Stated choice analysis of preferences for COVID-19 vaccines using the Choquet integral

Rico Krueger [*] Ricardo A. Daziano [†]

29 November 2021

Report TRANSP-OR 211129
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

[*]École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, {rico.krueger}@epfl.ch
[†]School of Civil and Environmental Engineering, Cornell University, United States, {daziano}@cornell.edu

Abstract

We investigate preferences for COVID-19 vaccines using data from a stated choice survey conducted in the US in March 2021. To analyse the data, we embed the Choquet integral, a flexible aggregation operator for capturing attribute interactions under monotonicity constraints, into a mixed logit model. We find that effectiveness is the most important vaccine attribute, followed by risk of severe side effects and protection period. The attribute interactions reveal that non-pecuniary vaccine attributes are synergistic. Out-of-pocket costs are independent of effectiveness, incubation period, and mild side effects but exhibit moderate synergistic interactions with other attributes. Vaccine adoption is significantly more likely among individuals who identify as male, have obtained a bachelor's degree or a higher level of education, have a high household income, support the democratic party, had COVID-19, got vaccinated against the flu in winter 2020/21, and have an underlying health condition.

Keywords: COVID-19, vaccines, patient preferences, stated choice, discrete choice.

1 Introduction

The COVID-19 pandemic continues to pose significant risks to public health. Worldwide, more than 250 million COVID-19 cases have been reported, and more than 5 million deaths have been associated with the disease as of 18 November 2021 (Dong et al., 2020b). In the United States (US), more than 3 million hospital admissions between 1 August 2020 and 15 November 2021 were linked to COVID-19 (Centers for Disease Control and Prevention, 2020). Initial non-pharmaceutical interventions such as lockdowns, social distancing, and work-from-home orders to slow the spread of the disease have led to substantial social and economic disruptions.

Pharmaceutical interventions in the form of vaccines are now viewed as the most effective way out of the pandemic. Several COVID-19 vaccines have been developed and authorised for use at a rapid pace (Basta et al., 2020, Wouters et al., 2021). COVID-19 vaccines are safe and effectively prevent symptomatic and asymptomatic infections with SARS-CoV-2, the virus that causes the COVID-19 disease (e.g. Baden et al., 2021, Polack et al., 2020). Vaccinated individuals are significantly less likely to develop severe symptoms that require hospitalisation and to die from the disease (Tenforde et al., 2021). Thus, mass immunisations with COVID-19 vaccines are crucial for ending the pandemic and the associated public health crisis. To that end, mass vaccination campaigns have been launched in many countries (Mathieu et al., 2021). The success of these campaigns depends critically on the decisions of individual members of society to get vaccinated. Aside from availability factors, the individual decision to get vaccinated is likely influenced by the attributes of the available vaccines and person-specific characteristics.

Understanding how individual preferences influence the decision to get vaccinated against COVID-19 is essential for supporting a widespread adoption of COVID-19 vaccines both during initial roll-outs and booster campaigns. First, insights into preferences for COVID-19 vaccines can inform targeted information campaigns that emphasise the perceived benefits of the available vaccines in communications with the target group. Second, information about preferences for COVID-19 vaccines can support decision-makers in public procurement processes in selecting vaccines that are comparatively more likely to be adopted by the target group. Third, insights into preferences for COVID-19 vaccines can guide pharmaceutical companies in developing COVID-19 vaccines with features that maximise the likelihood of adoption by a target group.

Stated choice methods constitute a powerful framework for eliciting and analysing preferences for vaccines. In that vein, several studies have employed stated choice methods to investigate preferences for COVID-19 vaccines (Borriello et al., 2021, Dong et al., 2020a, Eshun-Wilson et al., 2021, Leng et al., 2021, McPhedran and Toombs, 2021). The discrete choice experiments in these studies elicit preferences for various vaccine attributes such as the effectiveness, the length of the protection period, the risk of developing side effects, the number of required doses, and the out-of-pocket costs as well as for other attributes such as the place of administration and social influence. For the analysis of the stated choice data, the studies employ multinomial, mixed and latent class logit models in which the systematic utility is specified as linear-in-parameters.

