peers in terms of Covid-19 knowledge over the course of the first two months of the pandemic. Another feature of the results in Table 1 is the strong negative correlation between individuals' cost of acquiring information and Covid-19 knowledge. A standard deviation decrease in this proxy for the information cost is associated with a 0.22-0.24 standard deviation increase in Covid-19 knowledge. This suggests that there is heterogeneity across individuals in their cost of acquiring new information, and that cost is strongly positively correlated across domains.

We next investigate whether high-mortality-risk individuals simply absorbed more information overall–not just specific to Covid-19–in the early months of the pandemic. This might occur if, for example, they were more likely to stay home during lockdown orders and consumed more media.

To explore this question, we compare the results for Covid-19 knowledge with separate surveys conducted by Angelucci and Prat (2021) that measured political news knowledge (not related to Covid-19) collected over the same timeframe. In order to make this comparison "apples-to-apples", we focus on the quiz questions asked in both sets of surveys, which were structured and incentivized identically. In columns 1-4 of Table 2, we replicate Table 1 using the index k_{it}^{q} based on only the Covid-19 quiz questions in our sample. In columns 5-8, we estimate the same regression specification to analyze general political news questions using the AP surveys. Columns 5 and 7 show results for the three AP surveys concurrent with the Covid surveys (see Figure A.1); columns 6 and 8 includes seven additional AP survey rounds taking place throughout 2018-2019. Whether we analyze all Covid-19 knowledge as in Table 1 or only Covid-19 news questions as in Table 2 (which are very similar, though the latter are slightly noisier), we find consistent evidence that high-mortality-risk individuals had lower Covid-19 knowledge at the start of the pandemic, and substantially closed that gap within two months. We see no such pattern for general political news; in fact, the coefficient on $Days \times RMI$ is negative, though imprecisely estimated. These results suggest that high-risk individuals differentially acquired information about Covid-19, relative to general news, in the early weeks of the pandemic. The other coefficient of interest in Table 2 is the coefficient on information cost. The magnitude of the coefficient is larger for general news knowledge (column 5, coefficient of -0.378) than for Covid-19 knowledge (column 2, coefficient of -0.220) in spring 2020, and is larger still for general news knowledge over a longer time horizon (column 6, coefficient of -0.782).[18]

4.1 Knowledge Regressions: Robustness and Mechanisms

The above pattern of results is robust to alternative ways of modeling mortality risk and decisions regarding the flexibility of our trends. Recall that our baseline construction of RMI includes variation in mortality risk along the dimensions of age, gender, and race/ethnicity. To understand at a more granular level how the different dimensions of risk contribute to the knowledge acquisition results, Figure A.3 regresses Covid-19 knowledge k_{it} on $Days$, alone and interacted flexibly with each demographic characteristic that enters RMI. The top six estimates compare individuals by age, in 10-year age groupings. While the trend in knowledge is higher for each age group relative to the holdout sample of individuals under 25, the strongest association is for individuals aged 75 and over, among whom the Covid-19 mortality risk is highest. Columns 1-2 of Table A.4 echo these results; they present estimates of equation (3) using a relative mortality risk index RMI^a based only on age, and the pattern of coefficients are very similar to our results based on the full demographic index RMI; for example, we estimate a coefficient on $Days \times RMI^a$ of 0.074 in column (1). Turning back to Figure A.3, we observe a consistent pattern in which high-mortality-risk demographic groups increased their Covid-19 knowledge more over the course of our survey. With the exception only of the "Other" race/ethnicity group, we estimate higher knowledge trends

[18]The large coefficient in column 6 is partially mechanical, because the 2018-19 AP surveys are used to generate the predicted information cost proxy $\tilde{\lambda}_i$.

for each race/ethnicity group with higher relative mortality risk than the holdout group of White individuals. We also estimate slightly higher knowledge increases for men relative to the holdout group of women.

Regarding the time trends, the remaining columns of Table A.4 interact RMI with round dummies (columns 3-4) or a linear trend in survey rounds (columns 5-6) instead of a linear trend in days. In columns 3-4, we see that the largest period of catch-up among high-risk individuals was survey round 3 (April 8-13). In columns 5-6, we see results very similar to the *Days* trend in Table 1.

Lastly, we examine whether high-risk respondents' knowledge increases can be attributed to changes in their media consumption in spring 2020. To do so, we estimate equation (3), replacing the dependent variable k_{it} with a media consumption variable: the total minutes of time spent during the previous 24 hours consuming mainstream media (local television, national television, newspaper, and radio), social media, and messaging apps, winsorized at the 99th percentile.[19] The results are presented in Table A.5. The results are all estimated quite imprecisely, with standard errors exceeding the point estimates, and flip sign from specification to specification. Thus, there is no detectible association between relative mortality risk and media consumption. This means that the increased Covid-19 knowledge acquisition undertaken by high-risk individuals was achieved by other means, such as a shift in the informative nature or quality of media consumption or by paying closer attention to Covid-19 news.

4.2 Behavior Regressions

Lastly, we present evidence on compliance with Covid-19 social distancing guidelines. First, we investigate the correlation between knowledge and behavior, by regressing beh_{it} on k_{it}; the results are in Panel A of Table 3. There is a positive correlation between knowledge and behavior – as predicted by equation (1) in our simple model, those with greater knowledge regarding Covid-19 exhibit better compliance. However, the coefficients are modest in magnitude, particularly when we control for individual fixed effects. Next, we explore the relationship between relative mortality risk and the trend in social distancing behavior in the early weeks of the pandemic by replacing k_{it} in equation (3) with beh_{it}. The results are presented in Panel B of Table 3.[20] The pattern of coefficients follows the pattern observed for knowledge, with high-risk respondents beginning the survey horizon with worse social distancing compliance in most specifications, and improving over time. These results are imprecisely estimated. However, these patterns of coefficients are consistent with the model, in which high-risk individuals with knowledge of the properties of the virus would exhibit greater social distancing compliance.

[19] See Table B.5 for a description of the news consumption survey questions.

[20] The full results by sample are reported in Table A.6.

5 Discussion

Consistent with Rational Inattention, individuals with higher costs of acquiring information–as proxied by the inverse of their steady-state knowledge of the news–exhibited lower levels of knowledge about Covid-19. The role played by the mortality risk index is more subtle. On the one hand, individuals with higher mortality risk indexes largely closed the knowledge gap that existed during the first survey round. This positive association between the mortality risk index and knowledge is consistent with the theory, and it also suggests that the index itself is a valid proxy of individuals' benefit of acquiring information. On the other hand, individuals with a higher mortality risk index started out with less knowledge and ended up with a level of knowledge no higher than that of the rest of the population. (It is, of course, possible that individuals with a higher mortality risk index achieved a higher level of knowledge about Covid-19 after our fourth survey ended.) One possibility is that individuals' mortality risk is positively correlated with their costs of acquiring information and that steady-state knowledge of the news may not fully account for these information acquisition costs. Another possibility is that the relevant information about Covid-19 was finite, implying that there was limited room for individuals with greater mortality risk to end up with higher knowledge than the rest of the population.

These findings are consistent with a model of individual-level optimization. However, they also identify a set of individuals who were at high risk from Covid-19 but faced a high cost of acquiring information from regular media channels. These individuals took weeks to learn as much as the rest of the population. This finding indicates a role for policies aimed at providing alternative information channels for this set of high-risk/low-information individuals. For example, Alsan et al. (2021) show the efficacy of physician delivered messages to Black and Latinx respondents. Future research should develop methodologies to identify individuals in this high-risk, low-information set and identify the most effective communication strategies to reduce their initial information acquisition cost in emergency situations.

References

Abaluck, J. and J. Gruber (2011): "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *The American Economic Review*, 101, 1180–1210.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang (2020): "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic," Working paper.

Alsan, M., F. Cody Stanford, A. Banerjee, E. Breza, A. G. Chandrasekhar, S. Eichmeyer, P. Goldsmith-Pinkham, L. Ogbu-Nwobodo, B. A. Olken, C. Torres, A. Sankar, P.-L. Vautrey, and E. Duflo (2021): "Comparison of Knowledge and Information-Seeking Behavior After General COVID-19 Public Health Messages and Messages Tailored for Black and Latinx Communities," *Annals of Internal Medicine*, 174, 484–492, pMID: 33347320.

