Social Science & Medicine 332 (2023) 116093

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Evolution of preferences for COVID-19 vaccine throughout the pandemic – The choice experiment approach

Ricardo Daziano [a], Wiktor Budziński [b,*]

[a] School of Civil and Environmental Engineering, Cornell University, USA
[b] Faculty of Economic Sciences, University of Warsaw, Poland

A R T I C L E I N F O

Handling Editor: Ricardo A. Daziano

JEL classification:
C35
D12
D91
I12

Keywords:
COVID-19 vaccine
Choice experiment
Pandemic conditions
Hybrid choice model
Perceived vulnerability to disease

A B S T R A C T

In this study, we employ a choice experiment to study individual preferences for COVID-19 vaccines in the US. A unique characteristic of the microdata (N = 5671) is that the survey was conducted in five distinct waves from October 2020 to October 2021. Because of this dynamic feature, it is possible to control for evolving pandemic conditions such as the number of COVID-19 active cases, vaccination uptake, and the frequency of Google searches related to the vaccines. Furthermore, we employ a hybrid choice model to incorporate respondents' attitudes related to their perceived vulnerability to diseases, as well as their perceived health status. The hybrid choice model was extended to incorporate latent classes as well as random effects. We find that the rate of vaccinated individuals in the population actually increases the probability of vaccine hesitancy, and therefore may discourage people to get vaccinated. This may be evidence of free-riding behavior. On the other hand, the number of COVID-19 cases has a positive effect on the probability of getting vaccinated, suggesting that individuals react to the pandemic conditions by taking some protective measures. Google trend data do not seem to have a straightforward effect on the vaccination demand, but it increases consumers' willingness to pay for several vaccine characteristics. With respect to the analyzed attitudes, we find that perceived uninfectability is a significant driver of vaccine hesitancy, probably related to the frequent "natural immunity" argument. In turn, germ aversion has a positive effect on the probability of getting vaccinated as well as on the marginal willingness to pay. Finally, health status has a limited effect on whether the individual will decide to vaccinate or not.

1. Introduction

COVID-19 constitutes one of the biggest challenges that the world had to face in modern times. In the US alone, by May of 2022, deaths attributed to the disease surpassed 1 million. From the start, preventive measures such as social distancing, wearing masks, and frequent handwashing were advised by health professionals and governments to limit the spread of the virus. Nonetheless, these measures were perceived as temporary, with the hope of the pandemic ending once a vaccine becomes available. An unprecedented global effort led to the quick development and deployment of vaccines, with the first vaccines being distributed in the US in December 2020 - just eleven months after SARS-CoV-2 – the virus that causes COVID-19 – was officially identified (Forni and Mantovani, 2021). At the same time, the topic of vaccination has become very polarizing for the public, with a steady, or even increasing share of vaccine-hesitant and anti-vaccine individuals (Johnson et al., 2020; Pullan and Dey, 2021). It has been noted that without reaching the herd immunity threshold, the effort put into the rapid development of vaccines may not be sufficient to stop the pandemic (Chevallier et al., 2021). It is therefore important to understand individuals' choices regarding vaccination from the social and behavioral perspective (Looi, 2022), especially to identify lessons for future waves with potential new variants or for future pandemics.

In this study, we employ the choice experiment (CE) method to study individuals' preferences for hypothetical COVID-19 vaccine in the US, throughout the pandemic. Using CEs allows us to analyze trade-offs that respondents make between different attributes of the vaccine, as well as assess the effect of other variables, such as attitudes, pandemic conditions, and health status on individuals' choices. From the start of the pandemic, researchers have been using CEs to identify attributes of the vaccine that may be important to the public and investigate how these attributes affect vaccine uptake (e.g., Borriello et al., 2021). Such studies have been conducted all around the world, including China, France, India, Malaysia, the Netherlands, the US, and the UK (just to give a few

* Corresponding author.
E-mail addresses: daziano@cornell.edu (R. Daziano), wbudzinski@wne.uw.edu.pl (W. Budziński).

https://doi.org/10.1016/j.socscimed.2023.116093
Received 30 December 2022; Received in revised form 7 June 2023; Accepted 11 July 2023
Available online 21 July 2023
0277-9536/

examples, see Dong et al., 2020; McPhedran and Toombs, 2021; Motta, 2021; Schwarzinger et al., 2021; Bansal et al., 2022; Mouter et al., 2022; Teh et al., 2022). For the review of recent studies consider Amani et al. (2022). The most common attributes considered are efficiency of the vaccine, risk of side-effects, and out-of-pocket cost. Other attributes considered include the number of doses (e.g., Chu and Liu, 2021), country of origin (e.g., Kreps and Kriner, 2021), whether the vaccine is recommended/approved by certain institutions (e.g., McPhedran and Toombs, 2021), place in which the vaccination will be administered (e.g., Schwarzinger et al., 2021), type of the vaccine (Borriello et al., 2021; Motta, 2021; Teh et al., 2022), duration of protection (e.g., Bansal et al., 2022), and share of friends/family that got vaccinated (Leng et al., 2021; Bansal et al., 2022). Nonetheless, most of these studies focus on identifying the effect of attributes of the CE and only consider basic socio-demographic variables which can be related to vaccine hesitancy (such as age, gender, ethnicity) as potential drivers of observed preference heterogeneity. Our study contrasts with this parallel research efforts by considering attributes that coincide with those of contemporary CEs, but also controlling for a wide variety of additional covariates that provide us with greater insight regarding factors that may affect individuals' vaccination choices. Specifically, we consider three types of factors: (i) pandemic conditions, (ii) perceived vulnerability to diseases, and (iii) health status.

Another unique feature of our study is that data collection did not focus only on the early stages of the pandemic. Some of the previous research shows that the preferences for vaccination may change over time. For example, Daly and Robinson (2021) identify decreasing vaccination intention in the US, whereas Sanders et al. (2021) find an opposite trend in the Netherlands. At the same time, Chambon et al. (2022) and Raciborski et al. (2021) conclude that such preferences may be difficult to change, and are rather stable over time (in the Netherlands and Poland, respectively). To account for this temporal dimension, our study was effectively conducted in five waves, with the first wave starting in October 2020, and the last wave ending in October 2021. Because of the dynamic nature of data collection, we can control for (i) pandemic conditions, by looking at the number of COVID-19 active cases during the week a respondent took the survey, the share of the vaccinated population, and the number of Google searches related to the vaccine in that given week. The longer timespan of the survey provides us with greater insight regarding changes in preferences depending on the evolving pandemic conditions, which is not possible to identify in other research efforts in which the CE survey was usually conducted within a single month.

The private benefits of taking a vaccine for COVID-19 depend on how vulnerable to a disease one considers oneself to be. As such, vulnerability is an important factor to control when analyzing individuals' preferences. We account for this factor by incorporating (ii) perceived vulnerability into our model by employing the Perceived Vulnerability to Disease Questionnaire (PVDQ, Díaz et al., 2016). To the best of our knowledge, this is the first empirical application that uses this scale in a stated preference setting and links it with individuals' preferences for vaccines. As the scale is not necessarily COVID-oriented, the results can be considered more general. Extant research regarding COVID-19 vaccination is not really concerned with the effect of attitudes. The only exceptions are Leng et al. (2021), who control for the perceived risk of infection with COVID-19, and Schwarzinger et al. (2021), who control for the perceived severity of COVID-19 if infected. To jointly model responses to the PVDQ scale and to the CE, we specified and derived parameter estimates of a state-of-the-art hybrid choice model (HCM, Ben-Akiva et al., 2002). This integrated choice and latent variable framework allows us to incorporate the PVDQ scale into the choice model in the form of latent factors while controlling at the same time for measurement error that is likely to arise (Budziński and Czajkowski, 2022). To assure high quality of our estimates, we combine a latent class specification with random effects in the choice component of the HCM.

Lastly, health status is an important risk factor that can affect the severity of COVID-19 once infected. We control for (iii) health status by considering several perceptions of the respondent's own health and then treating health status as an additional latent factor. In previous research, only Schwarzinger et al. (2021) and Teh et al. (2022) control for any health conditions, but their findings are mixed.

Overall, our study contributes to the literature on individual preferences for vaccination in many directions. The long period over which the study was conducted allows us to get a better insight into individuals' behavior. Specifically, we focus on different aspects of the pandemic such as vaccination rate and the number of COVID-19 active cases, to assess how they affect individuals' vaccination uptake, and their preferences for a variety of vaccine attributes. Furthermore, we incorporate individuals' attitudes to link vaccine preferences with health-related scales that have never been employed in stated preference research. As estimation strategy, we achieve this integration by coding a novel hybrid choice model specification with a latent class logit kernel for the identification of discrete preference clusters and a continuous random effect to account for unobserved heterogeneity from repeated responses by the same individual. Our results provide relevant insights for research and development of future vaccines and for policymakers who need to identify the correct incentives for individuals to vaccinate. Even though the pandemic may be over soon (e.g. due to the development of a pan-vaccine (Looi, 2022) or herd immunity built by the Omicron variant (Das et al., 2022)), seasonal COVID-19 vaccination may become necessary in a similar way as flu-vaccination campaigns. Furthermore, the results are also relevant for future pandemics that may arise (Halabowski and Rzymski, 2021).

The rest of the paper is structured as follows. In Section 2 we briefly discussed the methods used in this study. Specifically, we describe the survey data and the econometric model. In Section 3, we present the results of the analysis of the data. In Section 4 we discuss the limitations of the current study. The last section provides a discussion of the results and their implication for public policy.

2. Methods

In this section, we describe in detail the methods used in the current study. In Section 2.1 we provide general information regarding the survey. We then describe the CE data on which we base our analysis. In Section 2.3 we briefly describe the covariates related to evolving pandemic conditions. Finally, Section 2.4 provides a description of the adopted econometric model.

2.1. Survey data

The data that we use come from an online survey conducted in the five distinct waves across the United States. The objective of the survey was to investigate public opinion regarding several COVID-related topics. Specifically, how COVID-19 affected individuals' daily lives, their preferences, and attitudes toward hypothetical vaccines, as well as preferences regarding COVID-19 tests. The survey consisted of seven blocks, with a median time of 20.7 min to complete it. The first block involved general questions related to COVID-19 and vaccination. It was followed by the CE used in this survey, which is described in more detail in Section 2.2. Then respondents were asked about their health. This block included the PVDQ scale that we utilize to identify latent factors (consult Table 3 below). The fourth block consisted of a second CE regarding COVID-19 tests, which we do not analyze in the current study. Then respondents were asked about the effect of COVID-19 on their daily commute. The sixth block consisted of questions regarding the impact of COVID-19 on other aspects of daily life. Finally, the last block included standard socio-demographic questions.

The sample was based on a Qualtrics panel aimed at being representative of adults in the USA in age, gender, and income. In Table 1 we present a summary of control variables that we use in the econometric model. When compared to the general population of the US, our sample

Table 1
Summary of socio-demographic and worldview-related variables.

	Median (or share)	
	Sample	Population
Age	42.000***	49[2]
Age over 65 (share)	0.183***	0.212[2]
Males (share)	0.475***	0.492
No. of children in the household	1.722	
No. of elderly in the household	1.526	
Has at least a bachelor degree (share)	0.596***	0.321
Race: black (share)	0.158***	0.134
Race: other non-white (share)	0.095***	0.074
Household income (in thousands of $ per year)	67.500***	62.843
Employed (share)	0.573	
Has driving licence (share)	0.891	
Worldview-related variables		
Republican (share)	0.280	
Religion is fairly important (or more; share)	0.645	
Conservative (somewhat or more; share)	0.349	

Note: *** represent significant difference between sample and the population statistic.

is slightly older, better educated, and has a higher income, which is expected for a survey conducted online (Adriaan and Jacco, 2009; Szolnoki and Hoffmann, 2013). At the same time, we observe a slightly higher share of female and black respondents. As vaccine hesitancy is often associated with a political worldview (Fridman et al., 2021), we also control for self-reported measures of political ideology. We find that 28% of our sample identifies as republicans, 65% consider religion to be fairly important, and 35% identify with a conservative worldview.[1]

Detailed information regarding race, education and worldview variables are available in Table A1 in the Appendix.

The survey was conducted in five waves, with the first wave starting in October 2020, and the last one ending in October 2021. The sample size for a given wave varied from 400 to 1500 respondents. In total, the survey was completed by 5671 respondents. Dates and sample sizes were mostly a function of availability of funds from the agency sponsoring this research. That being said, wave 4 coincides with when it was clear that some vaccines would require two doses, and wave 5 was launched when the need for boosters was evident. In Table 2 we present the shares of individuals who are planning to get vaccinated, compared across different waves, as well as the share of the individuals who already got vaccinated. As can be seen, the portion of individuals who are planning to get vaccinated is increasing over time, which may suggest a decreasing vaccine hesitancy as the pandemic was progressing. At the same time, the share of individuals who would get vaccinated in the first year of the vaccine being available is larger than the share of those who would get vaccinated in the first 3 months. This suggests that at least some individuals would prefer to delay their vaccination, probably to see first whether others experience side effects or complications (Mouter et al., 2022). It is important to mention that during the first 3 waves, vaccines were not available yet to the American population. As waves 4 and 5 were conducted when the vaccine was already available, we also asked respondents whether they had been vaccinated. The shares in our sample are slightly larger than those reported for the general US population in the same time period. In wave 4, 26.5% of respondents reported being vaccinated, whereas in the fifth wave the corresponding number was 75.4%. In comparison, in the general population, 19.7% and 64% of individuals were vaccinated at the time, respectively. We also inquired whether respondents are planning to get a flu shot and whether they would volunteer for clinical trials for the vaccine. The former may control for some spillover effects between the two diseases, whereas the latter may help to identify respondents with a high level of trust in the vaccine development process.

