RIETI Discussion Paper Series 25-E-049

The Impact of Media Type on Belief Variance: Evidence from a Panel Study

IWASAWA, Masamune
Doshisha University

NAKANISHI, Hayato
Kanagawa University

ONOZUKA, Yuki
Otaru University of Commerce

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-049
May 2025

The Impact of Media Type on Belief Variance: Evidence from a Panel Study*

Masamune IWASAWA (Doshisha University)
Hayato NAKANISHI (Kanagawa University)
Yuki ONOZUKA (Otaru University of Commerce)

Abstract

This paper examines how traditional and nontraditional media influence people's beliefs about future states, focusing on the variance of these beliefs. Nontraditional media, such as Internet search engines, allow individuals to select information according to their interest more easily than traditional media, such as television and newspapers, which provide relatively homogeneous information. Consequently, the media type can affect not only the mean but also the variance of beliefs, resulting in greater variance among nontraditional media users compared to traditional media users. We utilize a unique panel dataset that asks respondents about their main sources of COVID-19-related information and their predictions about the end of the pandemic. Since the prediction variable in our data is categorical, we apply an interval censored fixed effects regression model. The estimation results show that the variance of the predictions is significantly smaller for traditional media users than for nontraditional media users. At the population average, the standard deviation is almost one month larger for nontraditional media users, leading to a prediction interval that is approximately 20% wider. Further analysis suggests that information sources predominantly influence infection prevention behaviors through their impact on subjective beliefs about the pandemic's end rather than perceptions of the disease risk.

Keywords: subjective beliefs, media type, belief variance, preventive behavior, COVID-19
JEL classification: D83, L82, I12

*This study is conducted as a part of the Project "Basic research for exploring ideal interventions in medicine and health" undertaken at the Research Institute of Economy, Trade and Industry (RIETI).
The draft of this paper was presented at the RIETI DP seminar. We thank participants of the RIETI DP Seminar for their helpful comments.
This study uses the micro data of the 2020 Continuing Survey on Mental and Physical Health during the COVID-19 Pandemic, which was conducted by RIETI.

1 Introduction

The impact of information provision and acquisition on behavior is a topic of significant interest across disciplines such as psychology, marketing, and economics. Individuals receive information through various channels, with media outlets such as television, newspapers, and social networking services (SNS) serving as primary sources for information on societal issues. Numerous studies have empirically examined how these media influence behavior (see, for example, DellaVigna and La Ferrara, 2015 for a review). Given the variability in the type and quality of information both across and within media, some research has investigated the heterogeneous effects depending on the type of media.

An important open question is the mechanism by which media influence behavior. To address this question, we shed light on people's subjective beliefs on future states, which serve as a vehicle of behavioral changes in economic models. Understanding this mechanism provides a unified framework for analyzing the relationship between information acquisition and various behaviors, enabling better extrapolation and prediction of behavioral outcomes. During the COVID-19 pandemic, for example, individuals with differing beliefs about the pandemic's end may have responded differently in areas such as savings, work, and infection prevention measures. If subjective beliefs are the primary channel through which media impact people's behaviors, it is critical for policymakers to maintain a healthy environment where people can critically assess information from various perspectives to make informed decisions.

Another important unresolved question is how different types of media affect outcome distributions in distinct ways. Researchers and policymakers often rely on average effects to assess media impacts due to their methodological tractability and ease of interpretation. However, if information transmitted by media shifts the distributions of outcome variables, such as utility or willingness to engage in certain behaviors, relying solely on average effects may provide an incomplete or even misleading picture. For instance, in the time of disaster, it is essential to ensure that people accurately perceive the danger and evacuate to a safe place. However, if information is not easily accessible or is contaminated with misinformation, some individuals may struggle to interpret the information appropriately, leading to a larger variance of people's beliefs on the states. A large variance of beliefs suggest that, even if people, on average, evacuate properly, a substantial amount of people may underestimate the severity of the disaster and fail to take appropriate actions, potentially resulting in substantial casualties. Thus, investigating additional distributional indicators, such as variances, is essential for a comprehensive understanding of media effects.

This study examines whether and how people's beliefs about the end of the COVID-19 pandemic, particularly in terms of variance, depend on their primary sources of information, such as traditional and nontraditional media. Traditional media is defined as providing unidirectional information, including television and newspapers, while nontraditional media encompasses interactive platforms, such as internet search engines, news apps/sites, and SNS.

Users of nontraditional media may form different predictions about the end of the pandemic compared to those who rely on traditional media. Specifically, the variance in predictions among individuals who obtain information from nontraditional media is likely to be larger than that among traditional media users for two main reasons.

First, the interactive nature of nontraditional media allows users to selectively access information that aligns with their pre-existing interests and beliefs, which vary across individuals. Gorodnichenko et al. (2021) demonstrate using Twitter data that information tends to diffuse more strongly among individuals with similar beliefs. For instance, individuals may seek out positive information when feeling optimistic and gravitate toward pessimistic content when feeling down. Indeed, Faia et al. (2024) show that people tend to prefer information consistent with their existing beliefs about the COVID-19 pandemic. Additionally, social media algorithms further reinforce users' existing views by curating content accordingly (see, e.g., Levy, 2021).

In contrast, traditional media in Japan typically presents more uniform information, irrespective of individual preferences. Given the diversity of pre-existing interests and beliefs, nontraditional media users as a whole may be exposed to a broader range of information.

Second, the reliability of information sources plays a crucial role in belief formation. For instance, Liang (2024) shows that people are less likely to react to information from uncertain sources. Moreover, people's trust in information sources varies by source, and Internet news and SNS are generally considered less reliable (see, e.g., Ali et al., 2020; Sakya et al., 2021; Uchibori et al., 2022; Buturoiu et al., 2022). Hence, information from these sources may have a weaker impact on belief formation. Consequently, the beliefs of traditional media users many exhibit lower variance due to the more reliable and homogeneous nature of the information they receive.

We utilize data from a unique panel survey to investigate the variance effects of media on predictions about the COVID-19 pandemic. This survey consists of five rounds conducted over one year, beginning in October 2020. In the initial round, respondents were asked about their primary sources of information on COVID-19, which are classified into traditional and nontraditional media categories. Additionally, each round solicits respondents' predictions regarding the pandemic's end, with options such as "pandemic has ended," "pandemic will end by March 2021," and so forth.

To handle the categorical nature of these predictions, we employ the interval censored fixed effects regression model proposed by Abrevaya and Muris (2020) for our primary analysis. This model captures variations in regressors at both the level and the variance of the latent prediction,[1] measured as the number of months from the survey period.[2] [3]

The COVID-19 pandemic provides a compelling context for studying the effects of information sources on subjective beliefs for two primary reasons. First, predictions about the pandemic's end were significantly influenced by the available information, as determining its conclusion was virtually impossible during the early stages. COVID-19 was first identified in China in December 2019 and rapidly spread globally. The World Health Organization (WHO) declared a pandemic on March 11, 2020 (see WHO, 2021). The pandemic lasted for nearly three and a half years until the WHO officially declared the end of the COVID-19 global health emergency on May 5, 2023 (see WHO, 2023). During this period, few individuals were confident about when normalcy would be restored.

The second reason is the diversity of information about COVID-19. Initially, limited knowledge about the disease led to considerable debate about its severity, even among experts. While some specialists warned of COVID-19's potential risks and the WHO declared it a public health emergency at the end of January 2020, others downplayed the threat (see, for example, Bernstein, 2020). Consequently, a wide range of information, including "fake" information, emerged (see, e.g., Moscadelli et al., 2020). This variability in information suggests that individuals may have received differing messages depending on their sources. Sakya et al. (2021) demonstrate that the accuracy of COVID-19 knowledge varies based on the reliability of news sources. Thus, individuals might have formed different predictions about the pandemic's end depending on the information sources they relied on.

Our main estimation results indicate that individuals who rely on traditional media as their primary

[1]he model assumes that the latent prediction is independent and identically distributed (i.i.d.) conditional on the fixed effects and regressors, and that predictions are serially independent. Variance is considered across individuals rather than within individuals. See Section 3 for details.

[2]Although the fixed effects model does not allow direct estimation of the coefficient for the traditional media dummy variable, it includes interactions between traditional media dummies and time-varying variables, such as prefecture-level infection rates and COVID-19 vaccination rates, as regressors.

[3]Due to the limitations of the model and data, this study does not investigate how the variance of beliefs about the end of the pandemic evolves over time. As a result, it does not contribute to the literature on the polarization effects of media, such as Levy, 2021; Faia et al., 2024.

source of information about COVID-19 exhibit a significantly smaller variance in their predictions about the pandemic's end.[4] At the population average, the standard deviation of predictions for nontraditional media users is approximately one month larger than for traditional media users, leading to a prediction interval about 20% wider for the former. In contrast, factors such as the number of infected individuals in their prefecture or personal infection with COVID-19 have similar effects on prediction levels for both groups. Our findings are robust to an alternative definition of traditional and nontraditional media, as well as to the inclusion of various controls in the regressions.

Understanding how differences in beliefs between traditional and nontraditional media users translate into behavioral differences is not straightforward. To theoretically explore these discrepancies, we employ an intertemporal utility model in which subjective beliefs influence the extent to which future utilities are considered in behavioral decision-making.

We further incorporate the perceived utility model proposed by Cohen (1984), which accounts for utilities associated with preventive behaviors. According to this model, engaging in preventive actions increases utility by alleviating fear of infection, which depends on perceptions of disease risk, while the discomfort or inconvenience associated with these actions decreases utility.

Our model suggests that if behavioral switching costs are substantial and the negative utility arising from discomfort or inconvenience grows over time, individuals may be less likely to engage in preventive behaviors when they anticipate a longer duration of the pandemic.

To examine behavioral differences between traditional and nontraditional media users, we estimate the variance effects of information sources on preventive behaviors. Our analysis, based on different preventive behaviors reported in each survey round, indicates that individuals who rely on traditional media as their primary source of information tend to exhibit more consistent behavior across various types of preventive measures.

To assess whether the effects of information sources on preventive behaviors also operate through the perception channel, we conduct an additional investigation by estimating the variance effects of information sources on the perception of COVID-19 risk, measured by the degree of fear. The results indicate that media type has no significant effect on the perception of COVID-19 risk.

These findings, combined with our main result that traditional media users exhibit smaller variance in their predictions about the pandemic's end, suggest that information sources primarily influence

[4] The substantial variance in predictions among nontraditional media users might be attributed to polarization (Levy, 2021). However, due to data limitations, we cannot directly investigate whether this variance is caused by polarization, leaving this as a topic for future research.

preventive behaviors through their effects on subjective beliefs rather than through the perception of COVID-19 risk.

Recent literature has examined the effects of information on perceptions and behaviors related to COVID-19. Ren et al. (2022) find that traditional media, compared to social media, are effective at promoting preventive behaviors among university students by enhancing perceptions of severity and efficacy. Similarly, Loomba et al. (2021) demonstrate that exposure to misinformation is associated with vaccine reluctance. Other studies highlight how different types of cable news affect behaviors related to various preventive measures (see, e.g., Ash et al., 2024, Simonov et al., 2021, Pinna et al., 2022, and Bursztyn et al., 2023). Our study extends this research by offering a potential explanation for how information affects behavior, emphasizing the role of subjective beliefs about the future states shaped by the type and quality of information.

Moreover, our study contributes to the growing body of research on how individuals form or update their beliefs based on the type and quality of information. While many studies in this area adopt experimental approaches (see, e.g., Eil and Rao, 2011, Coutts, 2019, Zimmermann, 2020, Barron, 2021, Hartzmark et al., 2021, and Liang, 2024), our work adds to the literature by providing evidence from a nonexperimental setting.

The remainder of this paper is organized as follows. Section 2 describes the data used in this study. Section 3 details the econometric model employed for the main analysis. The estimation results are presented in Section 4. Section 5 conducts an additional analysis of the behavioral effects, complementing the main analysis on the beliefs. Finally, Section 6 concludes the study.

2 Data

We use data from an internet panel survey called the Continuing Survey on Mental and Physical Health during the COVID-19 Pandemic. This survey consists of five rounds: October 27-November 6, 2020 (hereafter referred to as R1); January 19-26, 2021 (R2); April 23-May 6, 2021 (R3); July 20-27, 2021 (R4); October 20-27, 2021 (R5). This survey was conducted by the Research Institute of Economy, Trade and Industry, contracted to NTTCom Online Marketing Solutions Corporation.