Discrete choice models with linear-in-parameters utility specifications are limited in their ability to explain preferences for alternatives described by multidimensional attribute vectors. This is because a linear-in-parameters utility specification makes it difficult to repre-

sent attribute interdependencies while also maintaining interpretability and monotonicity (Dubey et al., 2021, Tehrani et al., 2012). Interpretability of preferences is a key desideratum in stated choice analysis. Monotonicity is a behaviourally meaningful constraint in discrete choice analysis. Monotonicity implies that all else being equal, an increase in the level of a desirable attribute does not lower the utility of an alternative, and vice versa, that a decrease in the level of an undesirable attribute does not increase the utility of an alternative.

It is easy to see why a linear-in-parameters utility specification is found wanting in these two regards. In its standard form, a linear-in-parameters utility specification corresponds to a weighted sum aggregation of the attributes of an alternative. The weighted sum aggregation is simple and easy to interpret, mainly because the marginal effect of an attribute on the utility is given by its estimated weight in the utility. Yet, the weighted sum aggregation lacks expressiveness due to its inability to capture dependencies between attributes. To overcome this limitation, analysts may include second- and higher-order interaction effects in a linear-in-parameters utility specification. However, utility specifications with interaction effects are inherently difficult to interpret, since the marginal effect of an attribute depends on the main effect and all interaction effects that include the relevant attribute (Tehrani et al., 2012). For the same reason, utility specifications with interaction effects may also violate monotonicity constraints (Tehrani et al., 2012).

In this paper, we aim to advance the understanding of preferences for COVID-19 vaccines by formulating and applying a discrete choice model in which a component of the systematic utility is represented using the discrete Choquet integral. The discrete Choquet integral is a flexible aggregation operator for interacting attributes under monotonicity constraints. It also provides a quantification of the relative importance of individual attributes and the degree of interaction of attributes (i.e. the Choquet integral identifies to what extent two or more attributes are independent, synergistic or redundant). We embed the Choquet integral into a normal error components mixed logit formulation. The resulting model is a useful and behaviourally meaningful decision support tool. First, the model is easy to interpret because the Choquet integral quantifies both attribute importance and the degree of interaction of attributes. Second, the Choquet integral ensures monotonicity. Third, the model preserves the usual benefits of mixed logit. The normal error components allow us to capture unobserved agent effects and define meaningful nesting structures that imply realistic substitution patterns.

We apply the proposed model to data from a nationwide stated choice survey (N=1421), which we conducted in the US in March 2021. The discrete choice experiment in the survey included two hypothetical COVID-19 vaccines and an opt-out alternative. The vaccines were described by nine attributes, namely the out-of-pocket costs, the effectiveness, the protection period, the incubation period, the risk of severe side effects, the risk of mild side effects, and the number of required doses, whether the vaccine has a booster against variants, and the origin of the vaccine. The proposed discrete choice model with a Choquet integral representation of the systematic utility allows us to quantify the importance of the attributes and characterise the interaction of the attributes. In our model specification, we also include an alternative-specific constant (ASC) for the opt-out alternative. Interactions of this ASC with socio-demographic attributes offer insights into the person-specific attributes that drive vaccine non-adoption.

We organise the remainder of this paper as follows: In the following section, we present

a review of the pertinent literature. In Section 3, we describe the stated choice data on preferences for COVID-19 vaccines. In Section 4, we introduce the modelling approach. In Section 5, we present the results. Finally, we conclude in Section 6.

2 Literature review

2.1 Stated choice analysis of preferences for COVID-19 vaccines

An ever growing number of studies have investigated preferences for COVID-19 vaccines using stated choice methods.

Borriello et al. (2021) conducted a stated choice survey in Australia in March 2020 and analysed the collected data using a latent class choice model. The authors find that preferences for vaccine effectiveness, price, mild side effects as well as the mode and location of administration are heterogeneous, whereas preferences for severe side effects and immediacy (i.e. the expected point in time when the vaccine becomes available) are homogeneous.

Eshun-Wilson et al. (2021) carried out a stated choice survey in the US in March 2021 and analysed the collected data using mixed and latent class logit models. The authors' mixed logit analysis reveals that on average, respondents prefer one vaccine dose as opposed to two and prefer to be vaccinated a single time rather than annually. The authors' latent class analysis identifies four preference segments with the first and largest segment valuing vaccine features (i.e. number of required does and required vaccination frequency) the most, a second segment being primarily concerned about vaccine administration aspects (i.e. wait time and administration at mass site, health centre or at home), a third segment valuing enforcement and social proof of vaccine safety, and a fourth segment that is indifferent to vaccine and administration features and is opposed to enforcement.