In Table 3 we provide information regarding indicator variables that we use to identify latent factors. The upper part of the table contains questions from the PVDQ scale, which were divided into two latent factors: Germs aversion and Perceived (un)infectability, based on Díaz et al. (2016). One question was edited to refer to "public bathrooms" instead of "public telephones" as the latter is not common in the US anymore. The third latent factor is the respondent's health status. We use two questions related to the perception of their own health before and during the pandemic, as well as a third question regarding whether they have any underlying disease that would put them at a higher risk. In the third question, respondents could opt out by saying that they "don't know" or "prefer not to tell". As such, the variable had four levels that were not ordered. To account for the nature of this variable, in the model we recoded this covariate as two binary variables. The first variable indicates whether the respondent opted out from the question or not, and the second indicates whether the respondent reported having an underlying disease or not (conditional on not opting out).

The factors utilized in the current study could be interpreted in a light of the Health Belief Model (Rosenstock, 1974; Janz and Becker, 1984), which was proposed to explain a preventive health behavior, such as vaccination. The basic formulation of the model identifies four dimensions of the health-relevant beliefs, namely: perceived susceptibility, perceived severity, perceived benefits, and perceived barriers. The perceived infectability factor is a measure of the susceptibility to infectious diseases, and therefore it should increase the probability of getting vaccinated. The poor health status can be associated with the perceived severity as individuals in poor health are more likely to experience severe COVID-19 symptoms. As such, it should also increase individuals' likelihood of vaccination. Finally, germs aversion measures individuals' discomfort in situations associated with a high potential for pathogen transmission (Duncan et al., 2009). As such, individuals with a higher level of germ aversion may have an increased discomfort caused by the pandemic, for example, increased level of anxiety and increased perception of the importance of social distancing measures (Makhanova and Shepherd, 2020). Because of that, we argue that such individuals may perceive vaccination as more beneficial than individuals with lower germs aversion level.

2.2. Choice experiment

The choice experiment elicited respondents' preferences for COVID-19 vaccines. Each choice situation consisted of two alternatives, each of them representing a different vaccine option. In Fig. 1 we provide an example of a choice card from the survey. Note that the given alternative does not necessarily refer to a single shot, as some vaccines require more than one dose to obtain full protection. Respondents also had an opportunity to opt out from the choice altogether in case they were not

[1] The importance of religion was measured on 5 point Likert ranging from "not at all important" to "extremely important". It was transformed to binary variable equal to 1 if respondent marked 3 or more. The conservative worldview was also measured on 5 point Likert scale ranging from "very conservative" to "very liberal". It was transformed to binary variable equal to 1 if respondent marked 2 or less.

[2] Population age-related variables were rescaled, so that they would represent population older than 18, as this correspond to our sample data.

[3] In the first three waves this question was framed as "Are you getting a COVID-19 shot within the first 3 months of a vaccine being introduced?" It was recoded to the binary variable equal to 1 if the respondent answered "Yes" or "Probably yes". In the fourth wave the wording of the question was changed from "being introduced" to "being available to you". In the fifth wave the question was dropped. In the data analysis, we assume that it is equal to 1 in the fifth wave if the respondents reported that he has already been vaccinated or is planning to. Analogous reasoning was applied to the next variable, regarding vaccination within 12 months period.

[4] These indicators were dropped when estimating HCM as the respective latent variables had the lowest effect on them. These effects were still significant, but the covariates were dropped to speed up the estimation. Dropping them did not significantly affect how latent factors were affecting preferences in the CE.

Table 2
Comparison of vaccination-related plans between different waves of the study.

	Wave 1: 10.22.20–11.13.20	Wave 2: 11.19.20–11.24.20	Wave 3: 12.19.20–12.22.20	Wave 4: 03.04.21–03.10.20	Wave 5: 09.27.21–10.04.20
Plans to get vaccinated in the first 3 months of the vaccine being available[3]	0.368	0.404	0.568	0.724	0.845
Plans to get vaccinated in the first year of the vaccine being available	0.448	0.520	0.705	0.782	0.845
Vaccinated with at least 1 dose	0.000	0.000	0.000	0.265	0.754
Plans to get a flu shot	0.616	0.608	0.705	0.700	0.674
Would volunteer for a clinical trial for COVID vaccine	0.315	0.360	0.454	0.424	0.403
Sample size	1260	1049	414	1421	1527

satisfied with any of the available vaccines. The choice to opt out was presented concurrently with the two hypothetical vaccines, at the bottom of every choice card. Every respondent completed 7 choice cards. Details of design of the experiment are reported in Daziano (2022), where data from the first two waves were used.

Hypothetical vaccines differed in terms of several attributes which are listed in Table 4. As the survey was conducted in several waves, the design of CE was slightly modified over time to better represent attributes of available vaccines. For example, and as mentioned above, wave 4 was launched when it was clear that some vaccines would need 2 doses (which made us introduce number of doses and time between doses as experimental attributes), and wave 5 was launched when the need for boosters was made evident (which made us introduce the availability of boosters as an additional attribute). In fact, roll out of the first wave started before any announcement of actual vaccines and their clinical studies. In line with most concurrent CE vaccination studies, we control for cost,[5] efficiency, and risk of side effects. These vaccine features were set at levels similar to those against the seasonal flu, findings from the literature review, and outcomes from an online focus group. Apart from these common vaccine features, we included in the design the number of months of protection (Bansal et al., 2022), number of doses (Chu and Liu, 2021), country of origin (Kreps and Kriner, 2021), and recommending institution (McPhedran and Toombs, 2021). Levels for some of these features, such as months of protection, were not known for COVID-19 vaccines when data collection started and in fact changed over time. We also included the number of months since the development of the given vaccine to investigate whether rapid development of the vaccines could be a factor increasing vaccine hesitancy. Furthermore, we included the number of days between doses to see whether that waiting time matters to individuals for the vaccines that require more than one dose. Especially in the US, there is a significant difference between the number of individuals who were vaccinated with only one dose, vs. those who were vaccinated with both. Finally, and as mentioned earlier in the later waves of the study, we also added booster availability, as it became clear that one-time vaccination would not be sufficient. With the addition of experimental attributes (number of doses, booster availability), some features were dropped due to decreasing relevance. For example, whereas number of months since development was a feature that individuals in our focus group were concerned about at the early stages of vaccination, when it was clear that vaccines were safe that worry decreased. The initial design (waves 1–3) followed a Bayesian efficient design with two-way interactions, with priors set from an online focus group (N = 20) and an online pretest (N = 150). Later experimental designs also optimized Bayesian D-efficiency with updated priors from previous waves. A total of 24 choice scenarios were generated in the designs, with each respondent being randomly assigned to a subset of 7 choice cards. All choice cards presented informative trade-offs across the attributes, meaning that the design did not include any dominating alternative. To ensure good quality of the responses, individuals who completed the choice experiments in less than half the median time were replaced and so did those who always answered the same alternative for all 7 choice occasions. The number of respondents that needed to be replaced was very small (in the tens for each wave). Ordering of attributes was not randomized, but the order of the alternatives was.

2.3. Pandemic conditions

As the choice experiment was conducted in different moments of time throughout the pandemic, this dynamic feature of data collection allows us to control for how the pandemic evolved over time. Specifically, we incorporate into the HCM three covariates that are plotted in Fig. 2. The first dynamic variable is the number of COVID-19 active cases, which has been the main indicator of development of the pandemic that was reported in the media.[6] A higher number of cases not only increases the probability of getting infected, but also may render more difficult to obtain medical help due to the overload of the healthcare system. Previous research shows that the number of COVID-19 cases is directly related to the fear of the disease (Raciborski et al., 2021). In Fig. 2 we observe that the peak of the number of cases was during the first three waves of the current study, then it heavily decreased for the fourth wave, and then it heightened again during the fifth wave, although to a lesser extent than before.

The next two variables that we employ relate directly to vaccination. First, we use the number of individuals who have completed the vaccination cycle.[7] This is of course equal to zero through the first three waves of our study, but then it starts increasing once vaccines became available. Investigating the effect of this covariate on vaccination preferences may reveal how individuals' behavior is affected by the preferences of other members of society. A positive effect may indicate that herding or social norms act as drivers of behavior, whereas a negative effect may suggest some free-riding behavior (Agranov et al., 2021). As

[5] Although the vaccine has been made available for free, the introduction of cost allows researchers to derive welfare measures that are not subject to scale issues. Whereas marginal utilities across studies (or waves, or designs within the same study) cannot be compared as are subject to potentially different scale (which is associated with the variance of unobserved effects), the analysis of scale-free estimates, such as willingness to pay, allows researchers to compare results across waves.

[6] We use a 7-day moving average data from https://covid.cdc.gov/covid-data-tracker. In the model we incorporated the number of cases in the given state that respondent lives in, but for brevity we report the number of cases for the whole US in Fig. 2. We also checked whether the number of cases relative to the population of a given state may work better, but we did not observe any qualitative differences in the model results.

[7] We obtain data from https://data.cdc.gov/Vaccinations/COVID-19-Vaccination-Trends-in-the-United-States-N/rh2h-3yt2. Completed cycle refers to individuals vaccinated with two doses of the vaccines that require two doses, or a single dose of those which require only one dose. We also compared the results with using only first dose data, but it did not change our results qualitatively.

Table 3
Summary of the indicator variables which were used to identify latent factors in the HCM.

Germs aversion		Perceived (un)infectability	
It really bothers me when people sneeze without covering their mouths	Measured on 7-point Likert scale, from "Strongly agree" to "Strongly disagree"	If an illness is going around, I will get it.	Measured on 7-point Likert scale, from "Strongly agree" to "Strongly disagree"
I am comfortable sharing a water bottle with a friend.		My past experiences make me believe I am not likely to get sick even when my friends are sick[3]	
I don't like to write with a pencil someone else has obviously chewed on		I have a history of susceptibility to infectious diseases	
I prefer to wash my hands pretty soon after shaking someone's hand		In general, I am very susceptible to colds, flu, and other infectious diseases	
I dislike wearing used clothes because you don't know what the last person who wore them was like[§]		I am more likely than the people around me to catch an infectious disease	
My hands do not feel dirty after touching money		I am unlikely to catch a cold, flu, or other illness, even if it is going around [3]	
It does not make me anxious to be around sick people		My immune system protects me from most illnesses that other people get[3]	
I avoid using public bathrooms because of the risk that I may catch something from the previous user[3]			
Poor health status			
How would you describe your current health status?	Measured on 5-point scale from "Excellent" to "Very poor"	Did the respondent answer: "Do you have an underlying health condition that puts you more at risk if contracting COVID-19?"	Equal to 1 if answered, equal to 0 if respondents marked instead "I don't know" or "Prefer not to tell"
How would you describe your health status in 2019?		Do you have an underlying health condition that puts you more at risk if contracting COVID-19?	Equal to 1 if "yes", equal to 0 if "not", treated as missing if the respondent did not want to answer

last dynamic variable, we utilize Google trend data related to weekly Google searches for the word "vaccine" in the US.[8] The variable is measured in relative terms to the peak frequency of Google searches in a given period (in our case the last week of march, 2021). The Google trend data measures the population's interest in the given topic and could be an indicator of demand (Chang and Yin, 2021). This variable has shown to be a useful tool in the epidemiological context (Anggraeni and Aristiani, 2016).

[8] Data were obtained from https://trends.google.com/. We also tried data for different phrases such as "COVID vaccine", "Vaccine side effects", and "Pfizer". We did not find much difference in the results of the model, as all variables were highly correlated.

2.4. Hybrid choice model

To combine choice experiment data with attitudinal questions from the survey we use the hybrid choice modeling framework (Ben-Akiva et al., 2002). Hybrid choice models utilize a structural modeling approach to link latent factors with observed or stated indicator variables as well as individuals' choices. The model used in this study consists of three components: (i) a random-utility-maximization-based choice model, (ii) structural equations, and (iii) measurement equations. We next describe each component in detail.

We employ a latent class logit specification for the choice model to account for discrete unobserved preference heterogeneity. The latent class logit model has the advantage of grouping the population according to their preferences into several discrete segments, which can be useful to inform and guide policy development (Borriello et al., 2021). Instead of assigning individuals deterministically to a given segment, a latent class choice model treats preference heterogeneity as an unobserved categorical variable. To derive further insight into these segments, one can link the probability of belonging to a given class with observed covariates as well as underlying attitudes.