Alsan, M., S. Stantcheva, D. Yang, and D. Cutler (2020): "Disparities in Coronavirus 2019 Reported Incidence, Knowledge, and Behavior Among US Adults," *JAMA Network Open*, 3, 533–534.

Angelucci, C. and A. Prat (2021): "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News," *mimeo, Columbia University*.

Bartos, V., M. Bauer, J. Chytilova, and F. Matejka (2016): "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review*, 106, 1437–75.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012): "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment," *Quarterly Journal of Economics*, 127, 1205–1242.

Bhargava, S. and D. Manoli (2015): "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *Quarterly Journal of Economics*, 105, 3489–3529.

Bhattacharya, V. and G. Howard (Forthcoming): "Salience and Taxation: Theory and Evidence," *American Economic Journal: Microeconomics*.

Brown, C. S. and M. Ravallion (2020): "Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties," Working Paper 27549, National Bureau of Economic Research.

Brown, J. R. and A. Goolsbee (2002): "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy*, 110, 481–507.

Brown, Z. Y. and J. Jeon (2020): "Endogenous Information and Simplifying Insurance Choice," Working paper.

Bundorf, M. K., J. DeMatteis, G. Miller, M. Polyakova, J. L. Streeter, and J. Wivagg (2021): "Risk Perceptions and Protective Behaviors: Evidence from COVID-19 Pandemic," Working Paper 28741, National Bureau of Economic Research.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017): "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," *American Economic Journal: Macroeconomics*, 9, 1–35.

CDC (2020a): "Are You at Higher Risk for Severe Illness?" Available at https://www.cdc.gov/coronavirus/2019-ncov/specific-groups/high-risk-complications.html <accessed 03/23/2020>.

——— (2020b): "Coronavirus Disease 2019 Basics," Available at https://www.cdc.gov/coronavirus/2019-ncov/faq.html <accessed 04/08/2020>.

——— (2020c): "Coronavirus Disease 2019 (COVID-19) Frequently Asked Questions and Answers," Available at https://www.cdc.gov/coronavirus/2019-ncov/faq.html <accessed 03/08/2020>.

——— (2020d): "People at Higher Risk for COVID-19 Complications," Available at https://www.cdc.gov/coronavirus/2019-ncov/specific-groups/high-risk-complications.html <accessed 03/08/2020>.

——— (2020e): "Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States, February 12–March 16, 2020," Available at https://www.cdc.gov/mmwr/volumes/69/wr/mm6912e2.htm <accessed 03/24/2020>.

——— (2020f): "Steps to Prevent Illness," Available at https://www.cdc.gov/coronavirus/2019-ncov/about/prevention.html <accessed 03/12/2020>.

——— (2020g): "Symptoms," Available at https://www.cdc.gov/coronavirus/2019-ncov/about/symptoms.html <accessed 03/08/2020>.

——— (2020h): "Use of Cloth Face Coverings to Help Slow the Spread of COVID-19," Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/diy-cloth-face-coverings.html, accessed 04/09/2021.

——— (2020i): "Weekly Updates by Select Demographic and Geographic Characteristics," Retrieved from https://www.cdc.gov/nchs/nvss/vsrr/covid_weekly/index.htm, accessed 05/08/2020.

CIDRAP (2020): "Italy expands COVID-19 lockdown to whole country," Available at https://www.cidrap.umn.edu/news-perspective/2020/03/italy-expands-covid-19-lockdown-whole-country <accessed 03/12/2020>.

Coibion, O. and Y. Gorodnichenko (2012): "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120, 116–159.

——— (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105, 2644–78.

da Costa, V. G., M. V. Saivish, D. E. R. Santos, R. F. de Lima Silva, and M. L. Moreli (2020): "Comparative epidemiology between the 2009 H1N1 influenza and COVID-19 pandemics," *Journal of Infection and Public Health*, 13, 1797–1804.

Davis, L. W. and G. E. Metcalf (2016): "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels," *Journal of the Association of Environmental and Resource Economists*, 3.

Dong, E., H. Du, and L. Gardner (2020): "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet infectious diseases*, 20, 533–534.

Faia, E., A. Fuster, V. Pezone, and B. Zafar (2021): "Biases in Information Selection and Processing: Survey Evidence from the Pandemic," Working Paper 28484, National Bureau of Economic Research.

Fan, Y., A. Y. Orhun, and D. Turjeman (2020): "Heterogeneous Actions, Beliefs, Constraints and Risk Tolerance During the COVID-19 Pandemic," Working Paper 27211, National Bureau of Economic Research.

FDA (2020): "Coronavirus (COVID-19) Update: Daily Roundup April 14, 2020," Available at https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-daily-roundup-april-14-2020 <accessed 04/20/2020>.

Fenichel, E. P., C. Castillo-Chavez, M. G. Ceddia, G. Chowell, P. A. G. Parra, G. J. Hickling, G. Holloway, R. Horan, B. Morin, C. Perrings, M. Springborn, L. Velazquez, and C. Villalobos (2011): "Adaptive human behavior in epidemiological models," *Proceedings of the National Academy of Sciences*, 108, 6306–6311.

Fuster, A., R. Perez-Truglia, M. Wiederholt, and B. Zafar (forthcoming): "Expectations with Endogenous Information Acquisition: An Experimental Investigation," *Review of Economics and Statistics*.

Gaglianone, W. P., R. Giacomini, J. V. Issler, and V. Skreta (2020): "Incentive-driven inattention," *Journal of Econometrics*.

Galasso, V., V. Pons, P. Profeta, M. Becher, S. Brouard, and M. Foucault (2020): "Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries," *Proceedings of the National Academy of Sciences*, 117, 27285–27291.

Hastings, J. S. and J. M. Weinstein (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 123, 1373–1414.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016): "A Rational Theory of Mutual Funds' Attention Allocation," *Econometrica*, 84, 571–626.

Knittel, C. R. and B. Ozaltun (2020): "What Does and Does Not Correlate with COVID-19 Death Rates," Working Paper 27391, National Bureau of Economic Research.

Kritz, F. and P. Huang (2020): "Coronavirus Symptoms: Defining Mild, Moderate And Severe," *National Public Radio*.

Lacetera, N., D. G. Pope, and J. R. Sydnor (2012): "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 102, 2206–36.

Lovelace, B. J. (2020): "The coronavirus may be deadlier than the 1918 flu: Here's how it stacks up to other pandemics," *CNBC*.

Mackowiak, B., F. Matejka, and M. Wiederholt (2018): "Survey: Rational Inattention, a Disciplined Behavioral Model," CEPR Working Paper.

Morrison, W. and D. Taubinsky (2021): "Rules of Thumb and Attention Elasticities: Evidence from Under- and Overreaction to Taxes," Accepted at Review of Economic Studies.

NIH (2020): "Study suggests new coronavirus may remain on surfaces for days," Available at https://www.nih.gov/news-events/nih-research-matters/study-suggests-new-coronavirus-may-remain-surfaces-days <accessed 03/24/2020>.

Philipson, T. (2000): "Chapter 33 Economic epidemiology and infectious diseases," Elsevier, vol. 1 of *Handbook of Health Economics*, 1761 – 1799.

Polyakova, M., V. Udalova, G. Kocks, K. Genadek, K. Finlay, and A. N. Finkelstein (2021): "Racial Disparities in Excess All-Cause Mortality During the Early COVID-19 Pandemic Varied Substantially Across States," *Health Affairs*, 40.

Raifman, J., K. Nocka, D. K. Jones, J. Bor, S. K. Lipson, J. Jay, and P. A. C. <accessed 04/24/2020> (2020): "COVID-19 US state policy database," Available at: www.tinyurl.com/statepolicies.

Resnick, B. (2020): "12 things everyone needs to know about the coronavirus pandemic," Available at 12 things everyone needs to know about the coronavirus pandemic <accessed 04/02/2020>.

Scott Morton, F., F. Zettelmeyer, and J. Silva-Risso (2001): "Internet Car Retailing," *Journal of Industrial Economics*, 49, 501–519.