McPhedran and Toombs (2021) conducted a stated choice survey in the United Kingdom (UK) in August and September 2020. The authors' multinomial (conditional) logit analysis of the collect data reveals that respondents perceive vaccine effectiveness as the most important attribute and that the sensitivity for high vaccine effectiveness is comparatively larger of individuals aged 55 years old or older.

Dong et al. (2020a) collected data via a stated choice survey in China in June and July 2020. The authors' mixed logit analysis finds respondents value vaccines that are highly effective, offer a long protection period, have a low risk of side effects, and are manufactured overseas.

Leng et al. (2021) also conducted a stated choice survey in China in 2020 and analysed the collected data using multinomial logit and latent class logit models. The authors find that high vaccine effectiveness, a low risk of side effects and social influence (i.e. the proportion of vaccinated acquaintances) are most important to respondents.

Prior to the COVID-19 pandemic, Determann et al. (2014) conducted a discrete choice experiment to investigate preferences for vaccine attributes in a hypothetical pandemic outbreak. The choice experiment considered several vaccine attributes, including effectiveness, safety, advice, media coverage and out-of-pocket costs. The hypothetical pandemic outbreak was described by two scenario attributes, namely the disease susceptibility and the disease severity. The authors' latent class logit analysis detects substantial preference heterogeneity with respect to the considered attributes. Vaccine effectiveness,

out-of-pocket costs and the nature of the body that advises the vaccine are found to be the most relevant attributes.
Furthermore, using data from a stated choice survey, de Bekker-Grob et al. (2018) investigate preferences for attributes of influenza vaccines. The considered attributes include vaccine effectiveness, risk of mild side effects, risk of severe side effects, the incubation period and the protection period. The authors analysis finds that both vaccine attributes and person-specific attributes influence the decision to get vaccinated.

2.2 Discrete choice models and the Choquet integral

The Choquet integral (Choquet, 1954) has found widespread application in operations research in the context of multi-criteria decision-making (Grabisch, 1996, Grabisch and Labreuche, 2010). Yet, the Choquet integral has received limited attention in discrete choice analysis. Aggarwal (2020) incorporate the Choquet integral into a multinomial logit model. Similarly, Tehrani et al. (2012) formulate a logistic regression model based on the Choquet integral. Both of these models succumb to the well known weaknesses of logit (i.e. the inability of logit to capture realistic substitution patterns and correlation in unobserved factors over time). Dubey et al. (2021) embed the Choquet integral into a multinomial probit model to accommodate unrestricted substitution patterns. However, the resulting model is computationally expensive, since the authors employ the GHK simulator to approximate multinomial probit choice probabilities. The computational burden of this model would increase even further, if an analyst wished to accommodate agent-specific effects using error components in a mixed multinomial probit formulation. This is because the model would require two layers of simulation, one for the agent-specific effects and another one for the choice probabilities. In this paper, we thus embed the Choquet integral into a normal error components mixed logit formulation to accommodate unobserved agent effects and realistic substitution patterns in a computationally efficient manner. Since the choice probabilities of the logit kernel are available in closed-form, only one layer of simulation is required during model estimation.

3 Data

We conducted a nationwide stated choice survey in the US from 4 to 10 March 2021 to investigate preferences for COVID-19 vaccines. The survey included a discrete choice experiment which involved a choice between two hypothetical COVID-19 vaccines and an opt-out alternative. In total, we collected 1,421 valid responses. Each respondent completed seven choice scenarios.
The vaccines in the discrete choice experiment were described by nine attributes, namely the out-of-pocket costs, the effectiveness, the protection period, the incubation period, the risk of severe side effects, the risk of mild side effects, the number of required doses, whether the vaccine has a booster against variants, and the origin of the vaccine. An example of choice task is shown in Figure 1. Table 1 enumerates the levels of the considered attributes. The attributes were selected based on a review of the literature and an online focus group. Five of the nine attributes, namely effectiveness, protection period, incubation period, risk of severe side effects, risk of mild side effects are taken from de Bekker-Grob et al. (2018). The attribute out-of-pocket costs was included to facilitate eventual welfare

calculations. We also included the attribute number of required doses with levels one and two, since at the time of survey design, vaccines that were approved or awaiting approval required one or two doses.
The survey also collected information about the respondents' socio-demographic and health-related characteristics. Table 2 describes the sample in terms of these characteristics.