Formally, we assume that the utility function of individual *i*, who belongs to class *c*, for the *j*-th alternative in the *t*-th choice task is given by

$$U_{ijt}^{c} = \delta_i^c ASC_{ijt}^{Opt-out} + \beta_2^c \left(\boldsymbol{\beta}_1^c \mathbf{X}_{ijt} - \text{Cost}_{ijt} \right) + \varepsilon_{ijt}. \tag{1}$$

In this setting, $ASC_{ijt}^{Opt-out}$ denotes an alternative specific constant for the opt-out alternative, Cost_{ijt} denotes an out-of-pocket cost of the vaccine, and $\mathbf{X}_{ijt}$ denotes a vector with all the other attributes utilized in the choice experiment, as summarized in Table 4. As usual, ε_{ijt} denotes a stochastic component, assumed to follow an i.i.d. type I extreme value distribution with constant variance. For convenience we specify utility in WTP-space (Train and Weeks, 2005). Working in WTP-space allows us to easily interpret parameters $\boldsymbol{\beta}_1^c$ in monetary terms. To account for the difference in error terms structure between the opt-out alternative and the vaccines, we model $\delta_i^c \sim N(\mu_{ASC}, \sigma_{ASC})$ as a random effect following a normal distribution.

The probability of choosing alternative *j*, conditional on belonging to class *c* and random effect δ_i^c , is then given by the standard multinomial logit formula

$$P\left(y_{it} = j \middle| C_i = c, \delta_i^c\right) = \frac{\exp\left(\delta_i^c ASC_{ijt}^{Opt-out} + \beta_2^c \left(\boldsymbol{\beta}_1^c \mathbf{X}_{ijt} - \text{Cost}_{ijt}\right)\right)}{\sum_l \exp\left(\delta_i^c ASC_{ilt}^{Opt-out} + \beta_2^c \left(\boldsymbol{\beta}_1^c \mathbf{X}_{ilt} - \text{Cost}_{ilt}\right)\right)}. \tag{2}$$

In formula (2), we use C_i to denote an unobservable variable that indicates to which class a given individual belongs. As this covariate is latent, we need to specify its distribution. Specifically, we model it as a discrete random variable with probability described by the following multinomial logit formula

$$P(C_i = c | \mathbf{LV}_i) = \frac{\exp\left(\boldsymbol{\alpha}_1^c \mathbf{X}_i^{SD} + \boldsymbol{\alpha}_2^c \mathbf{LV}_i\right)}{\sum_s \exp\left(\boldsymbol{\alpha}_1^s \mathbf{X}_i^{SD} + \boldsymbol{\alpha}_2^s \mathbf{LV}_i\right)} \tag{3}$$

where $\mathbf{X}_i^{SD}$ denotes a vector of observable covariates such as socio-demographic characteristics of the respondents or evolving conditions of the pandemic. On the other hand, $\mathbf{LV}_i$ denotes a vector of unobservable latent factors. For identification, we assume that for the last class, $\boldsymbol{\alpha}_1^c$ and $\boldsymbol{\alpha}_2^c$ are equal to 0. The specification in (3) allows us to find the effect of analyzed variables on the probability of belonging to a given segment.

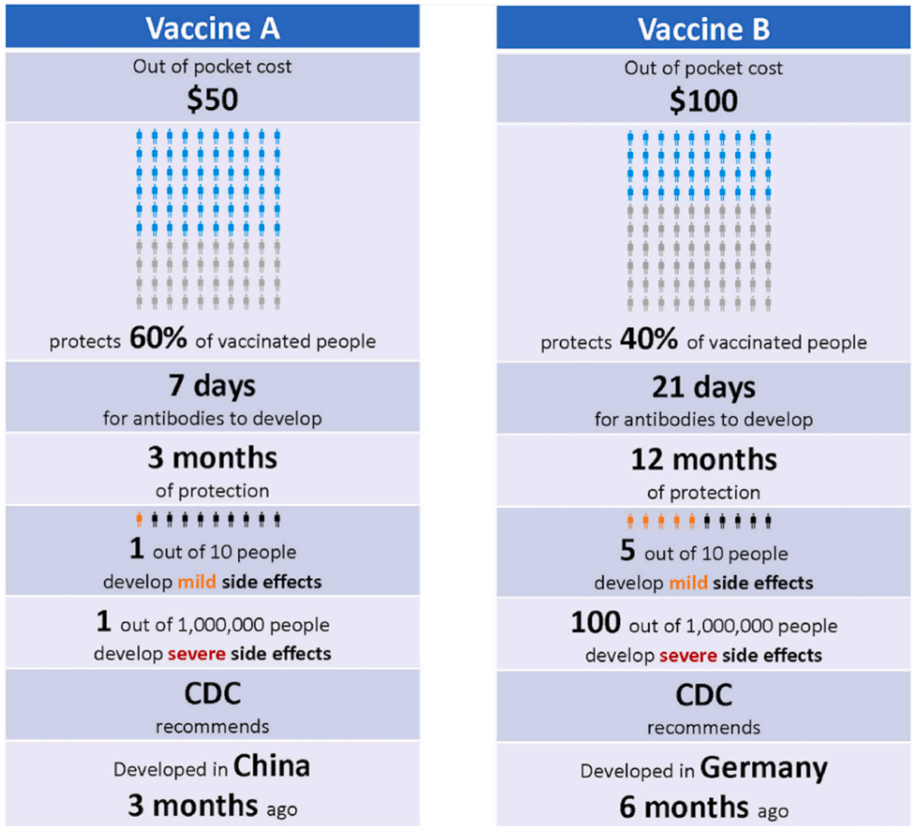

Fig. 1. Sample choice card.

Table 4
Choice experiment attributes and their levels.

	Attribute levels
Out-of-pocket cost ($)	0, 50, 100, 175
% of protected vaccinated people	20, 40, 60, 80, 95
No. of months of protection	3, 6, 12
No. of days for antibodies to develop	7, 14, 21
How many people develop mild side effects (out of 10)	1, 3, 5
How many people develop severe side effects (out of 1 mln.)	1, 10, 100
No. of doses[a]	1, 2
Whether booster is available[a]	Yes, no
Country of origin	USA, UK[c], Germany[c], China, Russia[a]
Recommended by*	Doctor, media, CDC, WHO
No. of months since developed*	3, 6
No. of days between doses[b]	0, 14, 21

[a] These attributes and levels were added after wave 3.
[b] This attribute was added after wave 4.
[c] these attributes and levels were dropped after wave 3.

The second component of the hybrid choice model consists of structural equations in which latent factors are explained by the observable variables. We assume that the k-th latent factor is a linear function of the variables in the vector $\mathbf{X}_i^{SE}$ and an unobservable stochastic term, η_{ik}.

$$LV_{ik} = \frac{\boldsymbol{\gamma}_k \mathbf{X}_i^{SE} + \eta_{ik}}{\delta_k} \tag{4}$$

The error term η_{ik} is assumed to follow a standard normal distribution, whereas δ_k is a normalizing factor, which assures that the variance of LV_{ik} is equal to 1. This normalizing factor facilitates interpretation (e.g., a researcher can easily compare coefficients for different latent factors, as they have the same scale) and estimation (e.g., coefficients for the latent factors do not change much upon adding variables to $\mathbf{X}_i^{SE}$). Covariates in vectors $\mathbf{X}_i^{SD}$ and $\mathbf{X}_i^{SE}$ can, and probably should, overlap (Budziński and Czajkowski, 2022).

The last component of the HCM consists of measurement equations that link answers to attitudinal questions with latent factors. The specific form of this part of the model depends on the distribution of the indicator variables. In the current study, all indicator variables are categorical, and therefore an ordered probit specification was utilized.[9] We denote individual i answer to the n-th item on the m-point scale by I_i^n. We then assume that there exists an unobserved variable, $\widetilde{I}_i^n$, such that

$$\widetilde{I}_i^n = \boldsymbol{\lambda}_n \mathbf{LV}_i + \xi_i^n \tag{5}$$

and

$$\begin{cases} I_i^n = 1 & \text{if} \quad \widetilde{I}_i^n \le \theta_1^n \\ I_i^n = 2 & \text{if} \quad \theta_1^n \le \widetilde{I}_i^n \le \theta_2^n \\ \vdots & \\ I_i^n = m-1 & \text{if} \quad \theta_{m-2}^n \le \widetilde{I}_i^n \le \theta_{m-1}^n \\ I_i^n = m & \text{if} \quad \theta_{m-1}^n \le \widetilde{I}_i^n \end{cases} \tag{6}$$

In (5), ξ_i^n is a measurement error following a standard normal distribution, and $\boldsymbol{\lambda}_\mathbf{n}$ is a vector of parameters to be estimated, measuring how strongly the latent variables affect the answer to the given indicator question. θ's in (6) are the usual threshold parameters that translate the values of the continuous variable, $\widetilde{I}_i^n$, to the ordinal one. The probability

[9] Some variables in Table 3 are actually binary, but it is a special case of the ordered model.

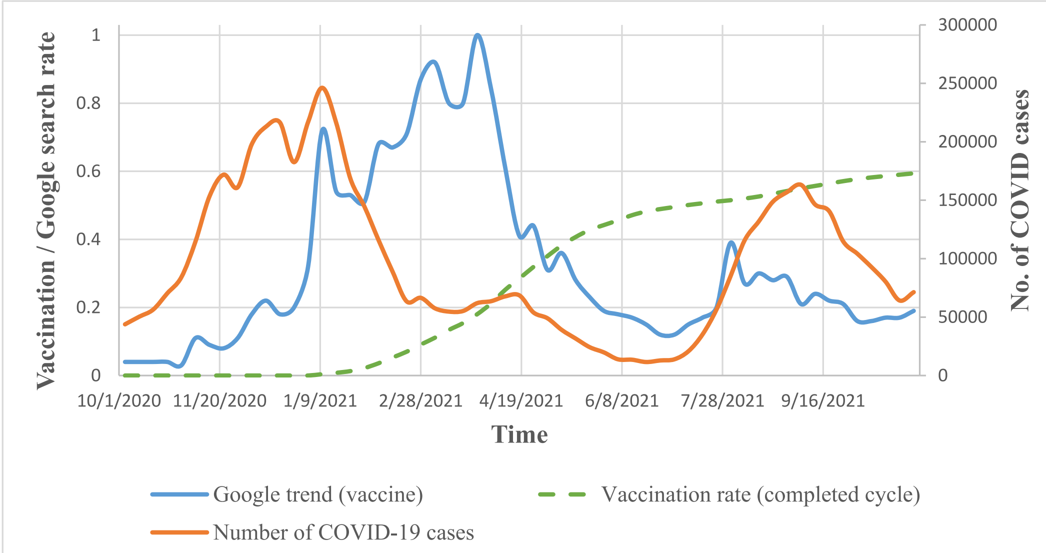

Fig. 2. Dynamic of the covariates describing the conditions of the pandemic.

of indicating in the survey that $I_i^n = k_n$ is then given by

$$P(I_i^n = k_n | \mathbf{LV}_i) = \Phi(\theta_{k_n}^n - \boldsymbol{\lambda}_n \mathbf{LV}_i) - \Phi(\theta_{k_n-1}^n - \boldsymbol{\lambda}_n \mathbf{LV}_i) \quad (7)$$

where Φ is a cumulative distribution function of the normal distribution (we assume $\theta_0^n = -\infty$ and $\theta_m^n = \infty$). Combining (2), (3) and (7) leads to the likelihood function

$$L_i = \int \sum_{c=1}^{C} \left[P(C_i = c | \mathbf{LV}_i, \delta_i^c) g(\delta_i^c) \prod_{t=1}^{T} P(y_{it} | C_i = c, \delta_i^c) \right] \prod_{n=1}^{N} P(I_i^n = k | \mathbf{LV}_i) f(\boldsymbol{\eta}_i) d\boldsymbol{\eta}_i d\boldsymbol{\delta}_i.$$

The likelihood function has a form of multidimensional integral, as the measurement errors, $\boldsymbol{\eta}_i$, as well as random parameters for the ASC, $\boldsymbol{\delta}_i$, are unobserved and therefore need to be integrated out. $f(\boldsymbol{\eta}_i)$ and $g(\delta_i^c)$ are pdf functions of these stochastic components, which follow normal distributions. As the integral above does not have an analytical solution, we use a maximum simulated likelihood method to approximate it by using 2000 scrambled and shuffled Sobol draws (Czajkowski and Budziński, 2019).

3. Results

In this section, we present the results of the estimated hybrid choice model. First, in subsection 3.1, we describe in detail the estimates of the class-specific utility functions. Estimation of the model in WTP-space allows us to conveniently interpret coefficients as marginal willingness to pay (WTP). Next, we describe the effects of a variety of covariates on the class membership probability, as described by equation (3). We focus on variables related to pandemic conditions, plans regarding vaccination, as well as socio-demographic covariates. Finally, results for the latent factors are discussed in subsection 3.3.