Survey respondents were members of a research panel of NTTCom Online Marketing Solutions Corporation or its affiliates. The sample was designed to represent individuals aged 18 to 74 living in Japan, stratified by prefecture, gender, and age. A total of 16,642 individuals responded to the first-round survey. After excluding those who did not complete all five rounds or failed to answer the relevant questions, our final analysis sample consists of 10,133 individuals.

This panel survey includes several questions related to COVID-19. All variables used in this study are detailed in Appendix B. Below, we explain two key variables: information sources and predictions.

2.1 Information sources

In the first round of the survey, respondents were asked to report their top three main sources of information about COVID-19. We do not observe the amount or specific content of the information they were exposed to through these sources. For our analysis, we focus on their primary source.

There are 15 choices, which we categorize into two groups: traditional media and nontraditional media. Traditional media includes television (both public and commercial), newspapers, magazines, and radio. Nontraditional media includes internet search engines, news apps/sites, governmental/corporate/specialized institution websites, Facebook, Twitter, Instagram, LINE[5], academic information, other, and none.[6]

In our data, 66.1% of respondents reported that their primary source of information is traditional media. Within the traditional media group, television is the most common source: 44% of respondents cited public television, while 42% cited commercial television. Among the nontraditional media group, the largest share is attributed to internet search engines (44%), followed by news apps/sites (15%).

We briefly describe some demographic and mental health characteristics of each group. Table 1 presents the means and standard deviations for the first survey round. The proportion of women is higher in the traditional media group, and, as expected, this group is older than the nontraditional media group. Additionally, the rate of being married is higher in the traditional media group. These differences suggest that traditional media are more likely to be used by older individuals and those who spend more time at home. In contrast, there are no significant differences in years of education or household income between the two groups. Differences in place of residence exist but are not substantial.[7] In our dataset,

[5]LINE is a social media app popular in Japan.

[6]As a robustness check, we also estimate models with an alternative definition of traditional media, where governmental/corporate/specialized institution websites and academic information are included in the traditional media group. Since information from academic and government websites may be more coherent and less diffuse, this definition may affect the results. Estimation results based on this alternative definition are reported in Appendix C, but they do not significantly alter our findings.

[7]To evaluate the representativeness of our survey data, we compare it with data from the "FY2020 Survey on Usage Time of Information and Communications Media and Information Behavior," conducted by the Ministry of Internal Affairs and Communications (MIC) in January 2021. This survey employed a placement method and included 1,500 respondents, comprising men and women aged 13 to 69, who were asked to identify their most frequently used information source for

Table 1: Demographic differences by information sources

	Traditional		Nontraditional	
Female	0.48		0.43	
Age	55.78	(12.81)	48.84	(13.07)
Years of education	14.34	(1.97)	14.58	(2.01)
Married	0.69		0.55	
Household income (10,000 yen)	530.63	(291.43)	534.35	(307.23)
Average of PHQ-9	1.542	(0.845)	1.745	(1.009)
Average of GAD-7	1.320	(0.591)	1.451	(0.712)
Average of Trust	0.333	(0.391)	0.251	(0.360)
Variance of PHQ-9	0.328	(0.958)	0.473	(1.248)
Variance of GAD-7	0.145	(0.290)	0.187	(0.330)
Variance of Trust	0.087	(0.120)	0.073	(0.115)
Area				
Hokkaido	4.82		4.51	
Tohoku	6.60		6.22	
Kanto	34.85		35.95	
Chubu	17.45		15.10	
Kansai	19.04		19.34	
Chugoku	5.71		5.64	
Shikoku	2.84		3.03	
Kyusyu & Okinawa	8.69		10.21	
High density prefecture	0.342		0.375	
N	6,695		3,438	

Note: Standard deviations are displayed in the parentheses. We remove respondents whose time taken to answer the question is shorter than 1 minute or longer than 120 minutes. The definitions of variables are provided in Appendix C.

the proportion of individuals relying on traditional media is slightly higher (66% versus 62%). This discrepancy may reflect the inclusion of older individuals in our dataset, as older age groups are more likely to consume traditional media.

2.2 Predictions

In each round of the survey, respondents selected one of seven or eight options to predict when the pandemic would end. Here, the "end" of the pandemic refers to the point at which individuals can resume their pre-pandemic lifestyle without worrying about the risk of infection.[8]

current affairs. Given the differences in the targeted age groups between the MIC survey (13-69) and our dataset (18-74), a direct comparison is not feasible. However, to the best of our knowledge, no other survey captures information source related to COVID-19 in a manner directly comparable to our survey. Table 19 in Appendix D, which is constructed based on a published aggregate table, reports the averages of selected variables that can be compared with those presented in Table 1.

[8] In this definition, the "end" does not necessarily correspond to the WHO's declaration of the end of the global health emergency on May 5, 2023. The actual timing of the pandemic's end may depend on individual's subjective perceptions,

Table 2: Choices for the prediction on the end of COVID-19 pandemic

calendar date			months from survey dates		
R1 & R2	R3 & R4	R5	R1 & R3	R2 & R4	R5
Has ended	Has ended	Has ended	$(-\infty, 0]$	$(-\infty, 0]$	$(-\infty, 0]$
March 2021	September 2021	March 2022	$(0, 5]$	$(0, 2]$	$(0, 5]$
September 2021	March 2022	September 2022	(5,11]	(2,8]	(5,11]
March 2022	September 2022	March 2023	(11,17]	(8,14]	(11,17]
September 2022	March 2023	September 2023	(17,23]	(14,20]	(17,23]
March 2023	September 2023	March 2024	$(23, 29]$	$(20, 26]$	$(23, 29]$
September 2023	March 2024	Later	$(29, 35]$	$(26, 32]$	$(29, \infty)$
Later	Later		$(35, \infty)$	$(32, \infty)$	

Since the five rounds of the survey were conducted over the course of a year, the available options varied by round, as shown in Table 2. We convert the choices from calendar dates (year and month) to the number of months from the respective survey periods. For example, the first survey round was conducted from October 27 to November 6, 2020. The corresponding transformed interval choices for this round are $(-\infty, 0]$, $(0, 5]$, $(5, 11]$, $(11, 17]$, $(17, 23]$, $(23, 29]$, $(29, 35]$, $(35, \infty)$ (refer to Table 2 for the transformed intervals for other rounds).

Figure 1 displays histograms of predictions from round 1, segmented by media group. Both groups exhibit considerable variance in their predictions regarding the end of the pandemic. For instance, approximately a quarter of respondents in each group predicted that the pandemic would end within 11 to 17 months (the fourth category), while about a third estimated it would continue for nearly three more years (the last category, which is explained in the figure note). This wide variability in predictions is also observed in other datasets (see Morikawa 2020).

Although the shapes of the histograms for the two groups are similar, some differences are observable. Specifically, the histogram for the nontraditional media group is flatter compared to that of the traditional media group. Additionally, the nontraditional media group shows a greater tendency to choose the last category (which is not displayed in Figure 1). These differences persist across all survey periods.

These descriptive variations suggest that the source of information may influence how people make their predictions about the end of the pandemic. However, these observed differences are influenced by a range of factors, both observed and unobserved. To account for these factors, we apply the interval

which are unobserved. Hence, differences in the belief variance between media types may reflect differences in both predictions and perceptions regarding the end of the pandemic.

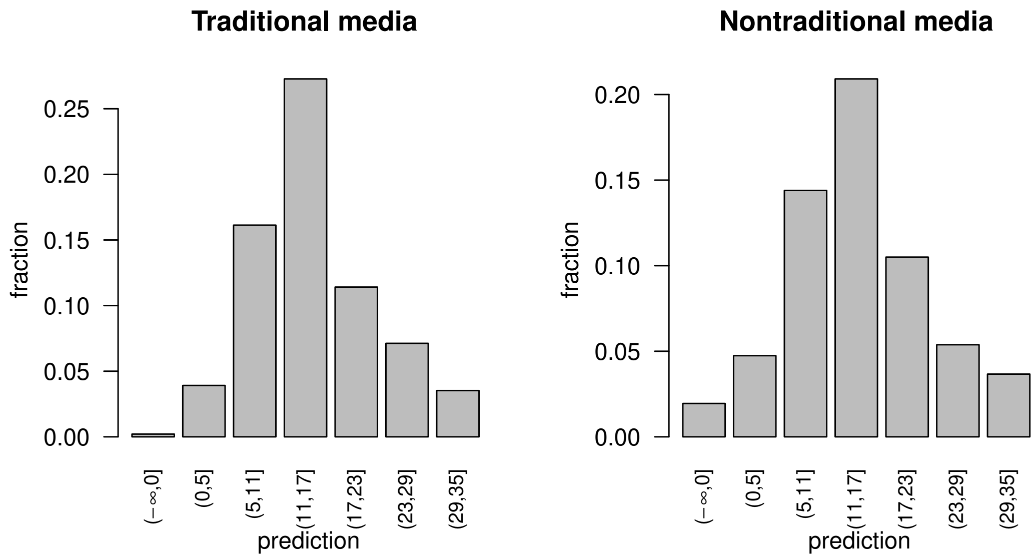

Figure 1: Predictions of the end of the pandemic by main information source in the first round. Among traditional media users, 30 percent choose the last category, and among nontraditional media users, 38 percent choose the last category.

censored regression model with fixed effects proposed by Abrevaya and Muris (2020).

3 Econometric models

Our prediction variable is categorical, indicating predictions about the end of the COVID-19 pandemic. While this variable does not provide precise numerical values for the predictions, it specifies thresholds (cutoff points) for these predictions. Consequently, we utilize the interval censored fixed effects regression model proposed by Abrevaya and Muris (2020) to analyze these beliefs.

Since the exact value of the predictions for the end of the pandemic is unobserved, we introduce a latent variable to represent it. Let $y^*_{pred,it}$ denote the latent variable such that respondent i predicts that the COVID-19 pandemic will end within $y^*_{pred,it}$ months at time t, where $t = \{\mathrm{R1}, \mathrm{R2}, \mathrm{R3}, \mathrm{R4}, \mathrm{R5}\}$. What we observe is the interval censored variable:

$$y_{pred,it} = \begin{cases} 1 & y^*_{pred,it} \leq c_{pred,1,t} \\ 2 & c_{pred,1,t} < y^*_{pred,it} \leq c_{pred,2,t} \\ \vdots & \\ J & c_{pred,J-1,t} < y^*_{pred,it}, \end{cases}$$

where $c_{pred,1,t}, \ldots, c_{pred,J-1,t}$ are the cutoff points, as listed in Table 2, corresponding to the right endpoints of the intervals.[9] For example, if t is R1 or R3, we have $c_{pred,1,t} = 0, c_{pred,2,t} = 5, \ldots, c_{pred,7,t} = 35$. Thus, the number of intervals, J, is 8.[10]

To investigate the formation of predictions, we employ an interval censored fixed effects model with heteroskedasticity. The latent variable model is specified as:

$$y^*_{pred,it} = \alpha_{pred,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{pred} - \sigma_{pred}(\boldsymbol{Z}_i)e_{it},$$

where $\alpha_{pred,i}$ represents the fixed effect, $\boldsymbol{X}_{it}$ is the vector of independent variables, $\boldsymbol{\beta}_{pred}$ is the vector of regression coefficients, e_{it} is the error term, and $\sigma_{pred}(\boldsymbol{Z}_i)$ is the error scaling function that introduces heterogeneity into the model. The error terms $e_i \equiv (e_{iR1}, e_{iR2}, e_{iR3}, e_{iR4}, e_{iR5})$ are assumed to be serially independent and follow a standard logistic distribution:

$$e_i | (\alpha_{pred,i}, \boldsymbol{X}_i) \sim \text{i.i.d. logistic},$$

conditional on $\alpha_{pred,i}$ and $\boldsymbol{X}_i \equiv (\boldsymbol{X}_{iR1}, \boldsymbol{X}_{iR2}, \boldsymbol{X}_{iR3}, \boldsymbol{X}_{iR4}, \boldsymbol{X}_{iR5})$.

In our estimation, $\boldsymbol{X}_{it}$ includes time dummies, the number of infections per 100,000 in the respondent's prefecture over the past four weeks, a COVID-19 infection dummy (equal to 1 if the respondent has ever been infected with COVID-19 and 0 otherwise), vaccination coverage (first vaccination) in the respondent's prefecture, and the interactions of the nontraditional media dummy with the above variables.