Figure 1: Example of a choice task

Attribute	Levels
Out-of-pocket cost [USD]	(0, 50, 100, 175)
Effectiveness [%]	(60, 80, 95)
Protection period [months]	(6, 12)
Incubation period [days]	(7, 14, 21)
Risk of severe side effects [out of 10^6]	(1, 10, 100)
Risk of mild side effects [out of 10]	(1, 3, 5)
No. required doses	(1, 2)
Booster against variants	(0, 1)
Origin	(China, Russia, USA)

Table 1: Attributes and levels

Variable	Sample proportion [%]
Gender: male	49.8
Cohort: Generation Z	4.2
Cohort: Millenial	28.9
Cohort: Generation X	22.1
Cohort: Baby Boomer	38.2
Cohort: older than Baby Boomer	6.6
Race: Asian or Asian-American	3.0
Race: Black or African-American	16.1
Ethnicity: Hispanic	15.1
Education: BSc	28.5
Education: PostGrad	29.1
Full-time worker	48.3
Household income	
less than $40,000	7.5
$40,000 to $74,999	35.7
$75,000 to $99,999	19.6
$100,000 to $124,999	10.8
$125,000 to $149,999	10.0
$150,000 to $199,999	9.2
$200,000 or more	7.2
Political views	
Democrat	49.9
Republican	26.0
independent or other	24.1
Has tested positive for COVID-19	17.4
Got vaccinated against flu in winter 2020/21	50.4
Has underlying condition	41.9
Division	
Pacific	16.7
Mountain	6.8
North West Central	5.4
West South Central	8.9
East North Central	12.6
East South Central	4.1
Middle Atlantic	18.9
South Atlantic	22.7
New England	3.9

Table 2: Sample description (N=1421)

4 Modelling approach

4.1 Set-up

We consider a standard set-up for a random utility model. We analyse a sample of N individuals indexed by $n = 1, \ldots, N$. Every individual is observed to choose an alternative y_{nt} from the set $\mathcal{M} = \{1, \ldots, J\}$ in T choice situations indexed by $t = 1, \ldots, T$. Each alternative is described by a set $X_{ntj} = \{x_{ntj1}, \ldots, x_{ntjK}\}$ of K attributes. Random utility theory (McFadden et al., 1973) posits that an individual selects the alternative with the highest random utility, i.e.

$$y_{nt} = j \text{ iff } U_{ntj} > U_{ntj'} \; \forall \, j' \in \mathcal{M} \setminus j, \tag{1}$$

where

$$U_{ntj} = V_{ntj}(X_{ntj}; \theta) + \varepsilon_{ntj} \tag{2}$$

is the random utility of alternative $j \in \mathcal{M}$. U_{ntj} is composed of a deterministic component $V_{ntj}(X_{ntj}; \theta)$ and a random component ε_{ntj}. The deterministic utility $V_{ntj}(X_{ntj}; \theta)$ is a score capturing the attractiveness of alternative j as a function of the attributes X_{ntj} and an unknown parameter vector θ. In general, $V_{ntj}(X_{ntj}; \theta)$ is calculated using an operator $\mathcal{H}$ that aggregates the attribute and parameter vectors into a scalar. Thus, we have

$$V_{ntj}(X_{ntj}; \theta) = \mathcal{H}(X_{ntj}; \theta). \tag{3}$$

The most common aggregation operator is the weighted sum W, i.e.

$$V_{ntj}(X_{ntj}; \theta = \beta) = W_\beta(X_{ntj}) = \sum_{k=1}^{K} \beta_k x_{ntjk} \tag{4}$$

with $\beta = (\beta_1, \ldots, \beta_K)$.

4.2 Choquet integral

In what follows, we outline the key features of the Choquet integral. For detailed discussions of the properties of the Choquet integral, the reader is directed to the literature (Grabisch, 1996, Grabisch et al., 2008, Grabisch and Labreuche, 2010, Marichal, 2002, Tehrani et al., 2012). For notational simplicity, we omit the individual- and choice situation-specific subscripts n and t in the subsequent exposition.
A fuzzy measure on a set of attributes $X = \{x_1, \ldots, x_K\}$ of cardinality K is a set function $\mu : 2^K \to [0, 1]$, which satisfies the following two conditions:

$$\mu(\emptyset) = 0, \quad \mu(X) = 1, \tag{5}$$

$$\text{for any } S, T \subseteq X, \quad S \subseteq T \Rightarrow \mu(S) \leq \mu(T). \tag{6}$$