3.1. *Choice model*

In Table 5 we report point estimates of the class-specific utility functions as per equation (1). This model was estimated in WTP-space, but for the sake of comparison we also report the preference-space equivalent in Table A2 in the Appendix. We decided to use a latent class logit model with four classes as it had the best fit to the data, based on AIC and BIC.[10] We also report the simulated average probability of belonging to a given class as well as predicted vaccine uptakes in each class. Class membership is distributed rather uniformly (on average), with a mean probability ranging from 19% to 33%, and the second class being the one with the largest expected share.[11] We note, however, that depending on the respondent's characteristics these probabilities vary substantially, ranging from almost 0% to over 95% for each class. Next, we present the predicted vaccination uptake for each class (probability of choosing either one of the vaccines presented in the given choice task). These were calculated for each individual and each choice task in the sample, but we present the minimum, median, and maximum to provide a summary of the distribution of the vaccination probability for each class. Classes 2 and 4 are characterized by high vaccine uptake. Especially for individuals in class 2, the probability of getting vaccinated is always above 90%. In turn, Class 1 is characterized by the highest variation in expected uptake, ranging from 15% to 96%, depending on vaccine characteristics. Lastly, Class 3 has the lowest probability of getting vaccinated, which never exceeds 35.3%. These probabilities are directly associated with estimates of an alternative-specific constant (ASC) for the opt-out alternative. Specifically, the estimate for class 3 is the highest, and the estimate for class 2 is the lowest. We also note that

[10] When compared with equivalent model with 2 or 3 classes. We couldn't obtain reliable estimates for the model with 5 classes, as some coefficients would explode to very large values, indicating some issues with identification or convergence. This could be caused by the very high number of coefficients needed to be estimated. The model with 4 classes has 308 parameters, whereas model with 5 classes would have 353 parameters.

[11] We also report posterior class membership probabilities, that are conditional on the sequence of choices that respondent made. Nonetheless, on average they are very similar to the ones calculated based on equation.

Table 5
Latent class model estimates.

	Class 1		Class 2		Class 3		Class 4	
			Average probability of class membership					
	0.199		0.327		0.192		0.282	
	[0.007]		[0.010]		[0.006]		[0.011]	
			Average posterior probability of class membership					
	0.181		0.346		0.204		0.269	
			Vaccine uptake					
Min.	0.149		0.904		0.050		0.562	
Median	0.616		0.940		0.167		0.862	
Max.	0.956		0.961		0.353		0.967	
			Model coefficients					
Opt out ASC	−0.047		−4.235	***	4.685	***	0.697	**
	[0.364]		[0.242]		[0.893]		[0.275]	
Opt out ASC (std.dev.)	1.511	***	2.383	***	5.486	***	2.169	***
	[0.086]		[0.174]		[0.368]		[0.135]	
Out-of-pocket cost (in 10$)	−0.076	***	−0.009	***	−0.441	***	−0.062	***
	[0.006]		[0.002]		[0.027]		[0.003]	
			Protection (WTP in $)					
% of protected vaccinated people (divided by 10)	33.757	***	27.065	**	2.958	***	74.696	***
	[3.702]		[10.869]		[0.991]		[4.112]	
= 1 if minimal protection satisfied	26.824	**	6.279		0.32		8.007	
	[13.273]		[52.692]		[4.954]		[11.025]	
Number of months of protection	11.815	***	5.592		1.264	***	14.884	***
	[1.424]		[3.732]		[0.451]		[0.969]	
			Side-effects (WTP in $)					
How many people develop mild side effects (1 in 10)	−11.98	***	−24.532	**	−2.294	**	−11.416	***
	[2.918]		[10.813]		[1.039]		[2.098]	
How many people develop severe side effects (1 in 1 mln)	−0.887	***	−0.703	**	−0.017		−1.068	***
	[0.104]		[0.308]		[0.031]		[0.071]	
	Origin (WTP in $)							
Country of origin: UK (base level: USA)	−134.463	***	−193.397	***	−29.583	***	43.823	**
	[15.963]		[69.847]		[6.701]		[18.949]	
Country of origin: Germany (base level: USA)	−145.622	***	−350.686	***	−11.212	**	−36.426	**
	[16.332]		[98.026]		[5.084]		[16.348]	
Country of origin: China (base level: USA)	−641.913	***	−544.65	***	−21.806	***	−56.489	***
	[46.153]		[118.139]		[4.369]		[8.130]	
Country of origin: Russia (base level: USA)	−650.139	***	−281.975	***	−14.876	***	−88.571	***
	[49.445]		[89.102]		[5.264]		[12.608]	
	Recommending institution (WTP in $)							
Recommended by: media (base level: doctor)	−166.268	***	−281.442	**	−11.308	*	−147.684	***
	[22.786]		[120.813]		[6.287]		[23.797]	
Recommended by: CDC (base level: doctor)	10.688		−148.588		13.554	**	18.374	
	[17.793]		[93.679]		[6.734]		[19.467]	
Recommended by: WHO (base level: doctor)	−46.891	***	−67.817		−6.87		−23.816	
	[17.684]		[93.008]		[6.280]		[21.953]	
	Other (WTP in $)							
Number of days for antibodies to develop	−2.802	***	−3.396		−0.564	**	−2.338	***
	[0.718]		[2.454]		[0.256]		[0.544]	
Number of months since developed	6.692	*	22.208	*	0.743		13.466	***
	[3.688]		[13.108]		[1.256]		[4.647]	
= 1 if booster is available	85.811	***	71.972	*	5.53		57.499	***
	[14.058]		[43.093]		[4.869]		[7.362]	
Number of doses	8.751		−102.783		2.601		−30.578	**
	[16.716]		[70.855]		[7.607]		[12.028]	
Days between doses (divided by 10)	0.085		36.562		0.76		8.176	
	[9.889]		[46.768]		[4.322]		[7.692]	

this ASC was modeled as a normally distributed random parameter. For all classes, we observe a significant standard deviation of this random effect, indicating that the latent class specification itself did not sufficiently account for the correlation structure of the stochastic part of the model.

The estimated marginal disutility of cost is significant and negative in all classes, accordingly to what we would expect, and supporting the fact that individuals were attentive to vaccines eventually having an associated expense. The cost parameter also varies across classes, with the third class being the most cost sensitive, and the second class being the least cost sensitive.

All other coefficients in Table 5 can be interpreted in monetary terms, as the model was estimated in WTP-space to ensure comparisons of estimates that are not subject to scale differences (for example, across classes), in addition to the provision of metrics that are measured in monetary terms. For example, individuals in Class 4 have the highest WTP for efficiency and would be willing to pay \$74.7 to increase vaccine efficiency by 10 percentage points. For convenience, we grouped the attributes into several types, so they would be easier to discuss.

First, there are attributes related to protection against COVID-19 as provided by the vaccine. These include the efficiency of the vaccine as well as the number of months of protection. We observe that individuals in class 4 have the highest WTP for both of these attributes, followed by the first class. Surprisingly, even though respondents in class 2 have the highest vaccine uptake, they seem to care much less about the protection provided by the vaccine. Class 3 is characterized by several times lower WTPs than the other classes. In the survey, we also asked respondents about minimal vaccine protection that they would find acceptable. This was incorporated into the model as a binary variable equal to 1 if the given vaccine's efficiency exceeds the reported level. As can be seen, this effect is significant only for respondents in Class 1, with individuals willing to pay \$26.8 more for the vaccine that provides at least minimal acceptable protection.[12]

For the attributes related to possible side-effects, we observe that preferences are relatively similar across classes 1, 2, and 4. In the case of mild side-effects, class 2 seems to have much lower WTP, but the large standard error indicates that it is actually not significantly different from the other two classes. WTP for reduced severe side-effects is also very similar across these classes and ranges from -$1 to -$0.7. As in the case of protection-related attributes, class 3 is much less concerned about potential side-effects.

For the vaccine origin attribute, we identified extremely negative preferences for those vaccines that originated outside of the US. This is especially apparent for classes 1 and 2. For example, individuals in class 1 would be willing to pay $640 less for a vaccine that originated in China, than for the equivalent vaccine from the US. Although a low level of trust in health products from some specific countries may be understandable,[13] we also find highly negative preferences for vaccines from the UK and Germany. This is surprising as the most popular vaccine in the US was developed mainly in Germany by the BioNTech company. However, it is possible that the respondents were not aware of this fact as it was mainly advertised as the "Pfizer vaccine", with Pfizer being a US-based company. On the other hand, in class 4 we actually find a positive effect of a UK-based vaccine, which suggests that at least some individuals would consider a vaccine developed outside of the US.

Regarding the recommendation attribute, we identify a strong distaste for vaccines recommended by the media. At the same time, WHO and CDC are mostly considered to be as reliable as a doctor. The exception is class 1, which revealed negative preferences for WHO recommendation, and class 3, which had positive and significant WTP for CDC recommendation.

Finally, with respect to other attributes, we see strong, positive preferences for booster availability and negative preferences for the number of days that it takes to develop antibodies. The number of doses, and how long ago the vaccine was developed, seem to only matter to individuals in class 4. The number of days between the two doses is the only attribute that we found to be not significant in any class, indicating that this waiting time is not an important inconvenience to the respondents.

To summarize, we observe substantial preference heterogeneity across the four classes that our model had identified. Surprisingly, the second class, which is characterized by the highest uptake, does not usually have the highest marginal WTP (even though this class is the least cost sensitive). This suggests that for these respondents it is important to get vaccinated, but they care much less about the specifics of the vaccine (as long as it originates within the US). Individuals in classes 1 and 4 usually have higher marginal WTP, even though their average probability of getting vaccinated is lower. Lastly, class 3 has usually both, the lowest marginal WTP, and the lowest uptake. We argue that class 3 represents the vaccine-hesitant part of the population.

3.2. Class membership probability

In this subsection, we discuss the effects of individual-specific covariates on class membership probabilities, as described by equation (3). As $\boldsymbol{\alpha}$ coefficients in this equation do not have a direct interpretation and need to be interpreted with respect to the reference class, we instead report marginal probability effects (on class membership probability). Furthermore, we omit the results for latent factors, which are discussed in the following subsection.

We first focus on the covariates related to the pandemic conditions described in detail in Section 2.3. The marginal effects are presented in Table 6. For convenience, we summarize each class with respect to uptake and marginal WTP based on the discussion in the previous subsection. The reported marginal effects have quantitative interpretation. For example, 10 mln. increase in the number of vaccinated individuals in the US is correlated with an increase in the probability of belonging to class 3 by 0.004. This is mostly compensated by the decreased probability of belonging to classes 1 and 2. Therefore, we observe that an increase in the vaccination rate is actually associated with an increase in individuals' vaccine hesitancy. Furthermore, we can infer that this switch in class membership probability will decrease the uptake and marginal WTPs as these are the lowest (in absolute values) in class 3 for most of the attributes. These effects are illustrated in Fig. 3, together with the effects of other pandemic-related variables. We observe a decreasing probability of getting vaccinated as the number of vaccinated individuals increases in the population, as well as decreased WTP for vaccine efficiency. On the other hand, for the attributes that were considered to be negative (e.g. experiencing mild side effects), we observe less negative preferences. For the number of COVID-19 cases in the state of respondents' residence, we find a negative association effect on the probability of belonging to class 1, and a positive effect on the probability of belonging to class 4. This would correspond to an increased uptake as the number of COVID-19 cases is increasing. As class 1 was also characterized by an extremely negative preference for vaccines that originated outside of the US, an increase in the number of COVID-19 cases could make such vaccines less undesirable (consider the lower-right panel of Fig. 3). Finally, the Google trend related to the vaccines is associated with a decrease in the probability of belonging to class 2, which is then compensated by an associated increase for class 4. This effect could therefore slightly decrease vaccine uptake, but could in turn increase most marginal WTPs. This result could suggest that increased interest revealed in Google searches be related to specific attributes of the vaccine, rather than just a demand for getting vaccinated.

In Table 7 we report marginal effects for the rest of the observable covariates. Specifically, we observe that individuals who are already vaccinated appear as more likely to belong to classes 1 and 4, rather than 2. This result could be interpreted as these individuals caring more about some specific attributes of the vaccine, rather than just about getting vaccinated. This is a similar effect to what we observed for the Google trend. On the other hand, individuals who were planning to get vaccinated in the first 3 months of the vaccine being available, appear as more likely to belong to class 2, and less likely to belong to class 3. This result means that these individuals are less likely to be vaccine-hesitant, and more likely to get vaccinated with limited care about the specific attributes of the vaccine. The difference in preferences between individuals who already got the vaccine, and individuals who are planning to get it in the near future may reflect the difference in the priorities of these two groups. The latter are interested in getting the vaccine quickly to decrease their risk of infection and care less about the specifics of the product itself. The former have already decreased their risk with the vaccine available on the market, so they can focus more on specifics to further improve their protection. On the other hand, it seems that individuals who were planning to get vaccinated in the first 12 months of the vaccine being available, and therefore would like to slightly delay the process, are also more likely to belong to class 4 (rather than 3). This suggests that these individuals could be more likely to get vaccinated, but also that they care more about specific attributes of the vaccine. Maybe these individuals need more time to get vaccinated to learn more about available vaccines or to wait for the vaccine that will be closer to their needs. We also find significant effects for individuals who plan to get a flu shot, and who would volunteer for the COVID clinical trials.