Unfortunately, information sources were only asked in the first round of the survey. In the early stages of the pandemic, people may have been particularly careful in choosing their information sources to obtain reliable information, suggesting that their preferred information sources could have changed over time. However, since the first round of the survey used in our analysis was conducted more than six months after the WHO's pandemic declaration, we assume that the respondents' information sources

[9] The latent prediction $y^*_{pred,it}$ is defined as the number of months remaining until the pandemic is predicted to end, as assessed at time t. For instance, $y^*_{pred,i1} = 3$ indicates that at round 1, respondent i predicts the pandemic will end within three months, whereas $y^*_{pred,i2} = 3$ signifies that at round 2, the same respondent predicts the pandemic will end within three months. This definition of $y^*_{pred,it}$ may be subject to debate, as the same numerical value at different times (e.g., $y^*_{pred,i1} = y^*_{pred,i2} = 3$) does not correspond to the same calendar month. To address concerns regarding the robustness of our main results to this definition, we re-estimate the interval censored regression model of predictions using an alternative definition of $y^*_{pred,it}$ that aligns with calendar dates, as detailed in Appendix C.

[10] Although the number of intervals is seven for $t = \text{R5}$, this does not affect the estimation strategy, as implied by the proof of Theorem 1 in Abrevaya and Muris (2020).

Table 3: Interval censored model of predictions: Estimates of $\boldsymbol{\beta}_{pred}$

	(1)		(2)	
Estimates of $\boldsymbol{\beta}_{pred}$				
Vaccination rate	4.211	(5.233)	2.139	(5.240)
Number of infections	0.011	(0.011)	0.008	(0.026)
Ever infected	-1.279	(1.925)	1.431	(1.848)
Round 2 dummy	9.640	(4.584)	12.912	(4.682)
Round 3 dummy	8.274	(2.891)	10.282	(3.096)
Round 4 dummy	2.741	(3.602)	4.723	(3.841)
Round 5 dummy	-3.868	(4.638)	-1.828	(4.918)
Nontraditional media ×				
Vaccination rate	6.372	(9.727)	6.216	(9.579)
Number of infections	0.036	(0.024)	0.014	(0.047)
Ever infected	-0.397	(2.681)	-0.636	(2.632)
Round 2 dummy	-11.221	(8.693)	-7.001	(8.074)
Round 3 dummy	-4.871	(6.717)	-3.218	(6.588)
Round 4 dummy	-8.288	(6.983)	-7.413	(6.943)
Round 5 dummy	-7.266	(8.910)	-7.190	(8.707)

Note: Standard errors are displayed in the parentheses. The definitions of variables are provided in Appendix C.

remained constant throughout the study period.

Following Abrevaya and Muris (2020), the error scaling function is specified as $\sigma_{pred}(\boldsymbol{Z}_i) = \exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{pred})$, where $\boldsymbol{\gamma}_{pred}$ is a parameter vector and $\boldsymbol{Z}_i$ includes a constant and the nontraditional media dummy. A positive coefficient of the nontraditional media user dummy in this error scaling function indicates that the variance of the belief distribution among nontraditional media users is larger than that among traditional media users.

As a robustness check, we also estimate a model in which age, years of education, a female dummy, a marital status dummy, the high population density prefecture dummy, and mental health characteristics are added to $\boldsymbol{Z}_i$.[11]

The parameters $\theta_{pred} \equiv (\boldsymbol{\beta}_{pred}, \boldsymbol{\gamma}_{pred})'$ are estimated using the composite likelihood estimator proposed by Abrevaya and Muris (2020). Appendix A provides a detailed description of the estimation strategy.

4 Estimation results

The estimation results for $\boldsymbol{\beta}_{pred}$ are presented in Table 3. As explained in Appendix A, standard errors are obtained using bootstrapping with 200 resamples. Column (1) includes only a constant and the nontraditional media dummy in $\boldsymbol{Z}_i$ of the error scaling function, while column (2) includes additional control variables in $\boldsymbol{Z}_i$ (see Table 4 for the estimates of $\boldsymbol{Z}_i$). The estimates of $\boldsymbol{\beta}_{pred}$ for the latent prediction model indicate no statistically significant effects for most variables. Furthermore, all interaction terms involving the nontraditional media dummy are insignificant. This suggests that the effects of the variables on average prediction do not differ by media group.

The dummy variables for rounds 2 and 3 have a significant positive effect on the prediction of the end of the COVID-19 pandemic. Compared to the survey period of round 1, people in round 2 and round 3 tend to believe that the pandemic will last longer, by nine months and eight months, respectively. These results may be explained by the epidemic nature of the virus. During the COVID-19 pandemic, the Japanese government declared a state of emergency four times.[12]The surveys for rounds 2, 3, and 4 were conducted were conducted during these declarations.

Another feature of the time dummies is that the estimates decrease for the dummies of later rounds. This suggests that, as time passes, people tend to believe that the pandemic will end in shorter periods. All interaction terms between the time dummies and the nontraditional media dummy remain insignificant. The estimation results are similar in column (2).

The estimates of $\boldsymbol{\gamma}_{pred}$ for the heterogeneity of the prediction variance are reported in Table 4. Nontraditional media has a significant positive effect on the variance of predictions about the end of the COVID-19 pandemic.[13] The variance of the predictions for respondents who rely on traditional media as their primary source of COVID-19 information is smaller than that for nontraditional media users.

Due to the limitations of the econometric model, we implicitly assume that the variance is constant over time. Thus, the result does not suggest that the variation in predictions over time is smaller for traditional media users. Rather, it indicates that the variation in predictions among traditional media users is smaller than that among nontraditional media users.

[11]Time-varying variables included in $\boldsymbol{X}_{it}$ cannot be added to $\boldsymbol{Z}_i$ due to the limitations of the estimation method.

[12]The four periods of declaration were April 7–May 25, 2020, January 8–March 21, 2021, April 25–June 20, 2021, and July 12–September 30, 2021. During these periods, people were asked to refrain from going out, and restrictions were imposed on the use of public facilities and restaurants, the holding of events, and other activities. (`https://corona.go.jp`).

[13]We confirmed that the estimate of the nontraditional media dummy for each period is positive, which refers to the value before the minimum distance procedure.

Table 4: Interval censored model of predictions: Estimates of γ_{pred}

	(1)		(2)	
Estimates of γ_{pred}				
Constant	1.748	(0.009)	2.527	(0.199)
Age			-0.005	(0.001)
Education			-0.036	(0.009)
Female			-0.089	(0.027)
Married			0.062	(0.024)
High density prefecture			-0.007	(0.023)
Average of GAD–7			-0.078	(0.058)
Average of PHQ–9			0.064	(0.044)
Average of Trust			-0.055	(0.039)
Variance of GAD–7			0.172	(0.070)
Variance of PHQ–9			0.004	(0.044)
Variance of Trust			0.147	(0.108)
Nontraditional media	0.193	(0.016)	0.171	(0.025)
Estimates of $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})$				
Nontraditional media	6.966		6.888	
Traditional media	5.743		5.807	
Difference in $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})$ between users of nontraditional and traditional media	1.223		1.081	
95% prediction interval of $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})e_{it}$				
Nontraditional media	[-25.530, 25.521]		[-25.244, 25.237]	
Traditional media	[-21.047, 21.040]		[-21.283, 21.277]	

Note: Standard errors are displayed in the parentheses. $\bar{\boldsymbol{Z}}_i$ denotes the average of $\boldsymbol{Z}_i$. The definitions of variables are provided in Appendix C.

At the population average, the estimated standard deviation for traditional media users is 1.22 months smaller than that for nontraditional media users in column (1), resulting in a prediction interval that is approximately 20% wider for nontraditional media users.

Column (2) includes demographic and mental health variables to control for sources of heteroskedasticity. All of the added demographic variables, are significant, yet the coefficient on the nontraditional media dummy remains positive and significant. The difference in the estimated standard deviation of predictions at the population mean is similar to that in column (1).

5 Effects on behavior

We have shown that information sources influence the variance of predictions, while these predictions are subjective beliefs about the future state of the COVID-19 pandemic. This finding suggests that information sources may shape COVID-19-related behaviors through the belief channel. However, the mechanism through which differences in belief translate into behavioral differences remains unclear.

In addition to beliefs, perceptions–such as the perceived severity of COVID-19 and the perceived efficacy of preventive behaviors–are also recognized in the literature (Ren et al., 2022; Bursztyn et al., 2023) as key determinants of behavior. Both beliefs and perceptions can be shaped by information sources. In particular, as suggested by the estimation results in the previous section, their variance may differ between users of traditional and nontraditional media.

The following subsection presents an intertemporal utility model to theoretically examine the relationship between beliefs and behavior. We then estimate the effect of information sources on infection prevention behavior and perceptions of COVID-19 risk to show suggestive evidence for the mechanism through which information sources influence behavior.

5.1 Model

We examine the mechanisms through which beliefs and perceptions influence preventive behaviors within the framework of the expected utility model. We then explore how differences in the variance of beliefs and perceptions may lead to variations in preventive behaviors.

Following the perceived utility model proposed by Cohen (1984) for health-related preventive behaviors,[14] utility is assumed to consist of two components:

$$u_{it} = u_{it}^a + u_{it}^u,$$

where u_{it}^a and u_{it}^u represent the utility in anticipation and the utility in use, respectively, for individual i at time t.

Utility in anticipation refers to the utility gained from a reduction or increase in anxiety experienced in the present. For instance, behaviors such as mask-wearing and hand washing can reduce anxiety by

[14] In economics, utility is typically a function of economic behaviors such as consumption and labor. Explaining non-economic behaviors is less common. In contrast, utility models are frequently used in other fields such as marketing, health, and epidemiology (see Moorman and Matulich, 1993 for a review of health behavior models).

lowering the perceived risk of infection and by aligning with prevailing social norms.

In contrast, utility in use pertains to the utility derived directly from the behavior itself. For example, while these preventive behaviors may contribute to health safety, they can also cause discomfort, such as irritation from wearing masks or rough skin from frequent hand washing, leading to negative utility. Although the sign of utility in use is not inherently negative, we simplify the analysis by assuming it to be negative. This assumption is based on the observation that most COVID-19-related preventive behaviors examined in this study are actions unlikely to occur in the absence of the pandemic. Additionally, we posit that utility in use diminishes over time, implying that the level of disutility increases with prolonged engagement in such preventive behaviors.

We assume that utility in anticipation is a function of perceptions. Since utility in anticipation is related to the level of fear, it is influenced by perceived severity and efficacy. For instance, physical distancing and hand washing may not reduce anxiety for individuals who do not fear infection. In contrast, utility in use pertains to the utility derived directly from the behavior itself, which is not influenced by the perceived severity of infection and the efficacy of the behavior.

Let us explicitly model COVID-19 preventive behaviors. Utility in anticipation is a function of behavior and perception, i.e., $u^a_{it} = u_a(c_{it}, a_{it})$, where c_{it} represents the implementation of infection prevention behavior and a_{it} denotes the perceived severity of COVID-19. Utility in use is a function of the history of the behavior, i.e., $u^u_{it} = u_u(c_{it}, c_{it-1}, \dots)$.[15]

To simplify the model, we restrict behavior to $c_{it} = 1$ (taking the behavior) and $c_{it} = 0$ (not taking the behavior). Perception a_{it} is defined as $a_{it} = 0$ if the pandemic ends at time t, and $a_{it} > 0$ if it does not. Since c_{it} represents COVID-19 preventive behaviors, it follows that $c_{it} = 0$ when the pandemic is over, i.e., $a_{it} = 0$. We also assume that a_{it} takes one of two values: a high value a^H or a low value a^L, where $a^H > a^L > 0$.

Furthermore, we specify that $u_a(0, a_{it}) = -a_{it} \leq 0$ and $u_a(1, a_{it}) = r$, where $-a_L < r < 0$. This implies that the utility in anticipation during the pandemic is negative in the absence of preventive behavior due to the individual's fear, and this fear can be reduced to r if infection prevention behavior is adopted. Similarly, we assume that $u_u(c_{it} = 0, \cdot) = 0$, which indicates that the utility in use at time t is zero when no behavior is taken at time t, regardless of past behaviors.

Subjective belief about the end of the pandemic for individual i is denoted by T_i. The "end" refers to

[15]Demographic variables can be included in both utility functions. However, we omit them to avoid complicated notation.

the point when they can resume their pre-pandemic lifestyle without worrying about the risk of infection. Thus, they expect $a_{it} = 0$ after period T_i. We assume that individuals determine the sequence of their preventive behaviors based on their current perceived severity of COVID-19, i.e., a_{it} remains constant for $t \leq T_i$. Since no preventive actions are taken against COVID-19 after the pandemic ends, future utility can be expressed as

$$u_{it} = \begin{cases} u_{it} & \text{if } \ t \leq T_i \\ 0 & \text{if } \ t > T_i. \end{cases}$$

The intertemporal utility[16] for those who believe that the pandemic will end at time T_i is given by

$$u_{it} + \sum_{k=t+1}^{T_i} \delta_k u_{ik}, \tag{1}$$

where $\delta_t \leq 1$ is the discount factor. An individual chooses a particular sequence of preventive behaviors, $\boldsymbol{c}_i = (c_{it}, c_{it+1}, \dots)$, that maximizes equation (1).