For any $S \subseteq X$, $\mu(S)$ represents the weight or importance of the coalition S of attributes in X.
The Choquet integral C is defined as

$$C_\mu(X) = \sum_{k=1}^{K} x_{(k)} \left[\mu\left(A_{(k)}\right) - \mu\left(A_{(k+1)}\right) \right], \tag{7}$$

where $(\cdot)$ is a permutation operator such that $x_{(1)} \leq \ldots \leq x_{(K)}$. Furthermore, $A_{(k)} = \{x_k, \ldots, x_K\}$ and $A_{(k+1)} = \emptyset$. For example, if $X = \{x_1, x_2, x_3\}$ and $x_3 \leq x_1 \leq x_2$, then

$$\begin{aligned} C_\mu(\{x_1, x_2, x_3\}, \mu) =& x_3 \left[\mu\left(\{x_3, x_1, x_2\}\right) - \mu\left(\{x_1, x_2\}\right)\right] + \\ & x_1 \left[\mu\left(\{x_1, x_2\}\right) - \mu\left(\{x_2\}\right)\right] + \\ & x_2 \mu\left(\{x_2\}\right). \end{aligned} \tag{8}$$

We let $\mathcal{F}_X$ denote the set of all fuzzy measures on X. Any fuzzy measure $\mu \in \mathcal{F}_X$ has an equivalent Möbius representation. The Möbius transform of a fuzzy measure $\mu \in \mathcal{F}_X$ is a set function $m_\mu : 2^K \to \mathbb{R}$. It is given by

$$m_\mu(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} \mu(T), \quad S \subseteq X. \tag{9}$$

The corresponding inverse transform is

$$\mu(S) = \sum_{T \subseteq S} m_\mu(T), \quad S \subseteq X. \tag{10}$$

For μ to be a valid fuzzy measure, m_μ must satisfy the following conditions:

$$m(\emptyset) = \emptyset, \quad \sum_{T \subseteq X} m(T) = 1, \tag{11}$$

$$\sum_{T \subseteq S | x_k \in T} m(T) \geq 0, \quad \forall\, S \subseteq X,\ \forall\, x_k \in S. \tag{12}$$

The Möbius transform simplifies the calculation of the Choquet integral. In terms of the Möbius representation m_μ, the Choquet integral can be equivalently expressed as:

$$C_{m_\mu}(X) = \sum_{T \subseteq X} m_\mu(T) \min_{i \in T} x_i. \tag{13}$$

The *Shapley importance index* $\phi_\mu(x_k)$ of attribute x_k on fuzzy measure μ measures the *relative importance* of attribute x_k. It is defined as

$$\phi_\mu(x_k) = \sum_{T \subseteq X \setminus x_k} \frac{(K - |T| - 1)!|T|!}{K!} \left[\mu(T \cup x_k) - \mu(T)\right]. \tag{14}$$

The Shapley importance index can also be calculated in terms of the Möbius transform m_μ:

$$\phi_{m_\mu}(x_k) = \sum_{T \subseteq X \setminus x_k} \frac{m_\mu(T \cup x_k)}{|T| + 1} \tag{15}$$

The Shapley importance index of x_k can be viewed as the average marginal contribution of x_k to all coalitions that exclude x_k. Shapley importance indices exhibit the properties $0 \leq \phi_{m_\mu}(x_k) \leq 1$ and $\sum_{k=1}^{K} \phi_{m_\mu}(x_k) = 1$. Thus, $\phi_{m_\mu}(x_k) < \frac{1}{K}$ implies that x_k is less important than the average, and $\phi_{m_\mu}(x_k) > \frac{1}{K}$ implies that x_k is more important than the average.

The *interaction index* characterises the *degree of interaction* of two attributes. The interaction index $\kappa_\mu(\{x_k, x_l\})$ of a set of two attributes $\{x_k, x_l\}$ with $k \neq l$ on fuzzy measure μ is

$$\kappa_\mu(\{x_k, x_l\}) = \sum_{T \subseteq X \setminus \{x_k, x_l\}} \frac{(K - |T| - 2)!|T|!}{(K-1)!} \left[\mu(T \cup \{x_k, x_l\}) - \mu(T \cup x_k) - \mu(T \cup x_l) + \mu(T)\right]. \tag{16}$$