With respect to other covariates, we find that individuals who

[12] We note that the limited effect of this variable could be attributed to the measurement error, that could arise if individuals were not certain about their preferences regarding vaccine efficiency. This could be especially relevant at the beginning of the pandemic, when the efficiency of the vaccines on the market was not really known.

[13] For example, COVID-19 vaccine developed in Russia was subjected to heavy, international criticism, due to lack of proper large-scale trials, hiding relevant data, and potential danger to vaccine recipients (Callaway, 2020; Moutinho and Wadman, 2021).

Table 6
Marginal effects of covariates related to pandemic conditions on class membership probability.

	Class 1		Class 2		Class 3		Class 4	
	Average uptake, high marginal WTP		High uptake, average marginal WTP		Low uptake, low marginal WTP		High uptake, high marginal WTP	
	Pandemic conditions							
Vaccinated individuals (in 10 mln.)	−0.003	**	−0.003	*	0.004	***	0.001	
	[0.001]		[0.002]		[0.001]		[0.002]	
No. of COVID-19 cases (7-day moving avg., in 1000)	−0.004	**	−0.001		0.001		0.005	**
	[0.002]		[0.002]		[0.001]		[0.002]	
Vaccine Google trend	−0.001		−0.012	***	0.002		0.011	***
	[0.002]		[0.002]		[0.002]		[0.002]	

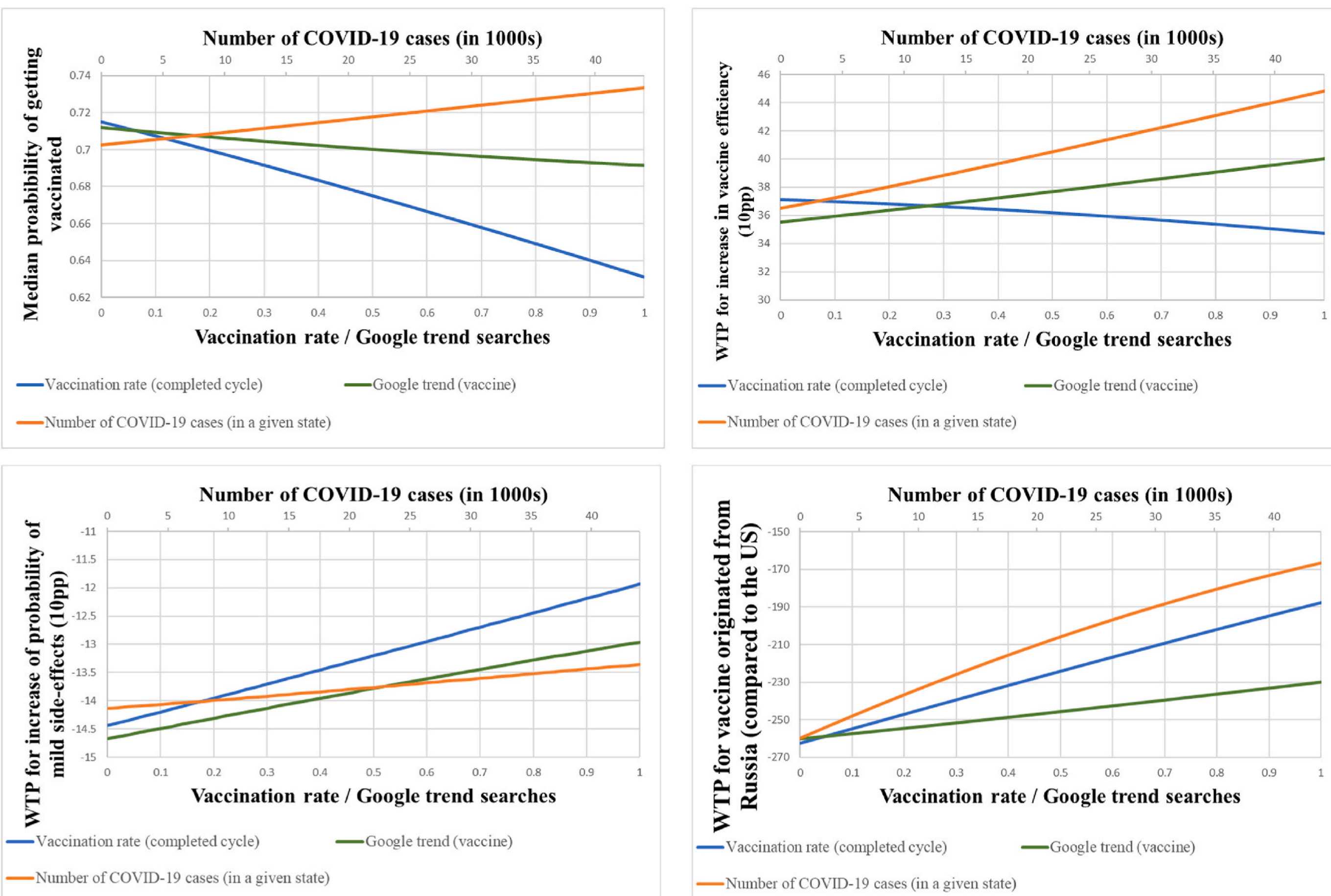

Fig. 3. Associated effect of pandemic conditions on: the median probability of getting vaccinated (upper-left panel), WTP for an increase in vaccine efficiency (upper-right panel), WTP for an increase of probability of mild side effects (lower-left panel), and WTP for vaccine originated from Russia (lower-right panel).

characterize themselves as republicans appear more likely to belong to classes 1 and 3. This is probably driven by the negative preferences toward vaccines from outside of the US in class 1, as well as general vaccine hesitancy in class 3. In contrast, individuals who describe themselves as somewhat conservative, or consider religion to be fairly important in their lives, appear as more likely to get vaccinated. Age is associated with an increase in the probability of belonging to class 2, which suggests that older individuals care more about getting vaccinated rather than the specifics of a given vaccine. Surprisingly, we do not find education to be correlated with the probability of belonging to class 3, which we identify as vaccine-hesitant.

3.3. Latent factors

We now consider the results for the latent factors. The estimates from the measurement equations, as per equation (5), were moved to Appendix A, Table A3. All estimates have expected signs. For the second latent factor, we obtained positive coefficients for items such as "*If an illness is going around, I will get it*", which indicates an increased probability of answering "*Strongly disagree*" for high values of this latent variable. Because of this result, we labeled this latent variable as "*Perceived uninfectability*", rather than "*Perceived infectability*" as in Díaz et al. (2016). Similarly, for the third latent factor we obtained positive coefficients for questions such as "*How would you describe your current health status?*". This means that the third latent variable increases the chance of answering "*Very poor*", so the latent variable was labeled "*Poor*

Table 7
Marginal effects of the rest of observable covariates on class membership probability.

	Class 1	Class 2	Class 3	Class 4
	Average uptake, high marginal WTP	High uptake, average marginal WTP	Low uptake, low marginal WTP	High uptake, high marginal WTP
Vaccination-related plans				
Already vaccinated	0.047* [0.024]	−0.086*** [0.029]	−0.013 [0.026]	0.052* [0.031]
Planning to get vaccinated in first 3 months	−0.016 [0.027]	0.093*** [0.026]	−0.057** [0.027]	−0.02 [0.030]
Planning to get vaccinated in first 12 months	0.003 [0.028]	0.071*** [0.026]	−0.25*** [0.030]	0.176*** [0.029]
Planning to get flu shot	0.046*** [0.015]	0.016 [0.018]	−0.101*** [0.013]	0.038** [0.019]
Would volunteer to clinical trial	−0.031** [0.015]	0.119*** [0.020]	−0.078*** [0.016]	−0.01 [0.019]
Other socio-demographic variables				
Republican	0.055*** [0.006]	−0.032*** [0.007]	0.03*** [0.005]	−0.053*** [0.008]
Religion is fairly important (or more)	−0.039* [0.022]	−0.106*** [0.036]	−0.027 [0.022]	0.172*** [0.041]
Conservative (somewhat or more)	−0.024* [0.014]	0.065*** [0.017]	−0.036*** [0.012]	−0.005 [0.018]
Age (divided by 10)	−0.011 [0.008]	0.022*** [0.007]	0.003 [0.005]	−0.013 [0.009]
Older than 65	0.003 [0.011]	0.024** [0.011]	−0.009 [0.009]	−0.019 [0.014]
Male	0.009 [0.015]	0.059*** [0.019]	−0.015 [0.011]	−0.053*** [0.020]
How many children	0.038* [0.022]	−0.048* [0.025]	−0.009 [0.014]	0.019 [0.026]
How many elderly	0.004 [0.014]	−0.002 [0.017]	−0.024** [0.011]	0.022 [0.018]
Currently employed	0.054*** [0.019]	−0.023 [0.019]	−0.009 [0.013]	−0.022 [0.023]
Has a driving licence	0.005 [0.014]	0.064*** [0.016]	0.003 [0.011]	−0.071*** [0.018]
Education (has a degree)	0.01 [0.017]	0.05*** [0.018]	0.003 [0.013]	−0.063*** [0.021]
Black	0.001 [0.019]	0.006 [0.021]	−0.005 [0.015]	−0.001 [0.023]
Non-white	−0.04* [0.022]	0.001 [0.023]	0.029 [0.018]	0.011 [0.026]
Income	0 [0.002]	0.003* [0.002]	−0.004*** [0.001]	0.001 [0.002]

Table 8
Estimates of structural equations for the latent factors.

	LV 1 (Germ aversion)	LV 2 (Perceived uninfectability)	LV 3 (Poor health)
Pandemic conditions			
Vaccinated individuals (in 10 mln.)	−0.02 [0.024]	0.083*** [0.021]	0.001 [0.021]
No. of COVID-19 cases (7-day moving avg., in 1000)	−0.006 [0.019]	0.008 [0.013]	−0.031** [0.016]
Vaccine Google trend	−0.029 [0.018]	0.038** [0.016]	−0.018 [0.016]
Vaccination-related plans			
Already vaccinated	0.051* [0.026]	0.074*** [0.024]	0.049** [0.023]
Planning to get vaccinated in first 3 months	−0.067** [0.031]	−0.147*** [0.028]	−0.051* [0.029]
Planning to get vaccinated in first 12 months	0.124*** [0.030]	−0.004 [0.028]	−0.004 [0.027]
Planning to get flu shot	0.043** [0.019]	−0.142*** [0.017]	0.047*** [0.017]
Would volunteer to clinical trial	−0.132*** [0.020]	−0.123*** [0.018]	−0.065*** [0.018]
Other socio-demographic variables			
Republican	−0.005 [0.019]	0.105*** [0.017]	0.006 [0.017]
Religion is fairly important (or more)	0.011 [0.017]	−0.121*** [0.015]	−0.068*** [0.016]
Conservative (somewhat or more)	−0.128*** [0.019]	−0.039** [0.017]	−0.071*** [0.017]
Age (divided by 10)	0.178*** [0.026]	0.2*** [0.024]	0.262*** [0.023]
Older than 65	−0.005 [0.026]	0.106*** [0.024]	−0.144*** [0.024]
Male	−0.313*** [0.018]	−0.012 [0.016]	−0.09*** [0.016]
How many children	−0.025 [0.018]	−0.085*** [0.017]	−0.024 [0.017]
How many elderly	−0.07*** [0.020]	−0.048** [0.019]	−0.026 [0.017]
Currently employed	−0.042** [0.019]	0.034* [0.019]	−0.176*** [0.018]
Has s driving licence	0.038** [0.016]	0.048*** [0.014]	−0.056*** [0.015]
Education (has a degree)	−0.047*** [0.018]	−0.002 [0.017]	−0.142*** [0.017]
Black	0.058*** [0.018]	0.065*** [0.016]	−0.044*** [0.016]
Non-white	0.05*** [0.016]	0.03** [0.015]	0.028* [0.015]
Income	−0.045** [0.021]	−0.046*** [0.017]	−0.143*** [0.017]

health".

To obtain a richer insight into the interpretation of the latent variables we provide in Table 8 the point estimates of the parameters of the structural equations (equation (4)). We note, however, that the design of the study does not allow us to conclude that these effects represent causal relationships, as they are only based on correlations.

With respect to germ aversion, we observe that all variables related to the pandemic conditions are insignificant. On the other hand, individuals who are planning to get vaccinated quickly and would volunteer for clinical trials are on average less germ averse, whereas for individuals with plans of getting vaccinated in a year, and with plans for getting a flu shot, we observe the opposite effect. This may indicate that individuals who are averse to germs may also be more concerned about the safety of the vaccine, and therefore they would rather delay their vaccination appointment. Furthermore, we observe that conservative, less educated, male, white, with higher income, and younger individuals are less germ averse.