Preventive behaviors such as hand washing and physical distancing have been shown to be habitual (Hussam et al., 2022; Zhang et al., 2022). Since the initial round of the survey used in our analysis was conducted more than six months after the WHO's declaration of the pandemic, COVID-19-related preventive measures were likely habitual at the time of the survey. Deviating from established habits to adopt different behaviors incurs substantial costs (Allcott and Rogers, 2014). Thus, we assume that individuals will not change their behavior unless their anticipation a_{it} and/or belief T_i are updated.

The intertemporal utility model suggests that both beliefs and perceptions are key determinants of behavioral decisions. Beliefs influence the extent to which future expected utilities are considered, while perceptions affect the magnitude of utility through anticipation effects.

Since deviating from established habits to adopt new behaviors entails substantial costs (Allcott and Rogers, 2014) and the disutility of use increases over time, individuals who expect the pandemic to persist longer are less likely to engage in preventive behaviors. Furthermore, because $u_a(0, a^H) < u_a(0, a^L)$, a greater level of fear leads to higher engagement in preventive behaviors.

These insights suggest that greater diversity in perceptions and beliefs results in more variation in preventive behaviors. The main estimation results in Section 4 indicate that the variance of beliefs is

[16] We consider a simple model in which the subjective belief about the future state is either "during the pandemic" or "after the pandemic". Additionally, we assume that the perceived severity of COVID-19 remains constant during the pandemic. Thus, once individual i has formed their subjective belief T_i and perception a_{it}, intertemporal utility is determined without any uncertainty.

Table 5: List of infection prevention actions

Numbering	Infection prevention actions
(I)	I do not go to poorly ventilated places.
(II)	I do not go to places where there are many people.
(III)	I do not talk or speak in close proximity to other people.
(IV)	I wear a mask.
(V)	I wash my hands.
(VI)	I sanitize my hands.
(VII)	I change my clothes frequently.
(VIII)	I gargle.
(IX)	I disinfect my belongings.
(X)	I keep distance from people when I go out.
(XI)	I refrain from visiting hospitals and clinics.
(XII)	I stay home as much as possible.

smaller among traditional media users. From this perspective, traditional media users are expected to exhibit more similar preventive behaviors compared to nontraditional media users.

5.2 Estimation results

The intertemporal utility model in the previous subsection demonstrates that traditional media users are expected to exhibit similar preventive behaviors compared to nontraditional media users, which is examined in this subsection.

The theoretical model also suggests that perceptions are key determinants of behavioral decisions. Accordingly, we conduct an additional investigation by estimating the variance effects of information sources on the perception of COVID-19 risk, measured by the degree of fear.[17]

The effects of information sources on preventive behaviors are presented in Subsection 5.2.1, while their effects on perceptions of COVID-19 risk are presented in Subsection 5.2.2.

5.2.1 Infection prevention behaviors

Each round of the survey asks respondents about their implementation of the COVID-19 infection prevention behaviors listed in Table 5. Recent studies increasingly suggest that information sources influence individual's COVID-19 preventive behaviors (see, e.g., Ash et al., 2024, Pinna et al., 2022, and Simonov et al., 2021). As a complementary investigation using our dataset, we estimate the effects

[17] The focus of this study is to examine the effects of media type. Therefore, direct investigations of the effects of beliefs on behavior and the effects of perceptions on behavior are beyond the scope of this study.

of information sources—whether traditional or nontraditional—on these preventive behaviors.

We estimate two models. The first model is a linear probability model with fixed effects, where the outcome variables are dummy variables for preventive behaviors. The explanatory variables include control variables and their interactions with the nontraditional media dummy. The second model is a regression model for the variance of preventive behaviors, in which the outcome variable is the variance of preventive behaviors. The variances are estimated using the residuals from the fixed effects estimates of the first model, which are then regressed on the nontraditional media dummy and other controls. Details of these models are provided in Appendix E.

The estimation results of the first model are presented in Tables 6 and 7, while those of the second model are reported in Tables 8 and 9. Most interaction terms of the nontraditional media dummy are statistically insignificant in Tables 6 and 7, indicating that the effects of explanatory variables on the level of preventive behaviors do not vary significantly by information source.

In contrast, Tables 8 and 9 show that the coefficient on the nontraditional media dummy is positive and significant for the variance of infection prevention behaviors in half of the cases, specifically from (I) to (VI). Meanwhile, the coefficients for behaviors (VII) to (XI) are not statistically significant. In the case of (XII), the coefficient is negative and significant.

The results indicate that, overall, the variance of a wide variety of preventive behaviors is larger among nontraditional media users than among traditional media users. In other words, given that most people in Japan have followed the government guidelines on preventive behaviors (see Table 13 in Appendix B), the proportion of individuals who do not engage in preventive behaviors is higher among those who primarily obtain information about COVID-19 from nontraditional media.

5.2.2 Perception of COVID-19 risk

We further estimate the effects of information sources on perceptions of COVID-19 risk. Although we do not explicitly decompose the differences in behavior between traditional and nontraditional media users into prediction and perception channels, the estimation in this subsection may provide some insight.

In this study, we measure perception using fear related to COVID-19. Each round of our survey includes questions assessing COVID-19-related fear, from which we construct an index called the FCV-19S (Ahorsu et al., 2022). The FCV-19S is a seven-item self-report questionnaire designed to measure the extent to which an individual fears COVID-19. FCV-19S scores range from 7 to 35, with higher

Table 6: Linear probability model of infection prevention behaviors

	(I)	(II)	(III)	(IV)	(V)	(VI)
Vaccination rate	0.055	0.063	0.053	0.000	0.035	0.068
	(0.061)	(0.060)	(0.070)	(0.028)	(0.037)	(0.059)
Infected	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	-0.018	-0.066	-0.034	-0.002	-0.031	-0.046
	(0.030)	(0.036)	(0.037)	(0.032)	(0.031)	(0.035)
Round 2 dummy	0.001	0.026	0.054	0.005	-0.004	0.010
	(0.006)	(0.006)	(0.007)	(0.003)	(0.004)	(0.005)
Round 3 dummy	0.025	0.022	0.057	0.007	-0.001	0.018
	(0.005)	(0.005)	(0.006)	(0.002)	(0.003)	(0.005)
Round 4 dummy	-0.002	-0.005	0.032	0.005	-0.019	-0.011
	(0.028)	(0.027)	(0.032)	(0.012)	(0.016)	(0.027)
Round 5 dummy	-0.019	-0.034	0.006	0.008	-0.033	-0.030
	(0.052)	(0.051)	(0.060)	(0.024)	(0.031)	(0.051)
Nontraditional media ×						
Vaccination rate	-0.034	0.016	-0.015	-0.055	-0.095	0.049
	(0.117)	(0.111)	(0.132)	(0.054)	(0.069)	(0.103)
Infected	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	-0.053	-0.050	-0.023	-0.062	-0.028	-0.022
	(0.052)	(0.051)	(0.058)	(0.046)	(0.048)	(0.053)
Round 2 dummy	0.009	0.004	0.006	0.003	0.005	0.001
	(0.012)	(0.011)	(0.013)	(0.006)	(0.007)	(0.010)
Round 3 dummy	0.014	0.017	0.007	0.000	-0.002	0.004
	(0.010)	(0.010)	(0.012)	(0.005)	(0.006)	(0.009)
Round 4 dummy	0.024	0.001	0.006	0.023	0.040	-0.015
	(0.053)	(0.050)	(0.059)	(0.024)	(0.031)	(0.046)
Round 5 dummy	0.041	-0.001	0.028	0.047	0.081	-0.036
	(0.101)	(0.095)	(0.113)	(0.046)	(0.059)	(0.088)

Note: Standard errors are displayed in the parentheses. Outcome variables are provided in Table 5.

Table 7: Linear probability model of infection prevention behaviors

	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Vaccination rate	0.184	0.046	0.034	0.096	0.136	0.089
	(0.087)	(0.073)	(0.086)	(0.073)	(0.090)	(0.086)
Infected	0.000	0.000	0.001	0.000	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.074	0.003	0.032	-0.023	0.034	0.024
	(0.048)	(0.043)	(0.049)	(0.042)	(0.047)	(0.043)
Round 2 dummy	0.024	0.018	0.013	0.027	0.032	0.136
	(0.008)	(0.007)	(0.008)	(0.007)	(0.009)	(0.008)
Round 3 dummy	0.046	0.011	0.046	0.035	0.005	0.085
	(0.007)	(0.006)	(0.007)	(0.006)	(0.008)	(0.007)
Round 4 dummy	-0.016	-0.024	0.029	-0.016	-0.090	0.027
	(0.039)	(0.033)	(0.039)	(0.033)	(0.041)	(0.038)
Round 5 dummy	-0.117	-0.030	0.009	-0.060	-0.156	-0.035
	(0.075)	(0.062)	(0.074)	(0.063)	(0.077)	(0.073)
Nontraditional media ×						
Vaccination rate	-0.400	0.154	-0.003	-0.169	-0.102	0.053
	(0.153)	(0.131)	(0.149)	(0.129)	(0.168)	(0.146)
Infected	0.001	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Experience	-0.079	-0.092	-0.017	-0.014	-0.123	-0.020
	(0.067)	(0.060)	(0.066)	(0.058)	(0.070)	(0.068)
Round 2 dummy	-0.013	0.011	0.018	-0.005	0.020	-0.018
	(0.014)	(0.012)	(0.014)	(0.013)	(0.017)	(0.015)
Round 3 dummy	-0.001	0.001	0.007	0.000	0.020	-0.012
	(0.013)	(0.011)	(0.013)	(0.011)	(0.015)	(0.013)
Round 4 dummy	0.175	-0.061	0.008	0.078	0.063	-0.046
	(0.069)	(0.059)	(0.067)	(0.058)	(0.076)	(0.066)
Round 5 dummy	0.358	-0.126	0.029	0.161	0.117	-0.042
	(0.131)	(0.112)	(0.128)	(0.110)	(0.144)	(0.125)

Note: Standard errors are displayed in the parentheses. Outcome variables are provided in Table 5.

Table 8: Regression model for the variance of prevention behaviors

	(I)	(II)	(III)	(IV)	(V)	(VI)
Intercept	0.135	0.107	0.137	0.031	0.043	0.068
	(0.011)	(0.010)	(0.012)	(0.006)	(0.007)	(0.010)
Nontraditional media	0.008	0.006	0.006	0.008	0.005	0.008
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Married	-0.004	-0.001	-0.004	-0.003	-0.004	-0.011
	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
Age	-0.001	-0.001	-0.001	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.034	-0.023	-0.020	-0.011	-0.016	-0.023
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Education	-0.001	-0.001	-0.001	0.000	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
Density	-0.005	0.000	-0.011	-0.001	-0.005	-0.007
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Average of GAD–7	-0.011	-0.005	-0.018	0.000	-0.004	-0.014
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)
Average of PHQ–9	0.011	0.007	0.016	0.000	0.004	0.011
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Average of Trust	0.008	0.014	0.016	-0.002	0.002	0.003
	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
Variance of GAD–7	0.013	0.018	0.017	0.019	0.016	0.016
	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Variance of PHQ–9	0.000	0.000	-0.002	0.001	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Variance of Trust	0.010	0.016	0.019	0.009	0.009	0.006
	(0.009)	(0.009)	(0.010)	(0.004)	(0.006)	(0.008)

Note: Standard errors are displayed in the parentheses. Outcome variables are provided in Table 5.