The interaction index can also be calculated in terms of the Möbius transform m_μ:

$$\kappa_{m_\mu}(\{x_k, x_l\}) = \sum_{T \subseteq X \setminus \{x_k, x_l\}} \frac{m_\mu(T \cup \{x_k, x_l\})}{|T| + 1}. \tag{17}$$

The interaction index of a set of two attributes $\{x_k, x_l\}$ can be viewed as the average marginal interaction between x_k and x_l. The interaction index is contained within $[-1, 1]$. A positive interaction index implies that two attributes x_k and x_l are *synergistic*, i.e. improving x_k and x_l jointly gives strictly more than improving either x_k or x_l. A negative interaction index implies that two attributes x_k and x_l are *redundant*, i.e. it is not necessary to improve x_k and x_l jointly. An interaction index of zero implies that two attributes x_k and x_l are *independent*.
An application of the Choquet integral requires that the attributes are normalised such that 0 corresponds to the worst possible levels of an attribute and 1 corresponds to the best possible value of an attribute. For desirable attributes, i.e. attributes for which more is better, the required normalisation is

$$x_{ntjk} = \frac{\tilde{x}_{ntjk} - \min \tilde{x}_k}{\max \tilde{x}_k - \min \tilde{x}_k}, \quad \forall\, n, t, j, k, \tag{18}$$

where $\tilde{x}_{ntjk}$ denotes the raw, unnormalised attribute. $\max \tilde{x}_k$ and $\min \tilde{x}_k$ denote the maximum and minimum of attribute k in its unnormalised form. For undesirable attributes, i.e. attributes for which less is better, the required normalisation is

$$x_{ntjk} = \frac{\max \tilde{x}_k - \tilde{x}_{ntjk}}{\max \tilde{x}_k - \min \tilde{x}_k}, \quad \forall\, n, t, j, k \tag{19}$$

As a consequence of the normalisation and the constraints imposed on the fuzzy measure, the output of the Choquet integral is constrained between 0 and 1.
In this paper, we exploit the Choquet integral as an aggregation operator for the specification of the deterministic utility in a random utility model. With

$$V_{ntj} = \lambda C_{m_\mu}(X_{ntj}), \tag{20}$$

we have

$$U_{ntj} = \lambda C_{m_\mu}(X_{ntj}) + \varepsilon_{ntj}, \tag{21}$$

whereby $\lambda > 0$ is an unknown precision parameter. We include λ because the output of the Choquet integral is constrained between 0 and 1 by design. λ sets the scale of the component of the deterministic utility that is represented by the Choquet integral with respect to other components of the deterministic utility and the error term.

4.3 Extensions

The deterministic utility can be constructed using a combination of different aggregation operators. Let $Z_{ntj} = \{z_{ntj1}, \ldots, z_{ntjK}\}$ denote a second set of L attributes. For example, we could specify

$$V_{ntj} = \lambda C_{m_\mu}(X_{ntj}) + W_\beta(Z_{ntj}) \tag{22}$$

such that

$$U_{ntj} = \lambda C_{m_\mu}(X_{ntj}) + \sum_{l=1}^{L} \beta_l z_{ntjl} + \varepsilon_{ntj}, \tag{23}$$

where m_μ and β are unknown parameters.
Furthermore, we can augment the stochastic utility component by adding normal error components. Suppose that there are B error components indexed by $b = 1, \ldots, B$. Then, the random utility becomes

$$U_{ntj} = \lambda C_{m_\mu}(X_{ntj}) + \sum_{l=1}^{L} \beta_l z_{ntjl} + \sum_{b=1}^{B} d_{jb}\sigma_b \xi_{nb} + \varepsilon_{ntj}, \tag{24}$$

where d_{jb} is one if error component b is associated with alternative j and zero otherwise. σ_b is the scale of error component b, and ξ_{nb} is a standard normal random variable.