For the latent perceived uninfectability, we find a positive relationship between no. of vaccinated individuals and the vaccine-related google trend. We also observe a similar effect for individuals who have already got vaccinated. On the other hand, this attitude is negatively associated with plans for getting a COVID-19 vaccine or a flu shot. This suggests that perceived uninfectability is closely related to the perceived risk of getting infected. Being vaccinated and higher rate of vaccinated individuals decrease the risk and therefore increase this attitude. At the same time, a low perception of this risk may be negatively associated with the propensity to get vaccinated. Apart from these results, we find, among others, that individuals identifying as republicans, older, black, and with lower income have a higher perception of uninfectability.

For the third latent factor, we find that the number of COVID-19 cases is correlated with an improvement in the perception of respondents' own health. It is not clear how to interpret this effect. Possibly, individuals rate their health relative to others. A high number of infections in the place of residence could then motivate respondents to think better about their own health status. Furthermore, we observe that vaccinated individuals report better health, similarly as individuals who plan to get a flu shot. With respect to other covariates, older individuals appear as more likely to report poor health, whereas conservatives, religious individuals, males, black, individuals with higher education, and individuals with higher income are more likely to report a better health status.

Analogously to the previous subsection, we report marginal effects in Table 9, to see how these latent factors are associated with class probability membership.[14] Contrary to the marginal effects for other variables (e.g., in Table 7), these estimates do not really provide much quantitative insight as the scale of latent factors is arbitrary and a result of the adopted normalization (Chorus and Kroesen, 2014). As such, a unit increase in the latent variable is not very informative. To tackle this issue we employ two approaches. First, we use Hess et al. (2018) sample enumeration approach, which is reported in the second part of Table 9. Here, we evaluate the difference in the probability of each class membership if the attitudes would change accordingly to some observed covariate. For example, from the results in Table 8, we can see that individuals of different ages hold different attitudes toward germs. With the sample enumeration approach, this variation can be utilized to say, for example, how would the class membership probability change if the germ aversion would increase from the level that is on average held by the 25 years old to the attitude level that is on average held by the 75 years old. For the second class, the effect is a decrease in the probability of 0.016 (1.6 percentage points). In Table 9 we report the marginal effects due to change accordingly to age and gender as these two covariates explained the most variation in the structural equations of the latent factors.

The second approach links the latent factors back to the indicator variables (consider Table 3). Specifically, we utilize individual-specific posterior distribution to estimate what is the expected value of the latent factor for the given individual, given their answers to the indicator questions (Sarrias, 2020). The marginal effects in the last part of Table 9 then estimate how the class membership probability would change if the value of the latent factor would increase accordingly to the expected change if all the relevant indicator variables would change by one unit.[15] Using this approach, the marginal effects are not dependent on the employed normalization and could be replicated if one would use the same scale in a different study. We believe that this technique could be a useful alternative to the sample enumeration approach, for example, when structural equations do not explain much variation in the latent factors. In Fig. 4 we use the same approach to illustrate how these effects translate to the predicted vaccine uptake and several chosen marginal WTPs. For example, for germ aversion, we consider 7 levels of the latent factor, which correspond to its expected value given that answers to all relevant indicator variables are of the same level. So level 1 corresponds to the expected value of the latent variable if all answers to the indicator variables that are positively (negatively) correlated with it would be equal to 1 (7). Level 2 corresponds to all the answers equal to 2 (6), and so on.[16]

Germ aversion is associated with a decrease in the probability of belonging to classes 2 and 3, which in turn increases the probability of belonging to the fourth class. Decreased probability of belonging to vaccine-hesitant class 3 correlated with an increase in germ aversion, ultimately seems to increase the probability of getting vaccinated, although very slightly. At the same time, an increased probability of belonging to the fourth class is related to a large increase in WTP for vaccine efficiency.

Perceived uninfectability is correlated with a decrease in the probability of belonging to class 2, which is then compensated by an associated increase in membership in classes 3 and 4. This result would translate to a strictly negative effect on vaccine uptake, mostly driven by increased membership in the vaccine-hesitant class. Nonetheless, the effect on the attributes would be mixed, as individuals in class 3 generally care less about the whole set of vaccine attributes, whereas in class 4 individuals appear to often care more than those in the second class.

Finally, poor health is associated with a decrease in the probability of belonging to the second class and an increase in the probability of belonging to the fourth class. Because of this result, poor health would have a very limited effect on vaccine uptake, but would mostly increase marginal WTP metrics.

4. Limitations

Although limitations have been acknowledged throughout the article, this section summarizes caveats of the data collection strategy and modeling approach that frame results in a very specific context. First, recruitment of participants was performed through a panel of respondents of on-line surveys acquired from Qualtrics. Compared to other studies by the authors with data from both professional panels of respondents and intercept surveys in other contexts, the proportion of covariates that resulted in statistically significant estimates is high, supporting the fact that respondents were highly motivated to participate and engaged in the survey. In fact, an open-ended question at the end of the survey asking for comments also supports that respondents were engaged and motivated to answer accurately. Responses were long and highlighted that respondents found the survey timely and engaging.

Choice experiments offer an insight on stated responses, but intended actions may differ from actual behavior. In this particular case, the presented vaccines were clearly framed within a hypothetical context where only two vaccines would be available, that individuals would have a choice about which vaccine to receive, and that out-of-pocket costs could be involved (although a level representing a free vaccine was considered). In practice, vaccines have been offered for free and individuals did not have a choice of which vaccine they would receive. Furthermore, over the first waves, vaccines were still under development, and it was not until data collection was completed that information was available about duration of protection being relatively short-lived. In this sense, there have been multiple sources of uncertainty regarding vaccines and their features. But in this context, the hypothetical framing of the experiment is a strength as it offers a scenario of presumed certainty about vaccine attributes. With respect to welfare measures, since the vaccine ended up being offered for free, a fact that was not clear when designing the experiment and over early data collection, the estimates of willingness to pay need to be interpreted with care as they may be biased. However, willingness to pay estimates are free of scale issues and thus can be compared across waves and studies. Furthermore, the estimates offer insights of ranges of willingness to pay that could be considered fair by individuals in a possible future scenario where people may face out-of-pocket expenses.

The use of a hybrid choice model that integrates latent factors into the utility function of individuals offers a behaviorally rich specification with gained insights in associations between attitudes and vaccine uptake, but transfer of results, as well as predictions, become more difficult. Specifically, as the scale of the latent factors is arbitrary, it is difficult to utilize these models for quantitative inference. To address this issue, we utilize the sample enumeration approach proposed by Hess

[14] Marginal effects for the latent variables were calculated for each set of quasi-random draws, and then averaged over them for each individual in the sample.

[15] We consider a direction of a change consistent with an increase in the latent factor. For example, germ aversion is negatively correlated with the first indicator and positively correlated with the second (consider Table A3). As such we consider a unit decrease of the former, and a unit increase of the latter. We also assume that the change cannot exceed the existing scale (for example, indicator cannot increase from 7 to 8 if it was measured on the 7-point scale).

[16] This approach is easier to use when all the indicator variables are on the same scale. In the case of the „poor health" some indicators are binary, and some use 5-point scale. Because of that we calculate only the two most extreme effects.

Table 9
Marginal effects of the latent factors on class membership probability.

	Class 1		Class 2		Class 3		Class 4	
	Average uptake, high marginal WTP		High uptake, average marginal WTP		Low uptake, low marginal WTP		High uptake, high marginal WTP	
LV 1 (Germ aversion)	0.015	*	−0.035	***	−0.022	***	0.043	***
	[0.009]		[0.010]		[0.007]		[0.011]	
LV 2 (Perceived uninfectability)	0.014	*	−0.09	***	0.025	***	0.052	***
	[0.008]		[0.009]		[0.006]		[0.010]	
LV 3 (Poor health)	0.002		−0.033	***	0.000		0.03	***
	[0.008]		[0.009]		[0.006]		[0.010]	
				Sample enumeration				
LV 1 (Germ aversion):								
Age (25 vs. 75)	0.007	*	−0.016	***	−0.01	***	0.019	***
	[0.004]		[0.005]		[0.003]		[0.006]	
Male (vs. female)	−0.009	*	0.021	***	0.013	***	−0.026	***
	[0.005]		[0.006]		[0.004]		[0.007]	
LV 2 (Perceived uninfectability):								
Age (25 vs. 75)	0.011		−0.07	***	0.018	***	0.04	***
	[0.007]		[0.009]		[0.005]		[0.009]	
Male (vs. female)	0.000		0.002		−0.001		−0.001	
	[0.000]		[0.003]		[0.001]		[0.002]	
LV 3 (Poor health):								
Age (25 vs. 75)	0.001		−0.01	***	0.000		0.01	***
	[0.002]		[0.004]		[0.002]		[0.004]	
Male (vs. female)	0.000		0.006	***	0.000		−0.005	***
	[0.001]		[0.002]		[0.001]		[0.002]	
				Unit increase of indicator variables				
LV 1 (Germ aversion)	0.008		−0.02	***	−0.012	***	0.024	***
	[0.005]		[0.006]		[0.004]		[0.006]	
LV 2 (Perceived uninfectability)	0.006		−0.045	***	0.012	***	0.026	***
	[0.004]		[0.004]		[0.003]		[0.006]	
LV 3 (Poor health)	0.002		−0.035	***	0.000		0.034	***
	[0.009]		[0.010]		[0.007]		[0.011]	

et al. (2018), as well as propose a new approach that employs individual-specific posterior distribution of the latent factors to link them back to the observed indicator variables. Furthermore, the interpretation of the latent variables is based on their correlation with the utilized indicator variables and is to some extent arbitrary. For the two of the latent variables we have employed an established psychological scale (PVDQ), and interpreted the obtained factors accordingly with the previous literature (Duncan et al., 2009; Díaz et al., 2016). Nonetheless, it is not clear how these factors relate to some other constructs utilized in the literature. For example, perceived uninfectability seems to be closely related to the perceived risk of infection, which was utilized for example by Leng et al. (2021). The main difference is that the construct that was used by us is more general, i.e. it is not directly associated with a specific disease like COVID-19. It should also be more stable in time, as it measures the general trait of (lack of) susceptibility to infectious diseases, whereas risk of infection with COVID-19 is closely related to the given context (for example, the number of COVID-19 cases). Nonetheless, we found that perceived uninfectability also depends on some contextual factors such as vaccination rate, so further research is needed to clearly distinguish between the different constructs.

Finally, as noted by one of the Reviewers, the results of the HCM could be affected by the order in the choice experiment and attitudinal data were collected. Specifically, the latter questions were asked after the choice experiment. The limitation of the current study is that its design does not really allow us to investigate this issue. As Chorus and Kroesen (2014) point out, individuals tend to align their attitudes with their choices to seem consistent when filling out the survey. It is not clear, however, whether such alignment is a significant issue in this case, as the PDVQ scale that we utilize is neither directly related to COVID-19 nor to the vaccination. So the “alignment” is not really straightforward. On the other hand, it could be the case that because of the context of the survey, individuals could interpret the questions differently than in the prior work. For example, the statement “*If an illness is going around, I will get it*”, could be interpreted in relation to COVID-19, rather than more generally to infectious diseases.

5. Conclusions

In the current study, we have analyzed individuals’ preferences for the COVID-19 vaccine in the US using a choice experiment survey. The fact that the study was conducted in five waves throughout the pandemic, and that the modeling framework combined choice experiment data with attitudinal questions, allowed us to control for a wide variety of factors that are absent in other work. Furthermore, our estimation strategy implements a novel hybrid choice model with a latent class logit kernel and a continuous random effect. In what follows we discuss the results presented in the previous section and provide some implications for public policy.

The latent class logit model that we employed revealed substantial preference heterogeneity within the analyzed sample. We identified four distinct segments of consumers: (i) vaccine-uncertain, with vaccination probability varying heavily depending on vaccine characteristics, especially, on the country of origin, (ii) pro-vaccine with limited interest in the specific characteristics of the vaccine, (iii) vaccine-hesitant, and (iv) pro-vaccine, but with high interest in the specific characteristics of the vaccine, such as vaccine efficiency. For policymakers who aim to incentivize individuals to get vaccinated, segments (i) and (iii) would be of the highest importance. Our results in Table 7 show that low-income republicans may be especially prone to belong to these classes. Estimates show that even for these two groups, trade-offs are still considered by individuals within these classes when presented with vaccination choices. As such, financial incentives could work well to convince consumers in these segments (Fishman et al., 2022). Furthermore, public campaigns for boosters and expected annual vaccination could focus on the specific attributes that individuals in these classes consider to be important, for example, that vaccine was developed in the US. We note, however, that effectiveness of such policies is unfortunately not clear (Sadaf et al., 2013).