Table 9: Regression model for the variance of prevention behaviors

	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Intercept	0.101	0.108	0.159	0.137	0.150	0.121
	(0.012)	(0.012)	(0.011)	(0.011)	(0.013)	(0.012)
Nontraditional media	-0.001	0.005	0.002	0.005	-0.002	-0.007
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Married	0.009	-0.001	0.015	0.000	0.004	0.014
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Age	0.000	0.000	-0.001	-0.001	-0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.004	-0.014	-0.003	-0.026	-0.006	-0.018
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Education	0.000	-0.002	-0.001	-0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Density	0.009	-0.012	-0.011	-0.006	0.003	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Average of GAD–7	0.020	-0.004	-0.018	-0.019	-0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Average of PHQ–9	-0.014	0.003	0.016	0.017	0.000	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Average of Trust	-0.012	-0.001	0.016	0.015	-0.004	0.011
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Variance of GAD–7	0.023	0.019	0.017	0.015	0.021	0.017
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Variance of PHQ–9	0.007	0.003	-0.002	0.000	0.000	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Variance of Trust	0.054	0.023	0.019	0.029	0.030	0.019
	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)

Note: Standard errors are displayed in the parentheses. Outcome variables are provided in Table 5.

Table 10: Ordered logit model for perception: Estimates ofv $\boldsymbol{\beta}_{per}$ and thresholds

	Estimate	Std. Error
Estimates of $\boldsymbol{\beta}_{per}$		
Vaccination rate	-0.031	(0.233)
Number of infections	0.001	(0.001)
Ever infected	0.126	(0.105)
Round 2 dummy	0.221	(0.022)
Round 3 dummy	0.174	(0.020)
Round 4 dummy	0.044	(0.105)
Round 5 dummy	-0.025	(0.200)
Nontraditional media ×		
Vaccination rate	-0.009	(0.431)
Number of infections	0.001	(0.001)
Ever infected	-0.107	(0.167)
Round 2 dummy	-0.088	(0.041)
Round 3 dummy	-0.095	(0.037)
Round 4 dummy	-0.031	(0.194)
Round 5 dummy	-0.028	(0.369)
Estimates of thresholds		
between 1 and 2	-1.043	(0.539)
between 2 and 3	-0.045	(0.540)
between 3 and 4	1.029	(0.542)

Note: The explained variable is a ordered variable based on the value of the FCV-19S scale (takes the largest value 4 if a respondent is afraid of COVID-19 and 1 if the respondent is not afraid of COVID-19).

scores indicating progressively higher levels of fear associated with COVID-19.

We apply the estimator for the ordered logit model with fixed effects developed by Muris (2017). The outcome variables takes the value 1 if the FCV-19S score is 7–15, 2 if the FCV-19S score is 16–20, 3 if the FCV-19S score is 21–25, and 4 if the FCV-19S score is 26–35.[18] The latent continuous value of FCV-19S, denoted by $y^*_{per,it}$ is modeled as $y^*_{per,it} = \alpha_{per,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{per} + \exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})\varepsilon_{it}$, where $\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})$ is the error scaling function that introduces heterogeneity. The parameters $\boldsymbol{\beta}_{per}$ and $\boldsymbol{\gamma}_{per}$ are estimated using the composite likelihood estimator that accounts for the likelihood function around selected thresholds.[19] The details of the estimator are explained in Appendix F.

[18]Theoretically, it is possible to apply Muris (2017) to the raw FCV-19S scores. However, such an estimation procedure requires estimating 28 (= 35−7) thresholds in addition to regression coefficients, which would be computationally intensive. To reduce estimation complexity, we define the explained variable for perception to take values 1,2,3 and 4.

[19]Tables 10 and 11, present the results obtained using randomly selected thresholds generated as 10 series of permutation of 1–2–2–2–3. The robustness of threshold selection is confirmed by estimating the model with an alternative set of thresholds, as described in Appendix F.

Table 11: Ordered logit model for perception: Estimates of $\boldsymbol{\gamma}_{per}$

	Estimate	Std. Error
Estimates of $\boldsymbol{\gamma}_{per}$		
Nontraditional media	0.033	(0.020)
Marriage	0.065	(0.020)
Age	-0.010	(0.001)
Female	-0.140	(0.018)
Education	-0.031	(0.003)
High density prefecture	0.003	(0.019)
Average of PHQ-9	-0.036	(0.024)
Average of GAD-7	0.009	(0.031)
Average of Trust	0.015	(0.026)
Variance of PHQ-9	0.047	(0.012)
Variance of GAD-7	0.155	(0.037)
Variance of Trust	0.264	(0.078)
Estimates of $\exp(\bar{\boldsymbol{Z}}_i \gamma_{per})$		
Nontraditional media	0.387	
Traditional media	0.374	
Difference in $\exp(\bar{\boldsymbol{Z}}_i \gamma_{per})$ between users of nontraditional and traditional media	0.013	
95% prediction interval of $\exp(\bar{\boldsymbol{Z}}_i \gamma_{per})\varepsilon_{it}$		
Nontraditional media	[-1.417, 1.417]	
Traditional media	[-1.372, 1.372]	

Note: The explained variable is a ordered variable based on the value of the FCV-19S scale (takes the largest value 4 if a respondent is afraid of COVID-19 and 1 if the respondent is not afraid of COVID-19). $\bar{\boldsymbol{Z}}_i$ denotes the average of $\boldsymbol{Z}_i$.

Estimation results are presented in Tables 10 and 11. The estimates of β_{per} show no statistically significant effects for most variables. However, the dummy variables for rounds 2 and 3 have significant positive effects on perception, measured as fear of COVID-19. Compared to the survey period of round 1, people tended to fear COVID-19 more in rounds 2 and 3. Furthermore, the interaction terms between the round 2 and 3 dummies and the nontraditional media dummy have significant negative effects on fear of COVID-19. This suggests that, compared to nontraditional media users, traditional media users are more likely to experience higher levels of fear related to COVID-19.

The estimates of γ_{per} for the heterogeneity of prediction variance indicate statistically significant effects for several variables. For example, marriage, variance of PHQ-9, GAD-7, and trust have statistically significant positive effects, while age, female, and education have statistically significant negative effects.

With regard to the nontraditional media dummy, the point estimate is negative, which is consistent with the results in Table 4, where the explained variable is the belief about the end of the COVID-19 pandemic. However, the estimate is not statistically significant, indicating that the effects of media on the variance of COVID-19-related fear do not differ by information source.

According to our model in Subsection 5.1, this estimation result suggests that the primary channel through which traditional and nontraditional information sources influence individual's infection prevention behavior is their prediction of the end of the pandemic. While previous studies have largely focused on risk perception as a determinant of behavior and disease-related responses, subjective belief may play a more crucial role than risk perception.

6 Conclusion

We are constantly exposed to information. Based on the information we receive, we form beliefs about future states and take various actions. In recent years, the role of information is attracting significant attention, leading to a growing body of research examining the effects of different types and qualities of information on belief formation in experimental settings.

This study estimates the effects of information sources on individual's subjective beliefs by leveraging a globally significant event: the COVID-19 pandemic. We apply an interval censored fixed effects regression model to a unique panel dataset that asks respondents about their primary sources of information on COVID-19 and their predictions about the end of the pandemic. Our objective is to determine whether individual's subjective beliefs are influenced differently by traditional and nontraditional media.

Our results show that the variance of predictions about the end of the pandemic is significantly smaller among individuals who rely on traditional media as their primary source of information. At the population average, the difference in standard deviation between nontraditional and traditional media users is almost one month, leading to a prediction interval approximately 20% wider for nontraditional media users. We also provide evidence that individuals adopt COVID-19 preventive behaviors differently depending on their main sources of information and that information sources predominantly influence preventive behaviors through their impact on subjective beliefs rather than perceptions.

Although we cannot observe the specific content individuals encounter or how it differs between nontraditional and traditional media users, our estimation results align with existing findings that nontraditional media enable individuals to acquire information based on their preferences.

Our findings imply three important points. First, policymakers should maintain a healthy environment where people can critically assess information. Although subjective beliefs about future states can affect a wide variety of behaviors, future states are inherently difficult to predict. Hence, ensuring such an environment is essential for enabling individuals to make informed decisions. Individuals may form misbeliefs, particularly when exposed to information platforms contaminated with "fake" content or misinformation.

Second, policymakers may need to adopt different strategies depending on the type of media to maintain a healthy information environment. For instance, policies aimed at facilitating the identification of misinformation, such as fact-checking initiatives, and increasing the transparency of information sources may be effective for nontraditional media. Enhancing the transparency of social media and search engine algorithms, and exposing individuals to a diverse range of information—not only content aligned with their preferences but also opposing perspectives— may be another effective policy.

Third, the way information is disseminated may need to differ depending on the type of media. Given that nontraditional media offer access to a much larger volume of content, simply conveying information in the same manner as in traditional media may be insufficient, as it can easily be buried among competing messages and fail to reach a broad population. Increasing the frequency of exposure or targeting "hub" individuals, who are highly connected within social networks, may be more effective in such environments.

The effects of media type on people's belief may change over time. On the one hand, advances in information technology could increase access to diverse information, potentially increasing the variance in people's beliefs. On the other hand, the rise of generative AI tools, such as ChatGPT, may change how people consume information. For instance, people might rely more on algorithmic summaries than on active searching. This could reduce the variance by narrowing exposure to alternative views. The effect of media on people's beliefs should continue to be examined, and strategies for maintaining a healthy information environment and approaches for conducting information will need to adopt as circumstances evolve over time.

Acknowledgments: This work was supported by JSPS KAKENHI Grant Numbers 23H00805; 24K04826. We thank Yifan Gong, Yuko Mori, Shosei Sakaguchi, and the participants at 2023 JEA Autumn meeting, Kansai Labor Workshop, and 2024 AMES-E/SE, and a workshop at RIETI for their useful and constructive comments.

References

Abrevaya, J., Muris, C., 2020. Interval censored regression with fixed effects. Journal of Applied Econometrics 35, 198–216.

Ahorsu, D.K., Lin, C.Y., Imani, V., Saffari, M., Griffiths, M.D., Pakpour, A.H., 2022. The fear of covid-19 scale: development and initial validation. International Journal of Mental Health and Addiction 20, 1537–1545.

Ali, S.H., Foreman, J., Tozan, Y., Capasso, A., Jones, A.M., DiClemente, R.J., 2020. Trends and predictors of covid-19 information sources and their relationship with knowledge and beliefs related to the pandemic: Nationwide cross-sectional study. JMIR Public Health Surveill 6, e21071.

Allcott, H., Rogers, T., 2014. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. American Economic Review 104, 3003–3037.

Ash, E., Galletta, S., Hangartner, D., Margalit, Y., Pinna, M., 2024. The effect of fox news on health behavior during covid-19. Political Analysis 32, 275–284.

Barron, K., 2021. Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains? Experimental Economics 24, 31–58.

Bernstein, L., 2020. The washington post: Get a grippe, america. the flu is a much bigger threat than coronavirus, for now. URL: https://www.washingtonpost.com/health/time-for-a-reality-check-america-the-flu-is-a-much-bigger-threat-than-coronavirus-for-now/2020/01/31/46a15166-4444-11ea-b5fc-eefa848cde99_story.html. accessed on April 17, 2023.

Bursztyn, L., Rao, A., Roth, C., Yanagizawa-Drott, D., 2023. Opinions as facts. The Review of Economic Studies 90, 1832–1864.

Buturoiu, R., Corbu, N., Oprea, D.A., Botan, M., 2022. Trust in information sources during the covid-19 pandemic. a romanian case study. Communications 47, 375–394.

Cohen, D.R., 1984. Utility model of preventive behaviour. Journal of Epidemiology & Community Health 38, 61–65.

Coutts, A., 2019. Good news and bad news are still news: experimental evidence on belief updating. Experimental Economics 22, 369–395.

DellaVigna, S., La Ferrara, E., 2015. Chapter 19 - economic and social impacts of the media, in: Anderson, S.P., Waldfogel, J., StrÃ¶mberg, D. (Eds.), Handbook of Media Economics. North-Holland. volume 1 of *Handbook of Media Economics*, pp. 723–768. URL: https://www.sciencedirect.com/science/article/pii/B978044463685000019X, doi:https://doi.org/10.1016/B978-0-444-63685-0.00019-X.

Eil, D., Rao, J.M., 2011. The good news-bad news effect: Asymmetric processing of objective information about yourself. American Economic Journal: Microeconomics 3, 114–38. URL: https://www.aeaweb.org/articles?id=10.1257/mic.3.2.114, doi:10.1257/mic.3.2.114.

Faia, E., Fuster, A., Pezone, V., Zafar, B., 2024. Biases in Information Selection and Processing: Survey Evidence from the Pandemic. The Review of Economics and Statistics 106, 829–847. URL: https://doi.org/10.1162/rest_a_01187, doi:10.1162/rest_a_01187, arXiv:https://direct.mit.edu/rest/article-pdf/106/3/829/2373201/rest_a_01187.pdf.