Sheikh, K., D. Watkins, J. Wu, and M. Grondahl (2020): “How Bad Will the Coronavirus Outbreak Get? Here Are 6 Key Factors,” *The New York Times*.

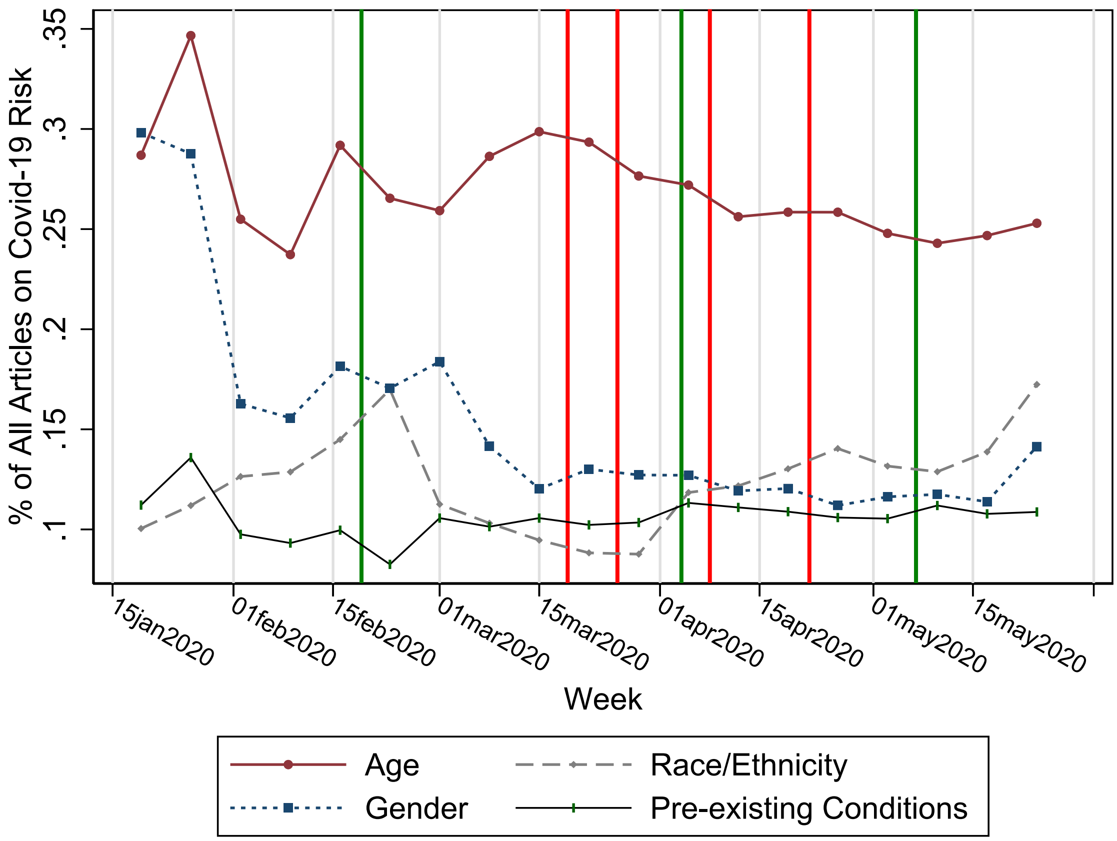
WHO (2020a): “Coronavirus (COVID-19),” Available at https://covid19.who.int/ <accessed 04/21/2020>.

——— (2020b): “Coronavirus disease (COVID-19) advice for the public: Myth busters,” Available at https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters <accessed 03/24/2020>.

——— (2020c): “Covid-19 Transmission Questions,” Available at https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-21-epi-win-covid-19-transmission-q-a.pdf <accessed 04/08/2020>.

Figures and Tables

Figure 1: Timeline of News Reports Regarding Covid-19 Mortality Risk Factors



Notes: Figure reports the weekly count of unique English-language news articles referencing Covid-19 mortality risk, between the the date of the first confirmed case of Covid-19 in the U.S. (January 19, 2020) and May 31, 2020. Counts are based on the results of searches for the terms: "(covid* OR coronavirus OR sars-cov-2) AND (mortalit* OR death OR risk)" and several different sets of characteristics in LexisNexis. Characteristics searched were those associated with elevated risk of Covid-19 mortality in CDC reports in spring 2020. The "Age" trend is based on the search terms: "age OR old*". The "Gender" trend is based on the search terms: "male OR men". The "Race/Ethnicity" trend is based on the search terms: "black OR hispanic OR latin* OR minorit* OR american indian OR native american OR indigenous OR alaska native". The "Conditions" trend is based on the search terms: "preexisting condition* OR comorbid* OR cancer OR kidney disease OR chronic obstructive pulmonary disease OR COPD OR heart failure OR heart disease OR cardiovascular disease OR coronary artery disease OR cardiomyopath* OR immunocompromise* OR obes* OR body mass index OR BMI OR pregnan* OR sickle cell OR smok* OR diabet*".

Table 1: Knowledge Index and Age-Sex-Race Mortality Risk

	(1)	(2)	(3)	(4)
RMI_i	-0.054**	-0.119***	-0.069***	-0.138***
	(0.022)	(0.037)	(0.018)	(0.036)
Days (norm) × RMI_i	0.062*	0.055	0.078**	0.082*
	(0.033)	(0.041)	(0.031)	(0.047)
Info. Costs ($\tilde{\lambda}_i$)	-0.241***	-0.215***		
	(0.016)	(0.029)		
Sample	xs+p	xs+p	xs+p	xs+p
Controls	†	†	‡	‡
Fixed Effects	{s,o}	{s,o,i}	{s,o}	{s,o,i}
R2	0.20	0.59	0.27	0.62
N	8,004	8,004	8,004	8,004

Notes: Each column presents the results of a different regression of our knowledge index on relative mortality risk RMI_i, alone and interacted with a linear trend, and controls for information cost. The regression sample includes cross-section (xs) and panel (p) respondents. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; and for state and occupation ($\{s, o\}$) fixed effects. Specifications labeled $\{s, o, i\}$ also include individual fixed effects for panel respondents. † denotes specifications that control directly for the information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with our richest set of controls: local population density; marriage status; number of children and adults over 65 in the household; level of education attained; income bin; employment status; political party registration, strength of party affiliation, and liberal/conservative self-identification; political news interest; pre-pandemic time spent on mainstream media, messaging apps, and social media; receipt of the flu vaccine; importance of religion and frequency of church attendance; and a binary variable indicating respondents' answer to the question "Can people be trusted?". Standard errors are clustered by state.

Table 2: Knowledge Index and Mortality Risk – Covid-19 Knowledge vs. General Knowledge

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RMI_i	-0.040	-0.123**	-0.053**	-0.138***	-0.046	0.047**	-0.059*	-0.004
	(0.027)	(0.046)	(0.024)	(0.041)	(0.034)	(0.022)	(0.035)	(0.041)
Days (norm) × RMI_i	0.051	0.103**	0.061	0.118**	-0.045	-0.062	-0.020	-0.041
	(0.045)	(0.044)	(0.044)	(0.046)	(0.082)	(0.045)	(0.085)	(0.060)
Info. Costs ($\tilde{\lambda}_i$)	-0.220***	-0.220***			-0.377***	-0.782***		
	(0.014)	(0.029)			(0.055)	(0.010)		
Sample	xs+p	xs+p	xs+p	xs+p	AP Sample	AP Sample	AP Sample	AP Sample
Controls	†	†	‡	‡	†	†	‡	‡
Knowledge Qs	C19 Quiz	C19 Quiz	C19 Quiz	C19 Quiz	AP Quiz	AP Quiz	AP Quiz	AP Quiz
Rounds	1-4	1-4	1-4	1-4	8-10	1-10	8-10	1-10
Fixed Effects	{s,o}	{s,o,i}	{s,o}	{s,o,i}	{s}	{s}	{s}	{s}
R2	0.13	0.53	0.20	0.55	0.14	0.44	0.18	0.12
N	8,004	8,004	8,004	8,004	1,475	6,066	1,475	6,066

Notes: Each column presents the results of a different regression of Covid-19 knowledge or general knowledge on RMI, alone and interacted with a linear trend (adjusted for each sample's time horizon), and controls for information cost. Columns 1-4 replicate Columns 1-4 of Table 1, with the knowledge index formed only using "quiz" questions regarding Covid-19 political news. Columns (5)-(8) show the same specifications, with the knowledge index formed using the general knowledge quiz questions in the AP surveys. Columns (5) and (7) present results for three AP survey rounds contemporaneous with the Covid-19 surveys; columns (6) and (8) also include seven AP survey rounds taking place throughout 2018-2019. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with linear trend and for indicator variables for each possible question set survey respondents could have been administered. All Covid-19 survey regressions control for state and occupation $\{s, o\}$ fixed effects. Some specifications $\{s, o, i\}$ also include individual fixed effects for panel respondents. All AP sample regressions include state fixed effects. † denotes specifications with information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with the richest set of controls. Standard errors are clustered by state.