In the latent class portion of the model, we identified two consumer classes with extremely negative preferences towards vaccines that originate outside of the US. This is in stark contrast to the results

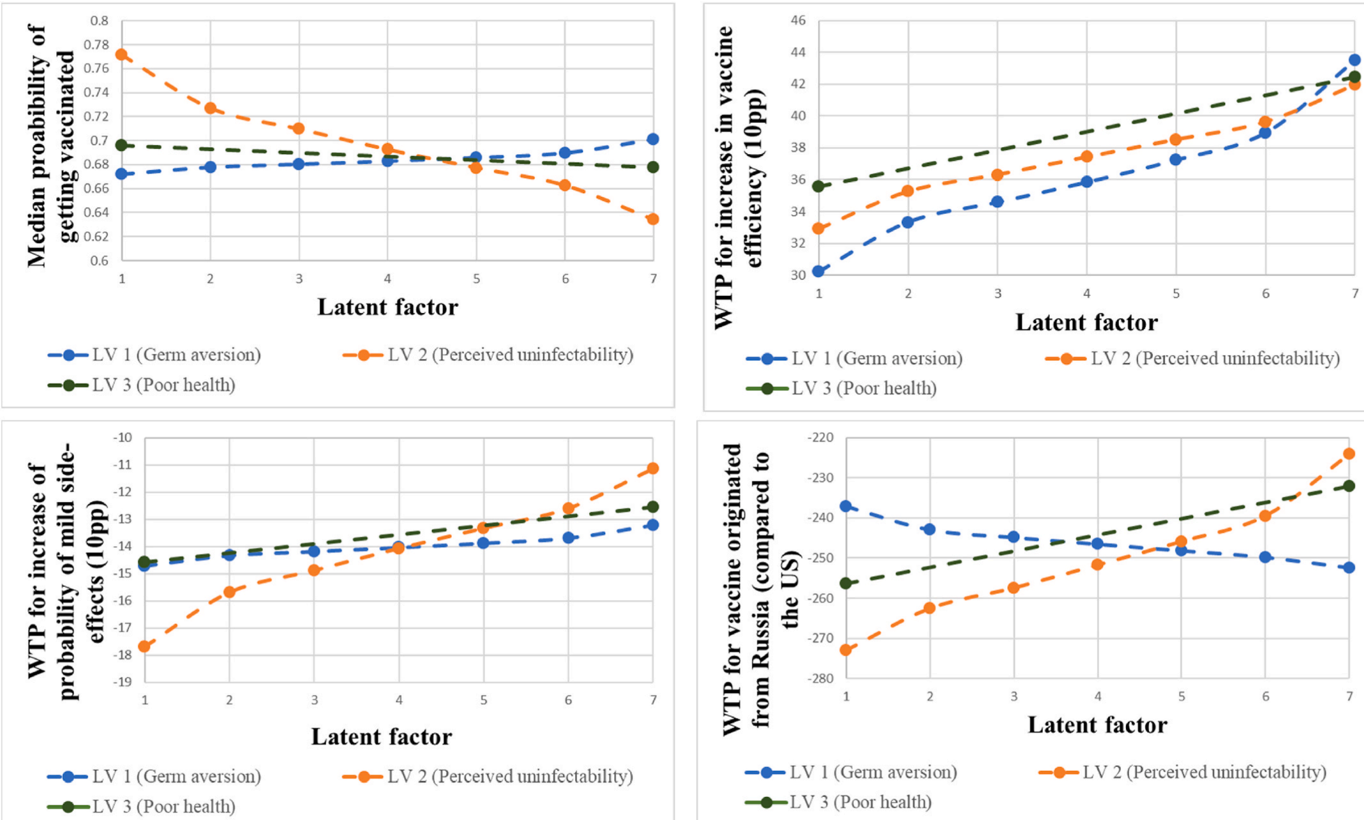

Fig. 4. The effect of latent factors on: the median probability of getting vaccinated (upper-left panel), WTP for an increase in vaccine efficiency (upper-right panel), WTP for an increase of probability of mild side effects (lower-left panel), and WTP for vaccine originated from Russia (lower-right panel). Note: The levels of the latent factor are related to its expected value given a specific answers to all related indicator variables. The first two LVs were based on the 7-point PVDQ scale, and therefore we consider 7 different levels. For the third LV we consider only two most extreme levels of the related indicator variables, which are visualized at points 1 and 7 in the graph.

reported recently by Kobayashi et al. (2021), who find no effect of the vaccine origin. Our results, therefore, seem more in line with Smith (2021), who reports that the US sample prefers US-based vaccines over those from other countries. Nonetheless, it seems that more research is needed to identify the reasons for these extreme preferences, especially with respect to the European countries which comply with international health standards (such as Germany and the UK).

Our analysis revealed that actual vaccination rates are positively related to vaccine hesitancy. This goes against the usual finding that vaccine intention is positively correlated with social norms (Brewer, 2021).[17] There could be several explanations for this effect. First, the general population may not be a relevant comparison group for vaccine-hesitant individuals. As such it may not be a good indicator of the social norm. Second, the vaccination rate in the US may not be high enough to actually “activate” the social norm. Lau et al. (2019) find that the vaccination rate needs to be at least 65% to start having a positive effect on individuals’ intentions. Still, these two arguments only explain the lack of the positive effect, rather than the negative effect that we observe. The third explanation could be that this effect is driven by free-riding behavior, where individuals feel less obliged to vaccinate when they believe to be protected by herd immunity (Hershey et al., 1994; Agranov et al., 2021). Finally, it could be an effect of the increased number of circulating information regarding vaccine side effects following the vaccine roll-out (Diaz et al., 2021). Nonetheless, the identified relationship shows that there is no “*snowball effect*”, where some public interest in vaccination encourages other individuals to vaccinate as well. Our result, therefore, highlights the importance of policy to manage and incentivize individuals to vaccinate. On the other hand, we observed that an increased number of COVID-19 cases may cause individuals to switch from the vaccine-uncertain segment to a pro-vaccine one. This illustrates that individuals react to the pandemic conditions and that they may be more likely to vaccinate when the infection rate is higher. Finally, the Google trend data does not seem to affect vaccine hesitancy much, but it increases individuals’ marginal willingness to pay. As such, Google trends do not necessarily seem to be an indicator of vaccination demand but could reveal public interest in some vaccine characteristics, such as effectiveness or side effects. One could try to use more detailed search queries, although we found these to be highly correlated with each other which renders the analysis difficult. In fact, there is more research needed on the framework of incorporating Google trend data into stated preference research.

The use of a hybrid choice modeling framework allowed us to incorporate the Perceived Vulnerability to Disease Questionnaire into our model. To our knowledge, this is the first application that combines this scale with choice experiment data. We find that both factors identified by this scale affect individuals’ vaccination choices. Germ aversion only slightly increases individuals’ likelihood of getting vaccinated, but has a large positive effect on some marginal WTP, especially for the vaccine efficiency. Previous research shows that this attitude is positively correlated with pandemic-related anxiety as well as social distancing behavior (Makhanova and Shepherd, 2020). As such, it is surprising that we find only a limited effect of germ aversion on vaccination uptake. Possibly, the vaccination appointment itself can be anxiety-inducing for germ-averse individuals, as it may require standing

[17] We note, however, that not all studies identify such an effect (e.g., Sinclair and Agerström, 2023).

in the queue in the pharmacy. This explanation is consistent with our finding that individuals who stated that they would like to get vaccinated quickly are on average less germ averse. Nonetheless, germ-averse individuals are willing to pay a premium to increase the level of protection, which is consistent with our expectations.

With respect to the second attitude, we find that the perceived uninfectability highly decreases vaccine uptake. This is an expected result, as it measures a belief in the lack of susceptibility to infectious diseases. Furthermore, this attitude is positively associated with probability of belonging to the vaccine-hesitant class. As such, perceived uninfectability may be related to the belief in “natural immunity”, which is one of the leading arguments of individuals hesitant to vaccinate (Taylor et al., 2020). Therefore, it seems that the PVDQ scale is a promising instrument to account for such an attitude in applied research. Nonetheless, even though we find that perceived uninfectability decreases the probability of getting vaccinated, it can also increase marginal WTP for certain attributes. As such, the relationship between this belief and individuals’ preferences may be more complex than expected. Nonetheless, in order to increase the vaccination rate, a public campaign addressing the “natural immunity” argument, maybe a worthwhile undertaking.

Finally, stated health status seems to have a very limited effect on vaccine uptake. As individuals in poor health are at higher risk of hospitalization and death from COVID-19, the lack of this effect shows that it should not be assumed that vulnerable citizens will voluntarily take necessary precautions. On the other hand, we see a positive effect of poor health status on marginal WTPs. Therefore, individuals in poor health would be willing to pay extra to get a vaccine that provides, for example, higher effectiveness.

From the modeling perspective, we find that an HCM specification with latent classes and random effect outperforms the analogous HCM specification with only a latent class structure. This result highlights the importance of combining different forms of unobserved preference heterogeneity to capture the complexity of consumers’ tastes.

Statement of exclusive submission

This paper has not been submitted elsewhere in identical or similar form, nor will it be during the first three months after its submission to the Publisher.

Data availability

The authors do not have permission to share data.

Acknowledgements

This research was supported by the National Science Foundation Award No. SES-2031841. WB extends his gratitude for the support of the Kosciuszko Foundation, the American Centre of Polish Culture.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2023.116093.

Appendix A

Table A1
Detailed information for the sample, regarding race, education level and worldview-related variables.

Race:	White	Black or African American	American Indian and Alaska Native	Asian	Native Hawaiian and Other Pacific Islander	Multiracial
	0.747	0.158	0.011	0.035	0.003	0.026
Education:	**Lower than bachelor degree**	**Bachelor or associate degree**	**Master degree**	**Higher than master degree**		
	0.404	0.368	0.176	0.051		
Generally speaking, do you usually think of yourself as:	**Very conservative**	**Somewhat conservative**	**Moderate**	**Somewhat liberal**	**Very liberal**	
	0.182	0.167	0.358	0.149	0.144	
How important is religion to you?	**Not at all important**	**Slightly important**	**Fairly important**	**Very important**	**Extremely important**	
	0.211	0.144	0.172	0.215	0.257	
In general, would you say your views in most political matters are:	**Republican**	**Democrat**	**Independent**	**Other**	**No preference**	
	0.28	0.451	0.218	0.011	0.04	

Table A2
Latent class model estimates (in preference-space).

	Class 1		Class 2		Class 3		Class 4	
				Model coefficients				
Opt out ASC	−0.047		−4.235	***	4.685	***	0.697	**
	[0.364]		[0.242]		[0.893]		[0.275]	
Opt out ASC (std.dev.)	1.511	***	2.383	***	5.486	***	2.169	***
	[0.086]		[0.174]		[0.368]		[0.135]	
Out-of-pocket cost (in 10$)	−0.076	***	−0.009	***	−0.441	***	−0.062	***
	[0.006]		[0.002]		[0.027]		[0.003]	
				Protection				
% of protected vaccinated people (divided by 10)	0.257	***	0.025	***	0.13	***	0.467	***
	[0.021]		[0.009]		[0.043]		[0.020]	
= 1 if minimal protection satisfied	0.204	**	0.006		0.014		0.05	

(*continued on next page*)

Table A2 (*continued*)

	Class 1		Class 2		Class 3		Class 4	
	Model coefficients							
	[0.101]		[0.049]		[0.218]		[0.069]	
Number of months of protection	0.09	***	0.005		0.056	***	0.093	***
	[0.010]		[0.004]		[0.020]		[0.006]	
	Side-effects							
How many people develop mild side effects (1 in 10)	−0.091	***	−0.023	***	−0.101	**	−0.071	***
	[0.022]		[0.008]		[0.046]		[0.013]	
How many people develop severe side effects (1 in 1 mln)	−0.067	***	−0.006	**	−0.008		−0.067	***
	[0.007]		[0.003]		[0.014]		[0.004]	
	Origin							
Country of origin: UK (base level: USA)	−1.022	***	−0.178	***	−1.303	***	0.274	**
	[0.121]		[0.054]		[0.329]		[0.116]	
Country of origin: Germany (base level: USA)	−1.107	***	−0.323	***	−0.494	**	−0.228	**
	[0.101]		[0.049]		[0.226]		[0.101]	
Country of origin: China (base level: USA)	−4.88	***	−0.502	***	−0.961	***	−0.353	***
	[0.149]		[0.034]		[0.213]		[0.049]	
Country of origin: Russia (base level: USA)	−4.943	***	−0.26	***	−0.655	***	−0.553	***
	[0.145]		[0.061]		[0.230]		[0.072]	
	Recommending institution							
Recommended by: media (base level: doctor)	−1.264	***	−0.259	***	−0.498	*	−0.923	***
	[0.140]		[0.094]		[0.275]		[0.145]	
Recommended by: CDC (base level is doctor)	0.081		−0.137	*	0.597	**	0.115	
	[0.136]		[0.074]		[0.304]		[0.122]	
Recommended by: WHO (base level is doctor)	−0.356	***	−0.063		−0.303		−0.149	
	[0.131]		[0.084]		[0.275]		[0.137]	
	Other							
Number of days for antibodies to develop	−0.021	***	−0.003		−0.025	**	−0.015	***
	[0.006]		[0.002]		[0.011]		[0.003]	
Number of months since developed	0.051	*	0.02	*	0.033		0.084	***
	[0.027]		[0.011]		[0.056]		[0.029]	
= 1 if booster available	0.652	***	0.066	*	0.244		0.359	***
	[0.100]		[0.039]		[0.214]		[0.044]	
Number of doses	0.067		−0.095		0.115		−0.191	**
	[0.126]		[0.064]		[0.333]		[0.075]	
Days between doses (divided by 10)	0.001		0.034		0.033		0.051	
	[0.075]		[0.042]		[0.190]		[0.048]	

Table A3

Estimates of coefficients of latent factors in the measurement equations. Each indicator variable is treated as an ordered probit model.