Gorodnichenko, Y., Pham, T., Talavera, O., 2021. Social media, sentiment and public opinions: Evidence from #brexit and #uselection. European Economic Review 136, 103772. URL: https:

//www.sciencedirect.com/science/article/pii/S0014292121001252, doi:https://doi.org/10.1016/j.euroecorev.2021.103772.

Hartzmark, S.M., Hirshman, S.D., Imas, A., 2021. Ownership, Learning, and Beliefs. The Quarterly Journal of Economics 136, 1665–1717.

Hussam, R., Rabbani, A., Reggiani, G., Rigol, N., 2022. Rational habit formation: experimental evidence from handwashing in india. American Economic Journal: Applied Economics 14, 1–41.

Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The phq-9: validity of a brief depression severity measure. Journal of General Internal Medicine 16, 606–613.

Levy, R., 2021. Social media, news consumption, and polarization: Evidence from a field experiment. American Economic Review 111, 831–70. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20191777, doi:10.1257/aer.20191777.

Liang, Y., 2024. Learning from unknown information sources. Management Science 0.

Loomba, S., de Figueiredo, A., Piatek, S.J., de Graaf, K., Larson, H.J., 2021. Measuring the impact of covid-19 vaccine misinformation on vaccination intent in the uk and usa. Nature Human Behaviour 5, 337–348.

Moorman, C., Matulich, E., 1993. A model of consumers' preventive health behaviors: The role of health motivation and health ability. Journal of Consumer Research 20, 208–228.

Morikawa, M., 2020. People's perception on the impacts of covid-19 and policy responses: Observations from a survey for individuals. RIETI Policy Discussion Paper Series 20-P-020.

Moscadelli, A., Albora, G., Biamonte, M.A., Giorgetti, D., Innocenzio, M., Paoli, S., Lorini, C., Bonanni, P., Bonaccorsi, G., 2020. Fake news and covid-19 in Italy: Results of a quantitative observational study. International Journal of Environmental Research and Public Health 17. URL: https://www.mdpi.com/1660-4601/17/16/5850.

Muris, C., 2017. Estimation in the fixed-effects ordered logit model. The Review of Economics and Statistics 99, pp. 465–477.

Pinna, M., Picard, L., Goessmann, C., 2022. Cable news and covid-19 vaccine uptake. Scientific Reports 12.

Ren, W., Zhu, X., Hu, Y., 2022. Differential effects of traditional and social media use on covid-19 preventive behaviors: The mediating role of risk and efficacy perceptions. Journal of Health Psychology 27, 1861–1874. doi:10.1177/13591053211003125. pMID: 33909510.

Sakya, S.M., Scoy, L.J.V., Garman, J.C., Miller, E.L., Snyder, B., Wasserman, E., Chinchilli, V.M., Lennon, R.P., 2021. The impact of covid-19-related changes in media consumption on public knowledge: results of a cross-sectional survey of pennsylvania adults. Current Medical Research and Opinion 37, 911–915. doi:10.1080/03007995.2021.1901679.

Simonov, A., Sacher, S., Dube, J.P., Biswas, S., 2021. Frontiers: The persuasive effect of fox news: Noncompliance with social distancing during the covid-19 pandemic. Marketing Science 41, 230–242.

Spitzer, R.L., Kroenke, K., Williams, J.B., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the gad-7. Archives of Internal Medicine 166, 1092–1097.

Uchibori, M., Ghaznavi, C., Murakami, M., Eguchi, A., Kunishima, H., Kaneko, S., Maruyama-Sakurai, K., Miyata, H., Nomura, S., 2022. Preventive behaviors and information sources during covid-19 pandemic: A cross-sectional study in japan. International Journal of Environmental Research and Public Health 19.

WHO, 2021. Listings of who's response to covid-19. https://www.who.int/news/item/29-06-2020-covidtimeline. Accessed on December 24, 2022.

WHO, 2023. Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic. Accessed on May 17, 2023.

Zhang, C., Adriaanse, M.A., Potgieter, R., Tummers, L., de Wit, J., Broersen, J., de Bruin, M., Aarts, H., 2022. Habit formation of preventive behaviours during the covid-19 pandemic: a longitudinal study of physical distancing and hand washing. BMC Public Health 22, 1588.

Zimmermann, F., 2020. The dynamics of motivated beliefs. American Economic Review 110, 337–61. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20180728, doi:10.1257/aer.20180728.

Appendix A: Estimation

To estimate the parameters $\theta_{pred} \equiv (\boldsymbol{\beta}_{pred}, \boldsymbol{\gamma}_{pred})' \in \Theta_{pred}$, we employ the composite-likelihood estimator by Abrevaya and Muris (2020). Let $T \equiv \{(R1, R2), (R1, R3), \ldots, (R4, R5)\}$ denote the set of all distinct pairs of time periods, where the cardinality of T is 10. Te error term is assumed to follow a standard logistic distribution:

$$(e_{it_1}, e_{it_2})|(\alpha_{pred,i}, \boldsymbol{X}_{it_1}, \boldsymbol{X}_{it_2}, \boldsymbol{Z}_i) \sim \text{i.i.d. logistic.}$$

for all $(t_1, t_2) \in T$.

For $(t_1, t_2) \in T$, let $\pi_{t_1 t_2} \equiv (\pi_{t_1}, \pi_{t_2})$ be a pair of values in $\{1, 2, \ldots, J-1\}$. For now, let $\pi_{t_1 t_2}$ be any one of s $(J-1)^2$ possible pairs. We define an indicator function as

$$d_{pred,it\pi_t} \equiv \mathbb{1}\{y_{pred,it} > \pi_t\} = \mathbb{1}\{y^*_{pred,it} \geq c_{pred,\pi_t,t}\} = \mathbb{1}\{e_{it} \leq (\alpha_{pred,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{pred} - c_{pred,\pi_t,t})/\sigma_{pred}(\boldsymbol{Z}_i)\}$$

for $t = t_1, t_2$. Letting $\Lambda(v) = \exp(v)/[1 + \exp(v)]$, Theorem 1 of Abrevaya and Muris (2020) shows that

$$\begin{aligned} p_{pred,1,\pi_{t_1 t_2}}(\boldsymbol{W}_{it}, \theta_{pred}) &\equiv P(d_{pred,it_1\pi_{t_1}} = 1, d_{pred,it_2\pi_{t_2}} = 0 | d_{pred,it_1\pi_{t_1}} + d_{pred,it_2\pi_{t_2}} = 1, \alpha_{pred,i}, \boldsymbol{X}_{it_1}, \boldsymbol{X}_{it_2}, \boldsymbol{Z}_i) \\ &= 1 - \Lambda\left((\boldsymbol{X}_{it_2} - \boldsymbol{X}_{it_1})\frac{\boldsymbol{\beta}_{pred}}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{pred})} - (c_{pred,\pi_{t_2},t_2} - c_{pred,\pi_{t_1},t_1})\frac{1}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{pred})}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} p_{pred,2,\pi_{t_1 t_2}}(\boldsymbol{W}_{it}, \theta_{pred}) &\equiv P(d_{pred,it_1\pi_{t_1}} = 0, d_{pred,it_2\pi_{t_2}} = 1 | d_{pred,it_1\pi_{t_1}} + d_{pred,it_2\pi_{t_2}} = 1, \alpha_{pred,i}, \boldsymbol{X}_{it_1}, \boldsymbol{X}_{it_2}, \boldsymbol{Z}_i) \\ &= \Lambda\left((\boldsymbol{X}_{it_2} - \boldsymbol{X}_{it_1})\frac{\boldsymbol{\beta}_{pred}}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{pred})} - (c_{pred,\pi_{t_2},t_2} - c_{pred,\pi_{t_1},t_1})\frac{1}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{pred})}\right), \end{aligned}$$

where $\boldsymbol{W}_{it} = (\boldsymbol{X}'_{it_1}, \boldsymbol{X}'_{it_2}, \boldsymbol{Z}_i)'$. For some $(t_1, t_2) \in T$, the composite maximum likelihood estimator is given by:

$$\begin{aligned} \hat{\theta}_{pred,t_1 t_2} = \underset{\theta_{pred} \in \Theta_{pred}}{\text{argmax}} \sum_{\pi_{t_1 t_2}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{d_{pred,it_1\pi_{t_1}} + d_{pred,it_2\pi_{t_2}} = 1\} \Big\{ d_{pred,it_1\pi_{t_1}} \ln\Big[1 - p_{pred,1\pi_{t_1 t_2}}(\boldsymbol{W}_{it}, \theta_{pred})\Big] \\ + d_{pred,it_2\pi_{t_1}} \ln p_{pred,2\pi_{t_1 t_2}}(\boldsymbol{W}_{it}, \theta_{pred}) \Big\} \end{aligned}$$

Since the cardinality of T is 10, we obtain 10 estimates of the $(k+1)$-dimensional parameter vector, denoted by $\hat{\theta}_{pred,t_1t_2}$ for $(t_1, t_2) \in T$. Let $\hat{\theta}_{pred,t_1t_2,j}$ denote the jth element of $\hat{\theta}_{pred,t_1t_2}$ and define $\hat{\theta}_{pred,j} = (\hat{\theta}_{pred,R_1R_2,j}, \hat{\theta}_{pred,R_1R_3,j} \ldots, \hat{\theta}_{pred,R_4R_5,j})'$ as the 10-dimensional vector consisting of estimates for the jth element of θ_{pred}. We then combine the estimates using the minimum distance estimator:

$$\hat{\theta}_{pred} = \underset{\theta \equiv (\theta_{pred,1},\ldots,\theta_{pred,k+1})' \in \Theta_{pred}}{\operatorname{argmin}} \begin{pmatrix} \hat{\theta}_{pred,1} - \theta_{pred,1} \\ \hat{\theta}_{pred2} - \theta_{pred,2} \\ \vdots \\ \hat{\theta}_{pred,k+1} - \theta_{pred,k+1} \end{pmatrix}' W \begin{pmatrix} \hat{\theta}_{pred,1} - \theta_{pred,1} \\ \hat{\theta}_{pred,2} - \theta_{pred,2} \\ \vdots \\ \hat{\theta}_{pred,k+1} - \theta_{pred,k+1} \end{pmatrix}$$

where W is the block diagonal matrix whose block matrices are $I_1, \ldots, I_{k+1}$, with I_j being the 10-dimensional identity matrix for all $j = 1, \ldots, k+1$. The standard errors for each element of $\hat{\theta}$ are obtained via bootstrapping.[20]

[20]Standard errors are computed through the following steps.

1. Randomly resample n observations from $\{y_{pred,it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_i\}_{i=1}^{n}$ with replacement for each $t \in \{\mathrm{R1}, \mathrm{R2}, \mathrm{R3}, \mathrm{R4}, \mathrm{R5}\}$.
2. Estimate θ_{pred} using the resampled observations, denoted as $\hat{\theta}^b_{pred}$.
3. Repeat steps 1 and step 2 M times to obtain $\hat{\theta}^b_{pred}, b = 1, 2, \ldots, M$.
4. Compute the standard errors as the sample variance of $\{\hat{\theta}^b_{pred}\}_{b=1}^{M}$.

Appendix B: Data

B-1: Definition of variables

This appendix gives the definitions of variables used in this study.

Table 12: Definition of variables

Variable name	Definition
Traditional media	Dummy variable for using television (both public and commercial), newspaper, magazine, and radio
Nontraditional media	Dummy variable for using internet search engine, news app/site, governmental/corporate/specialized institutions website, Facebook, Twitter, Instagram, LINE, academic information, others, and none
High density prefecture	Dummy variable for living in prefecture where population density is greater than 5,000 persons per square kilometer
Marriage	Dummy variable for married respondents
Age	Age of respondents
Female	Dummy variable for female
Education	Years of education
Number of infections	Number of infections per 100,000
Ever infected	Dummy variable for having been infected with COVID-19
Vaccination rate	Vaccination rate of the prefecture where respondents live
FCV-19S	The Fear of COVID-19 Scale to measure fear related to COVID-19
PHQ-9	Patient Health Questionnaire-9 scale: The measure for depression (Kroenke et al., 2001)
GAD-7	Generalized Anxiety Disorder-7 scale: The measure for anxiety (Spitzer et al., 2006)
Trust	Dummy variable for answering yes to the question "Do you think people are generally trustworthy?"
For robustness check	
Traditional media	Dummy variable for using television (both public and commercial), newspaper, magazine, radio, academic information, and governmental/corporate/specialized institutions websites
Nontraditional media	internet search engine, news app/site, Facebook, Twitter, Instagram, LINE, others, and none

B-2: Descriptive statistics for prevention behaviors

Tables 13 and 14 present the mean and standard deviation of the infection prevention behavior indicators.