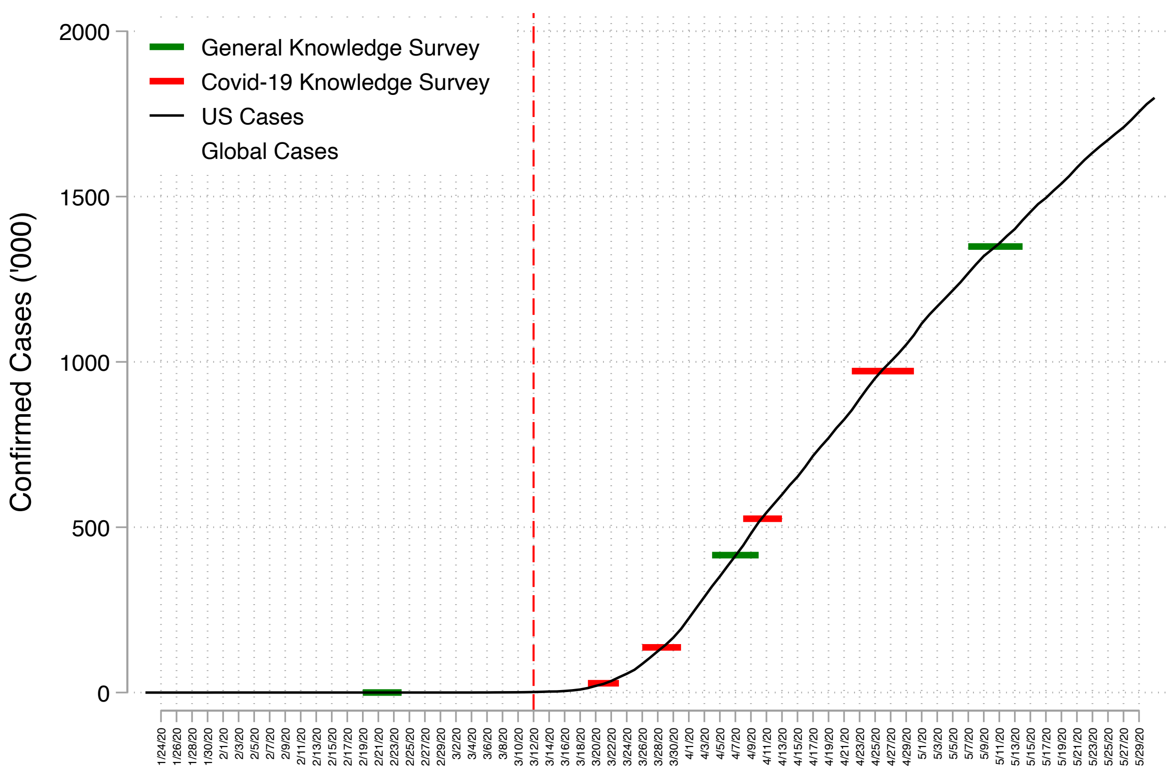
Table 3: Correlates of Behavior

	Panel A: Knowledge and Behavior Index			
	(1)	(2)	(3)	(4)
Knowledge Index	0.139***	0.041***	0.117***	0.025*
	(0.014)	(0.015)	(0.015)	(0.014)
Info. Costs ($\tilde{\lambda}_i$)	0.003	-0.008		
	(0.013)	(0.035)		
Sample	xs+p	xs+p	xs+p	xs+p
Controls	†	†	‡	‡
Fixed Effects	{s,o}	{s,o,i}	{s,o}	{s,o,i}
R2	0.13	0.61	0.18	0.63
N	7,844	7,844	7,844	7,844
	Panel B: Behavior Index and Mortality Risk			
	(1)	(2)	(3)	(4)
RMI_i	-0.040*	-0.011	-0.036	0.001
	(0.023)	(0.026)	(0.023)	(0.028)
Days (norm) × RMI_i	0.068	0.045	0.066	0.056
	(0.043)	(0.037)	(0.040)	(0.035)
Info. Costs ($\tilde{\lambda}_i$)	-0.029**	-0.016		
	(0.014)	(0.034)		
Sample	xs+p	xs+p	xs+p	xs+p
Controls	†	†	‡	‡
Fixed Effects	{s,o}	{s,o,i}	{s,o}	{s,o,i}
R2	0.12	0.61	0.17	0.63
N	7,844	7,844	7,844	7,844

Notes: Panel A shows the results of different regressions of our index of social distancing compliance beh_{it} on Covid-19 knowledge k_{it}. Panel A presents the results of regressions of beh_{it} on RMI, alone and interacted with a linear trend. The regression sample includes cross-section (xs) and panel (p) respondents. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with linear trend; for indicator variables for each possible question set survey respondents could have been presented with; and for state and occupation $\{s, o\}$ fixed effects. Specifications labeled $\{s, o, i\}$ also include individual fixed effects for panel respondents. † denotes specifications with information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with the richest set of controls. Standard errors are clustered by state.