	LV 1 (Germ aversion)		LV 2 (Perceived uninfectability)		LV 3 (Poor health)	
It really bothers me when people sneeze without covering their mouths	−0.679	***				
	[0.029]					
I am comfortable sharing a water bottle with a friend	0.912	***				
	[0.028]					
I don't like to write with a pencil someone else has obviously chewed on	−0.538	***				
	[0.023]					
I prefer to wash my hands pretty soon after shaking someone's hand	−0.521	***				
	[0.023]					
My hands do not feel dirty after touching money	0.836	***				
	[0.025]					
It does not make me anxious to be around sick people	0.771	***				
	[0.023]					
If an illness is going around, I will get it			1.289	***		
			[0.024]			
I have a history of susceptibility to infectious diseases			1.615	***		
			[0.032]			
In general, I am very susceptible to colds, flu, and other infectious diseases			1.719	***		
			[0.034]			
I am more likely than the people around me to catch an infectious disease			1.502	***		
			[0.028]			
How would you describe your current health status?					3.309	***
					[0.322]	
How would you describe your health status in 2019?					2.072	***
					[0.091]	
Did the respondent answer: "Do you have an underlying health condition that puts you more at risk if contracting COVID-19?"					−0.189	***
					[0.040]	
Do you have an underlying health condition that puts you more at risk if contracting COVID-19?					0.343	***
					[0.023]	

References

Adriaan, H., Jacco, D., 2009. Nonresponse in the recruitment of an internet panel based on probability sampling. Survey Research Methods 3.
Agranov, M., Elliott, M., Ortoleva, P., 2021. The importance of Social Norms against Strategic Effects: the case of COVID-19 vaccine uptake. Econ. Lett. 206, 109979.
Amani, A., Selenguai, H.K., Puepi, Y.D., Dairou, I.A., Kenmoe, S., Nouko, A., Cheuyem, F. Z.L., Koki, P.O.N., Mbacham, W., Ongolo-Zogo, P., 2022. A Systematic Review on COVID-19 Vaccine Preferences Using Discrete Choice Experiments. medRxiv.
Anggraeni, W., Aristiani, L., 2016. Using Google Trend data in forecasting number of dengue fever cases with ARIMAX method case study: surabaya, Indonesia. In: 2016 International Conference on Information & Communication Technology and Systems (ICTS). IEEE, pp. 114–118.
Bansal, P., Raj, A., Shukla, D.M., Sunder, N., 2022. COVID-19 vaccine preferences in India. Vaccine 40, 2242–2246.
Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., Polydoropoulou, A., 2002. Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges. In: Integration of Choice and Latent Variable Models, pp. 431–470.
Borriello, A., Master, D., Pellegrini, A., Rose, J.M., 2021. Preferences for a COVID-19 vaccine in Australia. Vaccine 39, 473–479.
Brewer, N.T., 2021. What works to increase vaccination uptake. Academic Pediatrics 21, S9–S16.
Budziński, W., Czajkowski, M., 2022. Endogeneity and measurement bias of the indicator variables in hybrid choice models: a Monte Carlo investigation. Environ. Resour. Econ.
Callaway, E., 2020. Russia's fast-track coronavirus vaccine draws outrage over safety. Nature 584, 334–336.
Chambon, M., Kammeraad, W.G., van Harreveld, F., Dalege, J., Elberse, J.E., van der Maas, H.L., 2022. Understanding change in COVID-19 vaccination intention with network analysis of longitudinal data from Dutch adults. npj Vaccines 7, 114.
Chang, Y.-F., Yin, S.-Y., 2021. Forecasting the Taipei House Prices Index: an Application of Factor Model with Google Trend Index. Available at: SSRN 3877924.
Chevallier, C., Hacquin, A.-S., Mercier, H., 2021. COVID-19 vaccine hesitancy: shortening the last mile. Trends Cognit. Sci. 25, 331–333.
Chorus, C.G., Kroesen, M., 2014. On the (im-)possibility of deriving transport policy implications from hybrid choice models. Transport Pol. 36, 217–222.
Chu, H., Liu, S., 2021. Light at the end of the tunnel: influence of vaccine availability and vaccination intention on people's consideration of the COVID-19 vaccine. Soc. Sci. Med. 286, 114315.
Czajkowski, M., Budziński, W., 2019. Simulation error in maximum likelihood estimation of discrete choice models. Journal of Choice Modelling 31, 73–85.
Daly, M., Robinson, E., 2021. Willingness to vaccinate against COVID-19 in the US: representative longitudinal evidence from April to October 2020. Am. J. Prev. Med. 60, 766–773.
Das, S., Samanta, S., Banerjee, J., Pal, A., Giri, B., Kar, S.S., Dash, S.K., 2022. Is Omicron the end of pandemic or start of a new innings? Trav. Med. Infect. Dis., 102332
Daziano, R.A., 2022. A choice experiment assessment of stated early response to COVID-19 vaccines in the USA. Health Economics Review 12, 1–16.
Díaz, A., Soriano, J.F., Beleña, Á., 2016. Perceived vulnerability to disease questionnaire: factor structure, psychometric properties and gender differences. Pers. Indiv. Differ. 101, 42–49.
Diaz, P., Reddy, P., Ramasahayam, R., Kuchakulla, M., Ramasamy, R., 2021. COVID-19 vaccine hesitancy linked to increased internet search queries for side effects on fertility potential in the initial rollout phase following Emergency Use Authorization. Andrologia 53, e14156.
Dong, D., Xu, R.H., Wong, E.L.y., Hung, C.T., Feng, D., Feng, Z., Yeoh, E.k., Wong, S.Y.s., 2020. Public preference for COVID-19 vaccines in China: a discrete choice experiment. Health Expect. 23, 1543–1578.
Duncan, L.A., Schaller, M., Park, J.H., 2009. Perceived vulnerability to disease: development and validation of a 15-item self-report instrument. Pers. Indiv. Differ. 47, 541–546.
Fishman, J., Salmon, M.K., Scheitrum, D., Schaefer, K.A., Robertson, C.T., 2022. Comparative effectiveness of mandates and financial policies targeting COVID-19 vaccine hesitancy: a randomized, controlled survey experiment. Vaccine 40, 7451–7459.
Forni, G., Mantovani, A., 2021. COVID-19 vaccines: where we stand and challenges ahead. Cell Death Differ. 28, 626–639.
Fridman, A., Gershon, R., Gneezy, A., 2021. COVID-19 and vaccine hesitancy: a longitudinal study. PLoS One 16, e0250123.
Halabowski, D., Rzymski, P., 2021. Taking a lesson from the COVID-19 pandemic: preventing the future outbreaks of viral zoonoses through a multi-faceted approach. Sci. Total Environ. 757, 143723.
Hershey, J.C., Asch, D.A., Thumasathit, T., Meszaros, J., Waters, V.V., 1994. The roles of altruism, free riding, and bandwagoning in vaccination decisions. Organ. Behav. Hum. Decis. Process. 59, 177–187.
Hess, S., Spitz, G., Bradley, M., Coogan, M., 2018. Analysis of mode choice for intercity travel: application of a hybrid choice model to two distinct US corridors. Transport. Res. Pol. Pract. 116, 547–567.
Janz, N.K., Becker, M.H., 1984. The health belief model: a decade later. Health Educ. Q. 11, 1–47.
Johnson, N.F., Velásquez, N., Restrepo, N.J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S., Lupu, Y., 2020. The online competition between pro-and anti-vaccination views. Nature 582, 230–233.
Kobayashi, Y., Howell, C., Heinrich, T., 2021. Vaccine hesitancy, state bias, and Covid-19: evidence from a survey experiment using Phase-3 results announcement by BioNTech and Pfizer. Soc. Sci. Med. 282, 114115.
Kreps, S., Kriner, D., 2021. Factors influencing Covid-19 vaccine acceptance across subgroups in the United States: evidence from a conjoint experiment. Vaccine 39, 3250–3258.
Lau, K., Miraldo, M., Galizzi, M.M., Hauck, K., 2019. Social norms and free-riding in influenza vaccine decisions in the UK: an online experiment. Lancet 394, S65.
Leng, A., Maitland, E., Wang, S., Nicholas, S., Liu, R., Wang, J., 2021. Individual preferences for COVID-19 vaccination in China. Vaccine 39, 247–254.
Looi, M.-K., 2022. The BMJ Interview: WHO Chief Scientist Optimistic for a Pan-Coronavirus Vaccine in Two Years, vol. 377.
Makhanova, A., Shepherd, M.A., 2020. Behavioral immune system linked to responses to the threat of COVID-19. Pers. Indiv. Differ. 167, 110221.
McPhedran, R., Toombs, B., 2021. Efficacy or delivery? An online Discrete Choice Experiment to explore preferences for COVID-19 vaccines in the UK. Econ. Lett. 200, 109747.
Motta, M., 2021. Can a COVID-19 vaccine live up to Americans' expectations? A conjoint analysis of how vaccine characteristics influence vaccination intentions. Soc. Sci. Med. 272, 113642.
Mouter, N., de Ruijter, A., de Wit, G.A., Lambooij, M.S., van Wijhe, M., van Exel, J., Kessels, R., 2022. "Please, you go first!" preferences for a COVID-19 vaccine among adults in The Netherlands. Soc. Sci. Med. 292, 114626.
Moutinho, S., Wadman, M., 2021. Brazil and Russia Face off over Vaccine Contamination Charge. American Association for the Advancement of Science.
Pullan, S., Dey, M., 2021. Vaccine hesitancy and anti-vaccination in the time of COVID-19: a Google Trends analysis. Vaccine 39, 1877–1881.
Raciborski, F., Jankowski, M., Gujski, M., Pinkas, J., Samel-Kowalik, P., 2021. Changes in attitudes towards the COVID-19 vaccine and the willingness to get vaccinated among adults in Poland: analysis of serial, cross-sectional, representative surveys. Vaccines 9, 832. January–April 2021.
Rosenstock, I.M., 1974. The health belief model and preventive health behavior. Health Educ. Monogr. 2, 354–386.
Sadaf, A., Richards, J.L., Glanz, J., Salmon, D.A., Omer, S.B., 2013. A systematic review of interventions for reducing parental vaccine refusal and vaccine hesitancy. Vaccine 31, 4293–4304.
Sanders, J.G., Spruijt, P., van Dijk, M., Elberse, J., Lambooij, M.S., Kroese, F.M., de Bruin, M., 2021. Understanding a national increase in COVID-19 vaccination intention, The Netherlands, November 2020–March 2021. Euro Surveill. 26, 2100792.
Sarrias, M., 2020. Individual-specific posterior distributions from Mixed Logit models: properties, limitations and diagnostic checks. Journal of Choice Modelling 36, 100224.
Schwarzinger, M., Watson, V., Arwidson, P., Alla, F., Luchini, S., 2021. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 6, e210–e221.
Sinclair, S., Agerström, J., 2023. Do social norms influence young people's willingness to take the COVID-19 vaccine? Health Commun. 38, 152–159.
Smith, M., 2021. How Much Difference Does it Make to People where a COVID Vaccine Was Developed?.
Szolnoki, G., Hoffmann, D., 2013. Online, face-to-face and telephone surveys—comparing different sampling methods in wine consumer research. Wine Economics and Policy 2, 57–66.
Taylor, S., Landry, C.A., Paluszek, M.M., Groenewoud, R., Rachor, G.S., Asmundson, G.J., 2020. A proactive approach for managing COVID-19: the importance of understanding the motivational roots of vaccination hesitancy for SARS-CoV2. Front. Psychol. 11, 575950.
Teh, H.S., Woon, Y.L., Leong, C.T., Hing, N.Y.L., Mien, T.Y.S., Roope, L.S., Clarke, P.M., Lim, L.-L., Buckell, J., 2022. Malaysian public preferences and decision making for COVID-19 vaccination: a discrete choice experiment. The Lancet Regional Health-Western Pacific 27, 100534.
Train, K., Weeks, M., 2005. Discrete choice models in preference space and willingness-to-pay space. In: Scarpa, R., Alberini, A. (Eds.), Applications of Simulation Methods in Environmental and Resource Economics. Springer Netherlands, pp. 1–16.