Table 13: Mean and standard deviation of preventive behaviors

		(I)	(II)	(III)	(IV)	(V)	(VI)
Nontraditional media							
	Round 1	0.809	0.830	0.771	0.944	0.944	0.844
		(0.393)	(0.375)	(0.420)	(0.231)	(0.229)	(0.362)
	Round 2	0.832	0.868	0.840	0.957	0.949	0.865
		(0.375)	(0.339)	(0.366)	(0.202)	(0.221)	(0.342)
	Round 3	0.859	0.877	0.839	0.955	0.942	0.877
		(0.349)	(0.329)	(0.367)	(0.207)	(0.235)	(0.328)
	Round 4	0.847	0.863	0.830	0.950	0.939	0.870
		(0.360)	(0.344)	(0.376)	(0.218)	(0.239)	(0.336)
	Round 5	0.851	0.864	0.839	0.955	0.940	0.876
		(0.356)	(0.343)	(0.367)	(0.208)	(0.237)	(0.330)
Traditional media							
	Round 1	0.886	0.896	0.831	0.980	0.973	0.907
		(0.318)	(0.305)	(0.374)	(0.139)	(0.161)	(0.290)
	Round 2	0.896	0.925	0.888	0.987	0.971	0.921
		(0.306)	(0.264)	(0.315)	(0.113)	(0.169)	(0.270)
	Round 3	0.917	0.921	0.893	0.988	0.973	0.928
		(0.277)	(0.270)	(0.310)	(0.109)	(0.162)	(0.259)
	Round 4	0.911	0.920	0.886	0.986	0.970	0.923
		(0.285)	(0.272)	(0.317)	(0.118)	(0.171)	(0.267)
	Round 5	0.915	0.917	0.882	0.987	0.970	0.925
		(0.279)	(0.276)	(0.322)	(0.113)	(0.170)	(0.263)

Note: Standard deviations are displayed in the parentheses.

Table 14: Mean and standard deviation of preventive behaviors

		(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Nontraditional media							
	Round 1	0.225	0.669	0.298	0.783	0.516	0.602
		(0.418)	(0.471)	(0.458)	(0.412)	(0.500)	(0.490)
	Round 2	0.249	0.705	0.330	0.827	0.587	0.735
		(0.432)	(0.456)	(0.470)	(0.379)	(0.492)	(0.441)
	Round 3	0.276	0.688	0.359	0.832	0.554	0.689
		(0.447)	(0.463)	(0.480)	(0.374)	(0.497)	(0.463)
	Round 4	0.292	0.680	0.356	0.821	0.511	0.653
		(0.455)	(0.466)	(0.479)	(0.384)	(0.500)	(0.476)
	Round 5	0.284	0.688	0.363	0.829	0.507	0.649
		(0.451)	(0.463)	(0.481)	(0.377)	(0.500)	(0.477)
Traditional media							
	Round 1	0.211	0.697	0.271	0.849	0.490	0.613
		(0.408)	(0.460)	(0.444)	(0.358)	(0.500)	(0.487)
	Round 2	0.231	0.722	0.299	0.884	0.545	0.775
		(0.422)	(0.448)	(0.458)	(0.321)	(0.498)	(0.418)
	Round 3	0.256	0.714	0.324	0.891	0.512	0.717
		(0.436)	(0.452)	(0.468)	(0.312)	(0.500)	(0.451)
	Round 4	0.276	0.694	0.318	0.879	0.469	0.689
		(0.447)	(0.461)	(0.466)	(0.326)	(0.499)	(0.463)
	Round 5	0.252	0.706	0.312	0.871	0.453	0.658
		(0.434)	(0.455)	(0.463)	(0.335)	(0.498)	(0.475)

Note: Standard deviations are displayed in the parentheses.

Appendix C: Robustness check

C-1: Robustness check for the definition of media

Classifying information sources on COVID-19 into traditional and nontraditional media may be subject to debate. To assess the robustness of our main results, we re-estimate the interval censored regression model for predictions using an alternative definition of media. In this re-estimation, we classify governmental, corporate, specialized institutions' websites, and academic information under the traditional media category (Table 12), as information from academic and governmental sources tends to be more coherent and less diffuse.

The estimation results, reported in Tables 15 and 16, confirm that the findings presented in Tables 3 and 4 are robust to this alternative definition of media, both in terms of statistical significance and the magnitude of significant estimates.

C-2: Robustness check for the definition of $y^*_{pred,it}$

The latent prediction $y^*_{pred,it}$ for respondent i is defined as the number of months remaining until the predicted end of the pandemic, as assessed at time t. For instance, $y^*_{pred,i1} = 3$ indicates that in round 1, respondent i predicts the pandemic will end within three months, whereas $y^*_{pred,i2} = 3$ signifies that at round 2, the same respondent predicts the pandemic will end within three months. This definition of $y^*_{pred,it}$ may be subject to debate, as the same numerical value at different time points (e.g., $y^*_{pred,i1} = y^*_{pred,i2} = 3$) does not correspond to the same calendar month.

To assess the robustness of our main results to this definition of $y^*_{pred,it}$, we re-estimate the interval censored regression model for predictions using an alternative definition of $y^*_{pred,it}$ that aligns with the calendar date.

Under this alternative specification, $y^*_{pred,it}$ is redefined as the number of months counted from the start of the first round of the survey. For example, $y^*_{pred,i1} = 3$ means a prediction that the pandemic will end within three months, whereas $y^*_{pred,i2} = 3$ means a prediction that the pandemic has already ended, given that round 2 was conducted three months after the first round.

The estimation results, reported in Tables 17 and 18, confirm that the findings are robust to this alternative definition of $y^*_{pred,it}$, both in terms of statistical significance and the magnitude of significant estimates.

Table 15: Interval censored model of predictions, robustness for the definition of traditional media: Estimates of $\boldsymbol{\beta}_{pred}$

	(1)		(2)	
Estimates of $\boldsymbol{\beta}_{pred}$				
Vaccination rate	2.480	(5.816)	4.623	(5.104)
Number of infections	0.012	(0.011)	0.052	(0.026)
Ever infected	-0.795	(1.956)	-0.064	(1.990)
Round 2 dummy	6.210	(4.374)	7.238	(4.484)
Round 3 dummy	6.552	(2.865)	5.940	(3.031)
Round 4 dummy	5.257	(3.765)	0.895	(3.769)
Round 5 dummy	-0.108	(4.896)	-4.312	(4.653)
Nontraditional media $\times$				
Vaccination rate	4.147	(11.850)	1.707	(10.524)
Number of infections	0.038	(0.025)	-0.058	(0.059)
Ever infected	-0.402	(3.240)	0.819	(2.875)
Round 2 dummy	-2.092	(8.876)	1.722	(8.856)
Round 3 dummy	-3.820	(6.752)	2.426	(6.702)
Round 4 dummy	-8.685	(7.356)	-5.541	(7.461)
Round 5 dummy	-8.007	(10.210)	-2.648	(9.339)

Note: Standard errors are displayed in the parentheses.

Table 16: Interval censored model of predictions, robustness for the definition of traditional media: Estimates of γ_{pred}

	(1)		(2)	
Estimates of γ_{pred}				
Constant	1.753	(0.207)	2.626	(0.186)
Age			-0.004	(0.001)
Education			-0.045	(0.009)
Female			-0.106	(0.027)
Married			-0.024	(0.024)
High density prefecture			-0.010	(0.023)
Average of GAD–7			-0.016	(0.064)
Average of PHQ–9			0.035	(0.049)
Average of Trust			-0.040	(0.037)
Variance of GAD–7			0.179	(0.065)
Variance of PHQ–9			-0.010	(0.043)
Variance of Trust			0.273	(0.111)
Nontraditional media	0.207	(0.019)	0.166	(0.028)
Estimates of $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})$				
Nontraditional media	7.101		6.953	
Traditional media	5.773		5.887	
Difference in $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})$ between users of nontraditional and traditional media	1.328		1.066	
95 percent prediction interval of $\exp(\bar{\boldsymbol{Z}}_i\gamma_{pred})e_{it}$				
Nontraditional media	[-26.026, 26.017]		[-25.484, 25.476]	
Traditional media	[-21.159, 21.152]		[-21.576, 21.569]	

Note: Standard errors are displayed in the parentheses. $\bar{\boldsymbol{Z}}_i$ denotes the average of $\boldsymbol{Z}_i$.

Table 17: Interval censored model of predictions, robustness for the definition of $y^*_{pred,it}$: Estimates of $\boldsymbol{\beta}_{pred}$.

	(1)		(2)	
Estimates of $\boldsymbol{\beta}_{pred}$				
Vaccination rate	3.663	(4.849)	2.003	(4.891)
Number of infections	0.004	(0.010)	0.020	(0.017)
Ever infected	-1.921	(1.607)	-0.695	(1.668)
Round 2 dummy	10.257	(4.315)	9.516	(4.346)
Round 3 dummy	8.488	(2.783)	8.046	(2.850)
Round 4 dummy	2.866	(3.555)	2.868	(3.636)
Round 5 dummy	0.227	(4.469)	1.979	(4.627)
Nontraditional media ×				
Vaccination rate	9.121	(8.701)	7.801	(8.630)
Number of infections	0.039	(0.021)	0.055	(0.035)
Ever infected	-0.289	(2.257)	-0.404	(2.205)
Round 2 dummy	-12.240	(6.893)	-3.909	(7.341)
Round 3 dummy	-6.420	(5.744)	-1.238	(5.750)
Round 4 dummy	-9.396	(6.283)	-9.067	(6.481)
Round 5 dummy	-9.757	(8.072)	-8.401	(8.009)

Note: Standard errors are displayed in the parentheses. In this estimation, we remove respondents by whom the time taken to answer the question is shorter than 1 minute or longer than 120 minutes; $y^*_{pred,it}$ is the number of weeks elapsed since the start of the survey.

Table 18: Interval censored model of predictions: time dummies, robustness for the definition of $y^*_{pred,it}$: Estimates of $\boldsymbol{\gamma}_{pred}$

	(1)		(2)	
Estimates of $\boldsymbol{\gamma}_{pred}$				
Constant	1.633	(0.008)	2.520	(0.187)
Age			-0.004	(0.001)
Education			-0.046	(0.009)
Female			-0.107	(0.024)
Married			0.016	(0.021)
High density prefecture			0.008	(0.020)
Average of GAD–7			0.040	(0.046)
Average of PHQ–9			-0.033	(0.036)
Average of Trust			-0.054	(0.033)
Variance of GAD–7			0.139	(0.056)
Variance of PHQ–9			0.020	(0.026)
Variance of Trust			0.264	(0.110)
Nontraditional media	0.173	(0.015)	0.145	(0.023)
Estimates of $\exp(\bar{\boldsymbol{Z}}_i\boldsymbol{\gamma}_{pred})$				
Nontraditional media	6.086		6.067	
Traditional media	5.119		5.248	
Difference in $\exp(\bar{\boldsymbol{Z}}_i\boldsymbol{\gamma}_{pred})$ between users of nontraditional and traditional media	0.967		0.819	
95 percent prediction interval of $\exp(\bar{\boldsymbol{Z}}_i\boldsymbol{\gamma}_{pred})e_{it}$				
Nontraditional media	[-22.304, 22.297]		[-25.484, 25.476]	
Traditional media	[-18.760, 18.754]		[-21.576, 21.569]	

Note: Standard errors are displayed in the parentheses. In this estimation, we remove respondents by whom the time taken to answer the question is shorter than 1 minute or longer than 120 minutes; $y^*_{pred,it}$ is the number of weeks elapsed since the start of the survey. $\bar{\boldsymbol{Z}}_i$ denotes the average of $\boldsymbol{Z}_i$.

Appendix D: Representativeness of data

As discussed in Section 2, we assess the representativeness of our survey data by comparing it with data from the "FY2020 Survey on Usage Time of Information and Communications Media and Information Behavior," conducted by the Ministry of Internal Affairs and Communications (MIC) in January 2021. This MIC survey employed a placement method and included 1,500 respondents, comprising men and women aged 13 to 69, who were asked to identify their most frequently used information source for current affairs. Table 19, which is constructed based on a published aggregate table, presents the average values of demographic variables in the MIC survey.