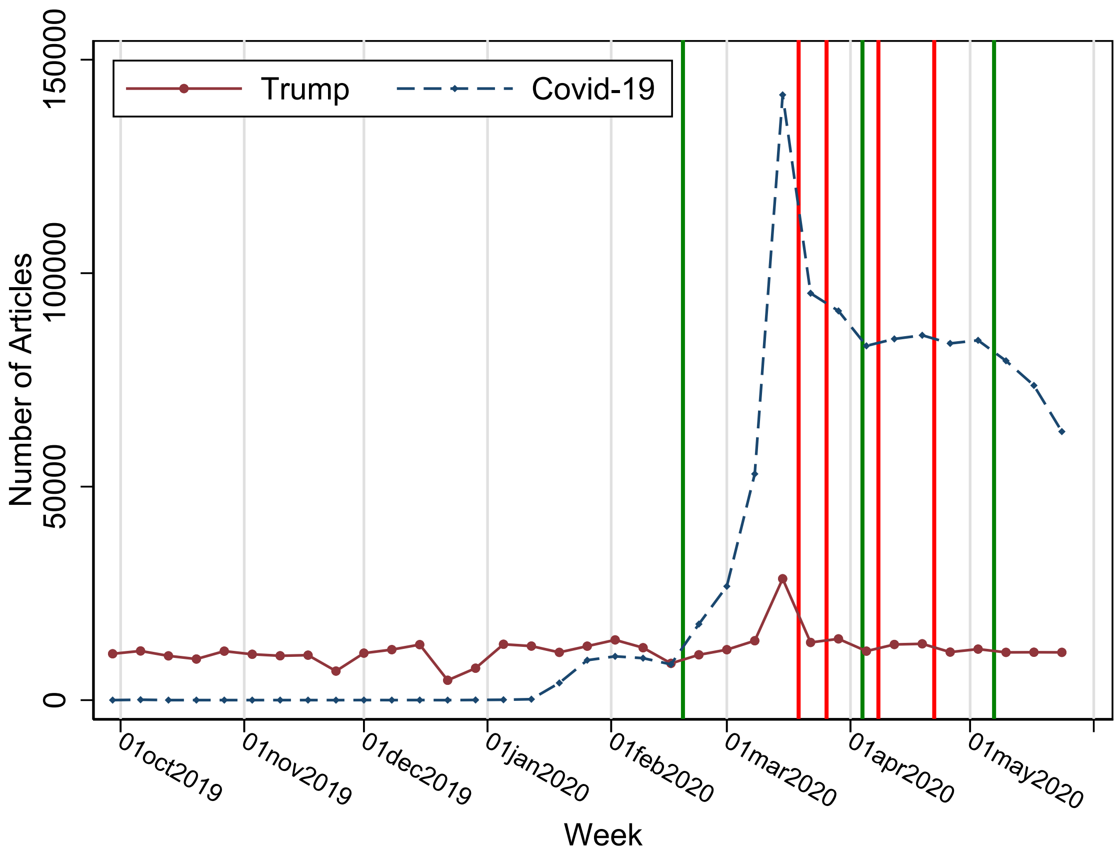
Appendix A: Extra Figures and Tables

Figure A.1: Survey and Case Timeline



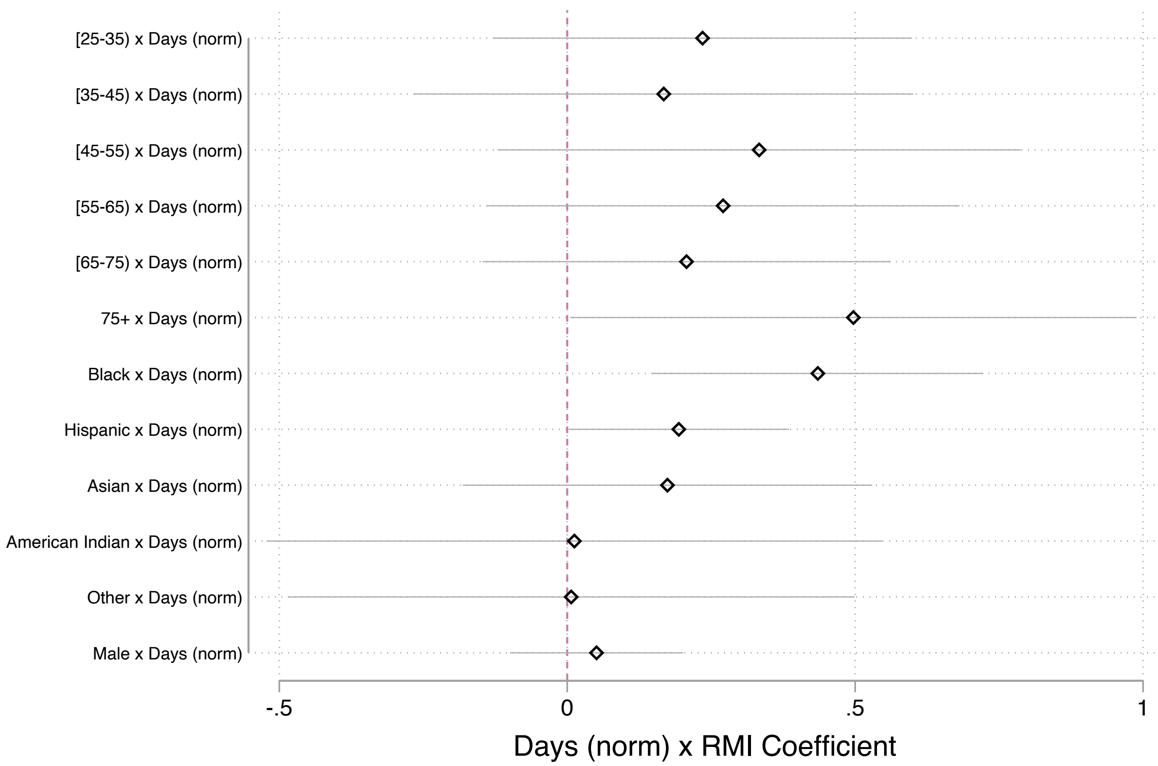
Notes: Timeline of Covid-19 and (Angelucci and Prat, 2021) General Knowledge Survey against total confirmed Covid-19 case counts in the U.S. taken from Dong et al. (2020).

Figure A.2: Timeline of Covid-19 and Trump News Coverage



Notes: Figure reports the weekly count of unique English-language news articles referencing Covid-19 and President Trump between October 1, 2019 and May 31, 2020, based on the results of searches for the terms: "covid* OR coronavirus OR sars-cov-2" and "trump," respectively, in Lexis-Nexis.

Figure A.3: Knowledge vs. Mortality Risk Factors



Notes: The figure presents the results of a regression of our Covid-19 knowledge index on a linear trend, alone and interacted with indicators for each demographic characteristic that enters RMI, and a control for the information cost proxy $\tilde{\lambda}_i$. The holdout groups for the age, race/ethnicity, and gender dimensions are under 25 years old, white, and female, respectively. The regression sample includes cross-section (xs) and panel (p) respondents. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; for state and occupation fixed effects; and for individual fixed effects for panel respondents. Each diamond in the figure represents the point estimate, while the solid grey braces represent 95% confidence intervals. Standard errors are clustered by state.

Table A.1: Relative Mortality Risk Index

Age Group	Female RMI_i	Male RMI_i	Total RMI_i	Race/ Ethnicity	RMI_i
under 5 years	0.002	0.002	0.002	White	0.868
5-14 years	0.000	0.001	0.001	Black	1.752
15-24 years	0.006	0.011	0.008	Hispanic	0.907
25-34 years	0.032	0.072	0.052	Asian	1.035
35-44 years	0.079	0.208	0.143	Native American	0.571
45-54 years	0.221	0.563	0.390	Other	1.167
55-64 years	0.618	1.334	0.963		
65-74 years	1.661	3.255	2.406		
75 years and over	7.869	10.992	9.143		

Notes: Table presents Relative Mortality Risk Index (RMI_i) constructed as described in text, for each demographic grouping $j \in \{ag, r\}$, where ag denotes age-gender combination, and r denotes race/ethnicity.

Table A.2: Summary Statistics

Panel A: Numerical variables

	Mean	SD
Age	49.54	17.34
# kids [0,4]	0.16	0.56
# kids [5,12]	0.23	0.63
# kids [13,18]	0.21	0.57
# adults > 65	0.37	0.69
Mainstream media (pre-covid)	96.33	152.81
Message media (pre-covid)	30.05	56.96
Social media (pre-covid)	65.59	102.51
Log population density	6.95	1.87

Notes: Table reports the mean and standard deviation of all numerical variables included as controls in the regressions. Media consumption is measured in minutes.

Panel B: Categorical variables

	Fraction
Race: White	0.67
Race: Black	0.11
Race: Hispanic	0.13
Gender: Female	0.45
Marital Status: Married	0.29
Believes most people can be trusted	0.33
Flu vaccine: Received	0.53
Religion: Not important	0.25
Religion: Not too important	0.15
Religion: Somewhat important	0.24
Interest in news: Hardly at all	0.10
Interest in news: Some of the time	0.63
Interest in news: Most of the time	0.26
Education: No HS	0.04
Education: High school graduate	0.31
Education: Some college	0.21
Education: 2-year college	0.11
Education: 4-year college	0.20
Income: Missing	0.13
Income: <$30k	0.25
Income: [$\$30k - \$50k$]	0.19
Income: [$\$50k - \$80k$]	0.21
Income: [$\$80k - \$150k$]	0.17
Party affiliation: Strong Democrat	0.26
Party affiliation: Not very strong Democrat	0.11
Party affiliation: Lean Democrat	0.10
Party affiliation: Independent	0.16
Party affiliation: Lean Republican	0.08
Party affiliation: Not very strong Republican	0.08
Party affiliation: Strong Republican	0.17
Labor force participation	0.48

Notes: Table provides descriptive statistics for all the categorical variables included as controls in the regressions. For each variable, the share of respondents belonging to all underlying categories except one is reported (for simplicity, we also exclude categories with few respondents).

Table A.3: Knowledge Index and Age-Sex-Race Mortality Risk, by Sample

	(1)	(2)	(3)	(4)	(5)	(6)
RMI_i	-0.057***	-0.036		-0.071***	-0.047*	
	(0.020)	(0.030)		(0.017)	(0.027)	
Days (norm) × RMI_i	0.076*	0.022	0.040	0.094**	0.026	0.037
	(0.038)	(0.050)	(0.045)	(0.036)	(0.045)	(0.045)
Info. Costs ($\tilde{\lambda}_i$)	-0.249***	-0.241***				
	(0.015)	(0.020)				
Sample	xs	p	p	xs	p	p
Controls	†	†	†	‡	‡	‡
Fixed Effects	{s,o}	{s,o}	{i}	{s,o}	{s,o}	{i}
R2	0.21	0.20	0.68	0.29	0.28	0.68
N	5,448	4,225	4,225	5,448	4,225	4,225

Notes: Each column presents the results of a different regression of our knowledge index on relative mortality risk RMI, alone and interacted with a linear trend, and controls for information cost, as in Table 1. The regression samples include cross-section (xs) respondents only, panel (p) respondents only, or both, as indicated for the column. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; and for state and occupation ($\{s, o\}$) fixed effects. Specifications labeled $\{i\}$ also include individual fixed effects for panel respondents. † denotes specifications that control directly for the information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with our richest set of controls. Standard errors are clustered by state.

Table A.4: Knowledge Index and Mortality Risk, Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
RMI_i	-0.111***	-0.082**	-0.121***	-0.136***	-0.125***	-0.143***
	(0.037)	(0.034)	(0.039)	(0.037)	(0.038)	(0.036)
Days (norm) × RMI_i	0.074**	0.081**				
	(0.033)	(0.034)				
Info. Costs ($\tilde{\lambda}_i$)	-0.219***		-0.216***		-0.215***	
	(0.028)		(0.029)		(0.029)	
round=2 × RMI_i			0.022	0.016		
			(0.022)	(0.022)		
round=3 × RMI_i			0.082***	0.093***		
			(0.029)	(0.028)		
round=4 × RMI_i			0.013	0.034		
			(0.039)	(0.043)		
Trend × RMI_i					0.053	0.072*
					(0.033)	(0.036)
Version	Age-Only	Age-Only	Round Dums.	Round Dums.	Round Trend	Round Trend
Sample	xs+p	xs+p	xs+p	xs+p	xs+p	xs+p
Controls	†	‡	†	‡	†	‡
Fixed Effects	{s,o,i}	{s,o,i}	{s,o,i}	{s,o,i}	{s,o,i}	{s,o,i}
R2	0.59	0.62	0.59	0.62	0.59	0.62
N	8,004	8,004	8,004	8,004	8,004	8,004

Notes: Each column presents the results of a different regression of our knowledge index on RMI, alone and interacted with a trend, and controls for information cost. Columns 1-2 use as the dependent variable a Relative Mortality Index RMI^a that is based only on age (rather than age, race/ethnicity, and gender as in all other results). Columns 3-4 specify the trend using survey round indicators (with round 1 as the holdout period) instead of a linear trend in *Days*. Columns 5-6 use a linear trend in survey round (ranging from 0 in round 1 to 1 in round 4) instead of a linear trend in *Days*. The regression sample includes cross-section (xs) and panel (p) respondents. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; for state and occupation fixed effects; and for individual fixed effects for panel respondents. † denotes specifications that control directly for the information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with our richest set of controls. Standard errors are clustered by state.

Table A.5: Media Consumption and Mortality Risk

	(1)	(2)	(3)	(4)	(5)
RMI_i	-0.001	0.005		-0.006	0.042
	(0.023)	(0.033)		(0.028)	(0.032)
Days (norm) × RMI_i	0.024	0.011	-0.016	0.024	-0.042
	(0.059)	(0.057)	(0.048)	(0.055)	(0.044)
Sample	xs	p	p	xs+p	xs+p
Controls	†	†	†	†	†
Fixed Effects	{s,o}	{s,o}	{i}	{s,o}	{s,o,i}
R2	0.06	0.08	0.71	0.06	0.60
N	5,436	4,222	4,222	7,989	7,989

Notes: Each column presents the results of a different regression of media consumption on RMI, alone and interacted with a linear trend, and controls for information cost. The regression samples include cross-section (xs) respondents only, panel (p) respondents only, or both, as indicated for the column. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; for the information cost proxy $\tilde{\lambda}_i$; and for state and occupation ($\{s, o\}$) fixed effects. Specifications labeled $\{i\}$ and $\{s, o, i\}$ also include individual fixed effects for panel respondents. Standard errors are clustered by state.

Table A.6: Correlates of Behavior, Details

	Panel A: Knowledge and Behavior Index					
	(1)	(2)	(3)	(4)	(5)	(6)
Knowledge Index	0.146***	0.116***	-0.032**	0.122***	0.100***	-0.035**
	(0.020)	(0.019)	(0.015)	(0.020)	(0.019)	(0.014)
Info. Costs ($\tilde{\lambda}_i$)	-0.001	0.006				
	(0.015)	(0.019)				
Sample	xs	p	p	xs	p	p
Controls	†	†	†	‡	‡	‡
Fixed Effects	{s,o}	{s,o}	{i}	{s,o}	{s,o}	{i}
R2	0.12	0.15	0.74	0.17	0.21	0.74
N	5,322	4,145	4,145	5,322	4,145	4,145
	Panel B: Behavior Index and Mortality Risk					
	(1)	(2)	(3)	(4)	(5)	(6)
RMI_i	-0.046*	-0.018		-0.040*	-0.027	
	(0.023)	(0.032)		(0.023)	(0.031)	
Days (norm) × RMI_i	0.087*	-0.003	0.049	0.083*	0.003	0.037
	(0.046)	(0.065)	(0.054)	(0.042)	(0.067)	(0.053)
Info. Costs ($\tilde{\lambda}_i$)	-0.035**	-0.021				
	(0.016)	(0.021)				
Sample	xs	p	p	xs	p	p
Controls	†	†	†	‡	‡	‡
Fixed Effects	{s,o}	{s,o}	{i}	{s,o}	{s,o}	{i}
R2	0.11	0.14	0.74	0.16	0.20	0.74
N	5,322	4,145	4,145	5,322	4,145	4,145

Notes: Panel A shows the results of different regressions of our index of social distancing compliance beh_{it} on Covid-19 knowledge k_{it}. Panel A presents the results of regressions of beh_{it} on RMI, alone and interacted with a linear trend. Each column presents the results of a different regression, as in Table 3. The regression samples include cross-section (xs) respondents only, panel (p) respondents only, or both, as indicated for the column. All regressions control for age, gender, and race/ethnicity indicators, each alone and interacted with the linear trend; for indicator variables for each possible question set survey respondents could have been presented with; and for state and occupation ($\{s, o\}$) fixed effects. Specifications labeled $\{i\}$ also include individual fixed effects for panel respondents. † denotes specifications that control directly for the information cost proxy $\tilde{\lambda}_i$. ‡ denotes specifications that replace $\tilde{\lambda}_i$ with our richest set of controls. Standard errors are clustered by state.