Table 19: Demographic differences by information sources: survey by MIC

	Traditional	Nontraditional
Female	0.52	0.45
Age		
13-19	0.07	0.13
20-29	0.10	0.21
30-39	0.13	0.23
40-49	0.20	0.24
50-59	0.24	0.12
60-69	0.25	0.08
Years of education	13.47	13.69
Household income (10,000 yen)	535.43	544.89
Proportion	0.62	0.38

Appendix E: Models for infection prevention behavior

In section 5.2, we estimate the effect of information sources on infection prevention behaviors. This appendix explains the two models used in the estimation.

The first model is a linear probability model with fixed effects, where the outcome variables are binary indicators for preventive behaviors. The explanatory variables include control variables and their interactions with the nontraditional media dummy.

The second model is a regression model for the variance of preventive behaviors. In this model, the outcome variable is the variance of preventive behaviors, estimated using the residuals from the fixed effects estimation of the first model. These residuals are then regressed on the nontraditional media dummy and other controls.

Let $y_{beh,it}$ denote a binary variable indicating whether an individual engages in the infection prevention behavior against COVID-19. The first model, a linear probability model with fixed effects, is specified as follows:

$$y_{beh,it} = \alpha_{beh,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{beh} + \mu_{it},$$

where $\alpha_{beh,i}$ is the individual fixed effect, $\boldsymbol{X}_{it}$ is the vector of control variables, $\boldsymbol{\beta}_{beh}$ is the vector of regression coefficients, and μ_{it} is the error term. To account for the fixed effect $\alpha_{beh,i}$, we employ a fixed effects estimator to estimate parameters $\boldsymbol{\beta}_{beh}$.

Let $\hat{\boldsymbol{\beta}}_{beh}$ denote the fixed effects estimator of $\boldsymbol{\beta}_{beh}$. Then, the residuals are obtained as

$$\widehat{\dot{\mu}}_{it} = \dot{y}_{beh,it} - \dot{\boldsymbol{X}}_{it}\hat{\boldsymbol{\beta}}_{beh},$$

where $\dot{\mu}_{it}$, $\dot{y}_{beh,it}$, and $\dot{\boldsymbol{X}}_{it}$ represent within-transformations of μ_{it}, $y_{beh,it}$, and $\boldsymbol{X}_{it}$, respectively. For example, $\dot{\mu}_{it} = \mu_{it} - \bar{\mu}_i$, where $\bar{\mu}_i$ is the individual-specific mean of μ_{it} over time.

The outcome variable for the second model is an unbiased estimator of the variance of μ_{it} for each i, defined as $\hat{\sigma}^2_{beh,i} \equiv \frac{1}{T-1}\sum_{t=1}^{T}\widehat{\dot{\mu}}^2_{it}$. Then, the variance of preventive behaviors is modeled as

$$\hat{\sigma}^2_{beh,i} = \boldsymbol{Z}_i\gamma_{beh} + \epsilon_i,$$

where $\boldsymbol{Z}_i$ includes the nontraditional media dummy and other control variables, γ_{beh} is the vector of regression coefficients, and ϵ_i is the error term. The parameter $\boldsymbol{\gamma}_{beh}$ is then estimated using the ordinary least squares (OLS) estimator.

Appendix F: Model for perception

In section 5.2, we estimate the effect of information sources on perception of the risk of COVID-19. This appendix explains econometric model used for the analysis of perception and presents robustness checks for the selection of thresholds values.

F-1: Econometric model

Perception is measured using the Fear of COVID-19 Scale (FCV-19S), which is a categorical variable based on ordered responses. Thus, it is natural to employ estimators for ordered response models. To account for heteroskedasticity, we modify the fixed effects estimator for ordered logit models developed by Muris (2017). Let $y_{per,it}$ denote the FCV-19S score for respondent i at time t, and let $y^*_{per,it}$ represent its latent response, which is assumed to follow:

$$y^*_{per,it} = \alpha_{per,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{per} + \exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})\varepsilon_{it}, \qquad t = R1, ..., R5$$

$$(\varepsilon_{iR1}, ..., \varepsilon_{iR5})|(\alpha_{per,i}, \boldsymbol{X}_{iR1}, ..., \boldsymbol{X}_{iR5}, \boldsymbol{Z}_i) \sim \text{i.i.d. logistic.}$$

where $\alpha_{per,i}$ represents individual fixed effects, $\boldsymbol{X}_{it}$ and $\boldsymbol{Z}_i$ are vectors of individual attributes, $\boldsymbol{\beta}_{per}$ and $\boldsymbol{\gamma}_{per}$ are corresponding vectors of parameters. The observed categorical response $y_{per,it}$ is linked to its latent counterpart $y^*_{per,it}$ through the following threshold model:

$$y_{per,it} = \begin{cases} 1 \quad \text{if} \quad y^*_{per,it} < c_{per,1}, \\ 2 \quad \text{if} \quad c_{per,1} \leq y^*_{per,it} < c_{per,2}, \\ \qquad \vdots \\ K \quad \text{if} \quad c_{per,K-1} \leq y^*_{per,it}, \end{cases}$$

where $c_{per,1}, ..., c_{per,K-1}$ are threshold values. In our study, we set $K = 4$.

To estimate the parameter vector $\theta_{per} \equiv (\boldsymbol{\beta}_{per}, \boldsymbol{\gamma}_{per}, \boldsymbol{c}_{per})' \in \Theta_{per}$, we employ the composite-likelihood estimator proposed by Muris (2017). For t $(= R1, ..., R5)$, let $\{\rho(t)\}_{t=R1}^{R5}$ denote a time series of cutoff categories, where each $\rho(t)$ takes values in $\{1, 2, \ldots, K-1\}$. We define the binary response variable:

$$d_{per,i\rho} = (d_{per,it\rho} = \mathbb{1}\{y_{per,it} \leq \rho(t)\}, t = R1, ..., R5)$$

where

$$d_{per,it\rho} \equiv \mathbb{1}\{y^*_{per,it} < c_{per,\rho(t)}\} = \mathbb{1}\{\alpha_{per,i} + \boldsymbol{X}_{it}\boldsymbol{\beta}_{per} + \exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})\varepsilon_{it} < c_{per,\rho(t)}\}.$$

Our specification differs from Muris (2017) in that we allow for heteroskedasiticity via $\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})$, whereas the original paper assumes homoskedasticity. Let $\bar{d}_{per,i\rho} \equiv \sum_{t=R1}^{R5} d_{per,it\rho}$ be the number of observations at or below the corresponding cutoff categories. Define the probability distribution of $d_{per,i\rho}$ conditional on $\bar{d}_{per,i\rho}$,as $p_{per,i\rho}(d|\boldsymbol{\beta}_{per}, \boldsymbol{\gamma}_{per}, \boldsymbol{c}_{per}) \equiv P\left(d_{per,i\rho} = d|\bar{d}_{per,i\rho} = \bar{d}, \boldsymbol{X}_i, \alpha_{per,i}\right)$. Let $F_{\bar{d}_{per}}$ denote the set of all binary series that set exactly $\bar{d}_{per}$ elements to 1: $F_{\bar{d}_{per}} = \left\{f \in \{0,1\}^5 \text{ such that } \bar{f} = \bar{d}_{per}\right\}$. Consequently, equation (17) of Muris (2017) is modified as follows:

$$p_{i\rho}(d_{per}|\boldsymbol{\beta}_{per}, \boldsymbol{\gamma}_{per}, \boldsymbol{c}_{per}) = \left\{\sum_{f \in F_{\bar{d}_{per}}} \exp\left\{\sum_{t=R1}^{R5} \frac{(f_t - d_{per,t})(c_{per,\rho(t)} - c_{per,\rho(1)})}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})} - \sum_{t=R1}^{R5} \frac{(f_t - d_{per,t})\boldsymbol{X}_{it}\beta_{per}}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})}\right\}\right\}^{-1}.$$

When ρ is appropriately chosen, we can estimate θ_{per} as following equation.

$$\hat{\theta}_{per,\rho} = \text{argmax} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[d_{per,i} = d_{per}] \ln p_{per,i\rho}(d_{per}|\theta_{per,\rho}).$$

However, $\hat{\theta}_{per,\rho}$ is often sensitive to choice of ρ. Therefore, we employ the composite maximum likelihood estimator, as given in equation (34) of Muris (2017). The estimator is expressed as follows:

$$\hat{\theta}_{per} = \text{argmax} \frac{1}{n} \sum_{i=1}^{n} q(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \theta_{per})$$

where

$$q(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \theta_{per}) = -\sum_{\rho} \mathbb{1}_{[d_{per,i} = d_{per}]} \ln \sum_{f \in F_{\bar{d}_{per}}} \exp\left\{\sum_{t=R1}^{R5} \frac{(f_t - d_{per,t})(c_{per,\rho(t)} - c_{per,\rho(1)} - \boldsymbol{X}_{it}\beta_{per})}{\exp(\boldsymbol{Z}_i\boldsymbol{\gamma}_{per})}\right\}$$

Table 20: Ordered logit model for perception, robustness for the thresholds selection: Estimates of $\boldsymbol{\beta}_{per}$ and thresholds

	Estimate	Std. Error
Estimates of $\boldsymbol{\beta}_{per}$		
Vaccination rate	-0.029	(0.231)
Number of infections	0.002	(0.001)
Ever infected	-0.196	(0.111)
Round 2 dummy	0.236	(0.022)
Round 3 dummy	0.141	(0.019)
Round 4 dummy	-0.002	(0.103)
Round 5 dummy	-0.072	(0.198)
Nontraditional media ×		
Vaccination rate	-0.006	(0.427)
Number of infections	-0.001	(0.001)
Ever infected	0.117	(0.175)
Round 2 dummy	-0.025	(0.040)
Round 3 dummy	-0.034	(0.035)
Round 4 dummy	0.064	(0.191)
Round 5 dummy	0.020	(0.365)
Thresholds		
between 1 and 2	-0.989	(0.631)
between 2 and 3	0.053	(0.632)
between 3 and 4	1.179	(0.634)

Note: The explained variable is a ordered variable based on the value of the FCV-19S scale (takes the largest value 4 if a respondent is afraid of COVID-19 and 1 if the respondent is not afraid of COVID-19).

The standard error of $\hat{\theta}_{per}$ is estimated using the asymptotic variance of the M-estimator. Let $s(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \theta_{per})$ and $H(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \theta_{per})$ denote the score and hessian of the objective function $q(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \theta_{per})$, respectively. Define A as a sample analogue of $E[s(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \hat{\theta}_{per})s'(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \hat{\theta}_{per})]$, and define B as a sample analogue of $E[H(d_{per,i}, \boldsymbol{X}_i, \boldsymbol{Z}_i, \hat{\theta}_{per})]$. We then estimate the standard error of $\hat{\theta}_{per}$ using the trace of the asymptotic variance matrix $n^{-1}A^{-1}BA^{-1}$.

F-2: Robustness of threshold selection

To examine the robustness of the threshold selection, Tables 20 and 21 present the results obtained using 10 randomly selected permutations of the threshold sequence 1–1–2–2–3. The results remain consistent with those reported in Tables 10 and 11, indicating that the interpretation does not change.

Table 21: Ordered logit model for perception, robustness for the thresholds selection: Estimates of $\boldsymbol{\gamma}_{per}$

	Estimate	Std. Error
Estimates of $\boldsymbol{\gamma}_{per}$		
Nontraditional media	0.034	(0.020)
Marriage	0.088	(0.021)
Age	-0.010	(0.001)
Female	-0.154	(0.019)
Education	-0.028	(0.003)
High density prefecture	0.007	(0.020)
Average of PHQ-9	-0.025	(0.025)
Average of GAD-7	0.017	(0.033)
Average of Trust	-0.011	(0.027)
Variance of PHQ-9	0.046	(0.013)
Variance of GAD-7	0.149	(0.038)
Variance of Trust	0.226	(0.081)
Estimates of $\exp(\bar{\boldsymbol{Z}}_i \boldsymbol{\gamma}_{per})$		
Nontraditional media	0.404	
Traditional media	0.391	
Difference in $\exp(\bar{\boldsymbol{Z}}_i \boldsymbol{\gamma}_{per})$ between users of nontraditional and traditional media	0.013	
95% prediction interval of $\exp(\bar{\boldsymbol{Z}}_i \boldsymbol{\gamma}_{per})\varepsilon_{it}$		
Nontraditional media	[-1.482, 1.482]	
Traditional media	[-1.432, 1.432]	

Note: The explained variable is a ordered variable based on the value of the FCV-19S scale (takes the largest value 4 if a respondent is afraid of COVID-19 and 1 if the respondent is not afraid of COVID-19). $\bar{\boldsymbol{Z}}_i$ denotes the average of $\boldsymbol{Z}_i$.