Appendix B: Knowledge and Behavior Survey Questions

Table B.1: Knowledge Questions

		Question	Sample	Round
Part A	1	In what city did Coronavirus originate? • Nanjing, Xian, Beijing, Hong Kong, **Wuhan**, Chongqing	xs1, xs2, xs3, xs4, p1	1,2,3,4
	2	What is the recommended distance to keep between yourself and another person to prevent airborne transmission of Coronavirus? Pick any value between 0-25 feet	xs1, xs2, xs3, p1	1,2,3,4
	3	**T**/F: Coronavirus is more contagious on average than the seasonal flu.	xs1, xs2, xs3, xs4, p1	1,2,3,4
	4	T/**F**: Coronavirus is more contagious on average than measles.	xs1, xs2, xs3, xs4, p1	1,2,3,4
	5	To the best of your knowledge, which of the following statements is most true? • The most common symptom of Coronavirus is shortness of breath. • **The most common symptom of Coronavirus is fever.** • Nearly all individuals with Coronavirus experience fever and shortness of breath.	xs1, xs2, xs3, xs4, p1	1,2,3,4
	6	To the best of your knowledge, which of the following statements is most true? • Young children and older adults have about the same risk of severe illness from Coronavirus. • **Older adults are at greater risk of severe illness from Coronavirus.** • Young children are at greater risk of severe illness from coronavirus.	xs1, xs2, xs3, xs4, p1	1,2,3,4
Part B	7	As of yesterday, how many confirmed cases of Coronavirus were reported in your state?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
	8	As of yesterday, how many confirmed deaths from Coronavirus were reported in your state?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
Part C	9	Which of the following groups is NOT thought to be high risk for severe illness from Coronavirus (choose one)? • People with immune deficiencies • **People with osteoporosis (low bone density)** • People who live in a nursing home • People with chronic lung disease	p2, xs3, xs4	2,3,4
	10	Which of the following countries was the first, other than China, to declare a lockdown in response to the Coronavirus threat (choose one)? • Germany, **Italy**, United States, Ireland, Japan	p2, xs3, xs4	2,3,4
	11	On which of the following surfaces is the Coronavirus detectable for longest (choose one)? • Copper, **Plastic**, cardboard, Cotton fabric	p2, xs3, xs4	2,3,4
	12	**T**/F: Vaccines against pneumonia do not provide protection against Coronavirus.	p2, xs3, xs4	2,3,4
	13	T/**F**: In addition to human-to-human transmission, the Coronavirus can be transmitted through mosquito and flea bites.	p2, xs3, xs4	2,3,4
	14	T/**F**: The majority of individuals with Coronavirus develop symptoms within 3 days of exposure.	p2, xs3, xs4	2,3,4
	15	T/**F**: Exposing yourself to the sun or to temperatures higher than 77 degrees Fahrenheit prevents the coronavirus disease	p3, xs4	3,4
	16	As of today, which country has the most confirmed Coronavirus cases? • **United States**, China, South Korea, Italy	p3, xs4	3,4
	17	**T**/F: Fewer than 25% of Coronavirus cases have resulted in serious illness.	p3, xs4	3,4
	18	According to experts' estimates, each infected person is expected to spread the virus to how many others, on average, if no control measures are put into place? • 1-1.25 • **2-2.25** • 3-4 • More than 5	p3, xs4	3,4
	19	T/**F**: Hand Dryers are effective in killing the coronavirus.	p3, xs4	3,4
	20	As of today, does your state have an order in place for citizens to stay home?	p3, xs4	3,4
	21	Y/**N:** Does the CDC (Centers for Disease Control and Prevention) recommend wearing cloth face coverings at home when in presence of people age 65 and above?	p3, xs4	3,4
	22	**Y**/N: Does the CDC (Centers for Disease Control and Prevention) recommend that children ages 2 to 6 wear cloth face coverings in public settings?	p3, xs4	3,4
	23	**T**/F: Other types of coronavirus usually cause mild illnesses like the common cold.	p3, xs4	3,4
	24	How many cases of coronavirus have been confirmed worldwide? • 0 to 500,000 • 500,000 to 1 million • 1 million to 2 million • **Over 2 million**	xs4, p4	4
	25	How many deaths due to coronavirus have been reported worldwide? • 0 to 10,000 • 10,000 to 50,000 • 50,000 to 100,000 • **Over 100,000**	xs4, p4	4
	26	T/**F**: The FDA has approved home tests for Coronavirus	xs4, p4	4
	27	To the best of your knowledge, what is the World health Organization's (WHO) recommendation regarding the distance to maintain from other people outside your home? • **3 feet (1 meter)** • 6 feet (2 meters) • 9 feet (3 meters) • 12 feet (4 meters)	xs4, p4	4
	28	T/**F**: Contact lens wearers are more at risk for acquiring COVID-19 than eyeglass wearers?	xs4, p4	4
	29	**T**/F: The "19" in COVID-19 stands for the year the virus was first detected: 2019	xs4, p4	4
	30	**T**/F: Symptoms of Coronavirus are similar in children and adults	xs4, p4	4

Note: The table reports the knowledge questions respondents were asked about Covid-19. The correct answers are typeset in bold (when applicable). For each question, the table specifies the surveys in which the question was included (Round 1-4) and lists the samples of respondents (cross-section "xs" and/or panel "p") who were administered the question in any given round. Sources: CDC (2020c) (Q1); CDC (2020f) (Q2); Sheikh et al. (2020) (Q3,Q4); Kritz and Huang (2020) (Q5); CDC (2020d) (Q6); Dong et al. (2020) (Q7,Q8); CDC (2020a) (Q9); CIDRAP (2020) (Q10); NIH (2020) (Q11); WHO (2020b) (Q12,Q13,Q15,Q19); CDC (2020g) (Q14); Resnick (2020) (Q16,Q18); CDC (2020e) (Q17); Raifman et al. (2020) (Q20); CDC (2020h) (Q21,Q22); CDC (2020b) (Q23,Q28,Q29,Q30); WHO (2020a) (Q24,Q25); FDA (2020) (Q26); WHO (2020c) (Q27).

Table B.2: Knowledge of Covid-19 News

Question	Sample	Round
The following list of statements contains three true statements and three false statements as of March 17 2020. To the best of your recollection, which three statements are true? Please select exactly three statements. You have 60 seconds to answer this question. • President Trump appointed Ben Carson to lead Presidential advisory council on coronavirus • **President Trump announced "The United States is in far better shape than other countries"** • President Trump announced all domestic flights cancelled due to Coronavirus • **President Trump declared the coronavirus a national emergency** • President Trump announced new vaccine against Coronavirus is available to the public • **President Trump tested negative for Coronavirus**	xs1, p1	1
The following list of statements contains three true statements and three false statements as of March 22 2020. To the best of your recollection, which three statements are true? Please select exactly three statements. You have 60 seconds to answer this question. • Senator Bernie Sanders ended his presidential campaign, citing coronavirus crisis • House of Representatives threatened a government shut down over coronavirus bill • **China said President Trump's "Chinese virus" tweet smears China** • **President Trump announced a plan to send Americans direct checks to cushion coronavirus economic shock** • **The coronavirus led Ohio to delay its Democratic Party presidential nominating election** • New York Governor Andrew Cuomo entered the presidential race, citing his successful handling of coronavirus crisis	xs2, p2	2
The following list of statements contains three true statements and three false statements as of April 7 2020. To the best of your recollection, which three statements are true? Please select exactly three statements. You have 60 seconds to answer this question. • **Federal judge refused to postpone this week's U.S. presidential primary in Wisconsin** • White House Staff required to stay inside West wing indefinitely to avoid exposure to coronavirus • **Democrats delayed presidential convention until august** • **President Trump tested negative after undergoing second coronavirus test** • Joe Biden and Jared Kushner started working together on healthcare worker relief bill • Former President Obama publicly criticized President Trump for handling of coronavirus outbreak	xs3, p3	3
The following list of statements contains three true statements and three false statements as of April 21 2020. To the best of your recollection, which three statements are true? Please select exactly three statements. You have 60 seconds to answer this question. • **Supporters of President Trump protested Michigan's stay-at-home orders at state capital** • **New York and six other Northeastern state extended coronavirus stay-at-home orders to May 15th** • **President Trump issued new federal guidelines for a partial reopening of the economy** • Advisor Anthony Fauci publicly said he would consider Surgeon General position if offered • President Trump threatened to half all shipments of Apple products from china in retaliation for lack of COVID-19 warning • President Trump blamed advisor Anthony Fauci for falling ratings on televised White House briefings	xs4, p4	4

Note: The table reports the news quizzes respondents were asked about Covid-19. The correct answers are typeset in bold (when applicable). For each question, the table specifies the surveys in which the question was included (Round 1-4) and lists the samples of respondents (cross-section "xs" and/or panel "p") who were administered the question in any given round.

Table B.3: Non-Covid-19 Political News Questions

	Question	Survey
	The following list of statements contains three true statements and three false statements. To the best of your recollection, which three statements are true? Please select exactly three statements. You have 60 seconds to answer this question.	
PART A	• **Former Trump lawyer Michael Cohen sentenced to three years prison.** • **U.S. Senate hands Trump rebuke on Saudi Arabia.** • **U.S. lawmakers to unveil revised criminal justice bill in push for final passage.** • Trump secures funding for border wall in meeting with top Democrats. • Federal Judge rules public funding for Planned Parenthood unconstitutional. • Saudi Crown Prince to address Senate in effort to clear his name in journalists murder.	Nov 2018
	• **The U.S. Government was partially shut down in fight over Trump's border wall with Mexico.** • **Democratic lawmakers called for further investigation into a revelation that in 2016 Paul Manafort gave polling data to a man linked to Russian intelligence.** • **The U.S. Supreme Court gave itself another chance to make a definitive ruling on electoral map disputes.** • Soybean farmed marched on Washington over Chinese tariffs' impacts. • Trump fired Federal Reserve Chairman Jerome Powell for raising interest rates. • Trump Threatened to Raise Border Wall Cost to $7 Billion if Stall by Democrats Continues.	Dec 2018
	• **Special Counsel Robert Mueller did not find Trump 2016 campaign knowingly conspired with Russia.** • **President Donald Trump vetoed a measure passed by Democrats and Republicans in Congress to end his emergency declaration over border wall with Mexico.** • **Vice President Mike Pence visited Nebraska to take stock of the devastation unleashed across the U.S. Midwest by floods.** • 2020 Presidential Candidate Elizabeth Warren took millions in Wall Street campaign contributions. • President Donald Trump diverted Puerto Rico aid to fund border wall with Mexico. • House Republicans Unveil Legislation To Significantly Limit Funding To Planned Parenthood Centers Nationwide.	Mar 2019
	• **Rod Rosenstein, U.S. deputy attorney general who appointed Special Counsel Robert Mueller, submits resignation.** • **Homeland Security Secretary Nielsen resigns amid Trump anger over border.** • **Trump and Democrats agree to pursue $2 trillion infrastructure Plan.** • Trump releases redacted version of his taxes to Congress. • Clinton Foundation loses nonprofit status. • The Virginia Bar Association disbars Attorney General Barr for lying to Congress.	Arp 2019
	• **Mexico agreed to take more migrants seeking asylum in the United States while they await adjudication of their cases.** • **Alabama's governor signed a bill to ban nearly all abortions in the state.** • **President Trump proposed plan to make U.S. immigration more merit-based.** • Attorney General Barr released text message from Special Counsel prosecutor Robert Mueller text: 'We're taking down Trump'. • US Border Patrol facility admitted to measles outbreak among migrant children in custody. • Trump administration to continue to allow U.S. research using fetal tissue from abortions.	May 2019
	• **Whistle-blower report complains of White House cover-up on Trump-Ukraine scandal.** • **Supreme Court granted a request by President Trump's administration to fully enforce a new rule that would curtail asylum applications by immigrants at the U.S. - Mexico border.** • **At a closed-door meeting at the White House, top envoy to China delivered evidence of rising Farm Belt frustration over bio-fuel policy.** • President Trump announces he will resume peace talks with Iran at UN General Assembly. • China blacklists Apple, Microsoft amid escalating trade war. • Vaping case to make its way to Supreme Court.	Sep 2019
	• **A whistleblower filed a complaint against President Trump, leading to an impeachment inquiry.** • **Republican lawmakers in the House of Representatives condemned President Trump's decision to withdraw troops from Syria.** • **The Trump administration credited cooperation from Mexico and Central American countries in cracking down on migrants.** • President Trump's Tax Returns showed billions given to various charities. • China and the United States agreed on a new comprehensive trade deal. • Isis beheaded three Americans in response to Al-Baghdadi's death.	Oct/Nov 2019
PART B	• **The U.S. Senate acquitted Trump of impeachment charges.** • **Attorney General William Barr said that President Trump's attacks on prosecutors, the judge and jurors in the trial of Roger Stone undermined the Justice Department's work.** • **The House of Representatives passed legislation seeking to rein in President Trump's ability to deploy U.S. forces to fight abroad.** • A Tape surfaced of President Trump supporting abortion. • Mitt Romney decided to run for president against Trump in the 2020 race after breakout role in impeachment. • President Trump took a week-long break from Campaigning to Deal with Coronavirus Outbreak.	Jan/Feb 2020
	• **U.S. Supreme Court allowed President Trump's 'Remain in Mexico' asylum policy.** • **President Trump declared coronavirus a national emergency.** • **President Trump notified Congress he is firing the inspector general of U.S. intelligence community.** • President Trump fired coronavirus advisor Dr. Anthony Fauci. • Nancy Pelosi under investigation by Justice Department over alleged insider trading during coronavirus outbreak. • Agriculture trade group marched in Washington to draw attention to export problems.	Apr 2020
	• **President Trump said he would address national debt if re-elected.** • **In win for President Trump, U.S. Supreme Court made deporting immigrants for crimes easier.** • **Senior U.S. House members vowed to pass major defense bill despite pandemic.** • President Trump's campaign saw steep rise in donations after press conferences. • President Trump announced tax returns to be released by Mid-May. • Around 20% of IRS stimulus checks bounced.	May 2020

Note: The table reports the news quizzes (Angelucci and Prat, 2021) administered from November 2018 to May 2020. For each news quiz, the date when the corresponding survey was administered is indicated.

Table B.4: Behavior Questions

Question	Sample	Round
In the past 7 days…		
Approximately how many hours per day have you spent outside your household (do not include time spent on your own property, e.g., moving the lawn)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
Of these .. hours per day that you spent outside the house… How many of these hours were work-related on average?	xs2, xs3, xs4, p2, p3, p4	2,3,4
How many distinct outings with other people (outside your household) have you had (e.g., going with friends to restaurants, bars, dates, coffee shops, playing sports, going to the mall or movies)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
Did you participate in gatherings of 50 people or more (e.g., concerts or movies)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many trips have you taken to store(s) to purchase any essential items (essential items include groceries, medical drugs, and/or basic household goods like toilet paper)?	xs2, xs3, xs4, p2, p3, p4	2,3,4
How many trips have you taken to store(s) to purchase any non-essential items (essential items include groceries, medical drugs, and/or basic household goods like toilet paper)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
In the past 24 hours…		
Approximately how many hours have you spent outside your household (do not include time spent outside your house but on your own property)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many distinct outings with other people (outside your household) have you had (e.g., going with friends to restaurants, bars, dates, coffee shops, playing sports, going to the mall or movies)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many hands have you shaken (with anyone outside your household)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many hugs have you given/received (with anyone outside your household)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many adults over the age of 65 (excluding those in your household) have you interacted with in person?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
How many "high-touch" surfaces have you disinfected (e.g., your phone, door knobs, etc.)?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
Other than the members of your household, how many people have you been within 6 feet of? (e.g., standing in line at the grocery store, passing by on the street, being in the same elevator)	xs2, xs3, xs4, p2, p3, p4	2,3,4
How many trips have you taken to store(s) to purchase any essential items (essential items include groceries, medical drugs, and/or basic household goods like toilet paper)?	xs2, xs3, xs4, p2, p3, p4	2,3,4
Excluding the trips described in the example above, how many additional trips have you taken to store(s) to purchase any non-essential items?	xs2, xs3, xs4, p2, p3, p4	2,3,4
In the past 4 hours, how many times have you washed your hands and/or used hand sanitizer?	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
Would you have close friends over to your house if they had no sign or symptoms associated with Coronavirus? [Absolutely not, Unlikely, Maybe, Likely, Definitely]	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4

Note: Table reports the questions respondents were asked about their recent behavior. For each question, the table specifies the surveys in which the question was include (rounds 1-4) and lists the samples (cross-section "xs" and/or panel "p") who were administered the question in any given round.

Table B.5: Media Consumption Questions

Question	Sample	Round
In the past 24 hours, how many total minutes have you spent on any of the following mediums: Radio, TV, Text message conversations, Social media, or Newspapers? Please break down the following by total minutes. • Messaging apps • Social media • Mainstream media (local television, national television, newspaper, radio)	xs1, xs2, xs3, xs4, p1, p2, p3, p4	1,2,3,4
In a typical day last November, how many total minutes did you spend on any of the following mediums: Radio, TV, Text message conversations, Social media, or Newspapers? Please break down the following by total minutes. • Messaging apps • Social media • Mainstream media (local television, national television, newspaper, radio)	xs1, xs2, xs3, xs4, p1	1,2,3,4

Note: The table reports the questions respondents were asked about their news media consumption habits. For each question, the table specifies the surveys in which the question was included (Round 1-4) and lists the samples of respondents (cross-section "xs" and/or panel "p") who were administered the question in any given round.