4.4 Final model

The logit model is obtained under the assumption that the random error terms ε_{ntj} are independently and identically distributed according to Gumbel$(0, 1)$. Then, the probability that individual n selects alternative j in choice situation t conditional on ξ_n is

$$P(j|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n) = \frac{e^{V_{ntj}}}{\sum_{j' \in \mathcal{M}} e^{V_{ntj'}}}, \tag{25}$$

with

$$V_{ntj} = \lambda C_{m_\mu}(X_{ntj}) + \sum_{l=1}^{L} \beta_l z_{ntjl} + \sum_{b=1}^{B} d_{jb}\sigma_b \xi_{nb}. \tag{26}$$

Furthermore, The probability of observing the sequence of choices $y_n = (y_{n1}, \ldots, y_{nT})$ is

$$P(y_n|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n) = \prod_{t=1}^{T} P(y_{nt}|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n). \tag{27}$$

The unconditional probability is

$$P(y_n|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma) = \int P(y_n|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n) f(\xi_n) d\xi_n, \tag{28}$$

where $f(\xi_n)$ is the density of ξ_n. The integral in (28) is not analytically tractable. Therefore, it is approximated using R simulation draws denoted by ξ_{nr}:

$$P(y_n|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma) \approx \frac{1}{R} \sum_{r=1}^{R} P(y_n|X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_{nr}). \tag{29}$$

Consequently, the simulated log-likelihood is given by

$$\mathcal{L}(\theta) = \sum_{n=1}^{N} \ln \left(\frac{1}{R} \sum_{r=1}^{R} P(y_n | X_{ntj}, Z_{ntj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_{nr}) \right) \tag{30}$$

with $\theta = \{\lambda, m_\mu, \beta, \sigma\}$. The maximum simulated likelihood estimator of θ is then given by the solution to the following constrained optimisation problem:

$$\hat{\theta} = \arg\max_{\theta} \mathcal{L}(\theta) \tag{31}$$

$$\text{s.t.} \qquad m(\emptyset) = \emptyset, \quad \sum_{T \subseteq X} m(T) = 1 \tag{32}$$

$$\sum_{T \subseteq S \setminus x_k} m(T \cup x_k) \geq 0 \qquad \forall\, S \subseteq X,\ \forall\, x_k \in S \tag{33}$$

$$\lambda, \sigma \geq 0. \tag{34}$$

We implement the constrained maximum simulated likelihood estimation problem defined in (30)–(34) in Python. The unconditional choice probabilities are simulated using 200 Halton draws (Bhat, 2001) per individual. The constrained maximisation of the simulated likelihood is performed using the sequential least squares programming provided in Python's SciPy library (Virtanen et al., 2020). Standard errors are bootstrapped using 100 resamples.

5 Results

5.1 Model specifications

We estimate two normal error components mixed logit (NECML) models, namely

i) a NECML model in which all alternative-specific attributes are aggregated using the weighted sum operator (henceforth, *WS-NECML*), and

ii) a NECML model in which a component of the systematic utility of the vaccine alternatives is represented using the Choquet integral (henceforth, *Choquet-NECML*).

Both models include an alternative-specific constant (ASC) for the opt-out alternative. The ASC is interacted with socio-demographic attributes to provide insights into the person-specific characteristics that are associated with vaccine non-adoption. In both models, the utility for the opt-out alternative includes a normal error component with an estimable scale parameter. This error component introduces an agent effect and segregates the alternatives into two nests, one containing the two vaccine alternatives and another one containing the opt-out alternative. The considered normal error components specification of the model satisfies the non-trivial identification conditions of NECML models (see Walker et al., 2007).
In the model labelled Choquet-NECML, seven attributes, namely out-of-pocket costs, effectiveness, protection period, incubation period, risk of severe side effects, risk of mild

side effects and the number of required doses, are aggregated using the Choquet integral. The normalisation of the attributes (see equations (18) and (19)) requires us to identify which attributes are desirable and which attributes are undesirable. Consistent with common sense, we treat effectiveness and protection period as desirable attributes, while all remaining attributes are treated as undesirable. Since the Choquet integral only aggregates continuous attributes, the origin of the vaccine is included in a weighted sum aggregation. Specifically, we define a dummy variable indicating whether the vaccine is from the US. The attribute booster against variants is not included in both model specification, as the attribute was not found to have statistically significant influence on the utilities of the vaccine alternatives. In the model labelled WS-NECML, all alternative-specific attributes (with the exception of booster against variants) are aggregated using a weighted sum. A normalisation of attributes is not required to estimate WS-NECML.

5.2 Model fit

Table 3 compares the goodness-of-fit of the two models. We observe that Choquet-NECML provides a substantially better fit than WS-NECML, since the log-likelihood of Choquet-NECML is more than 100 units higher than the log-likelihood of WS-NECML. However, the improvement of fit appears to come at the cost added complexity. Whereas WS-NECML includes 20 unknown parameters, Choquet-NECML includes 140 unknown parameters.[1] Note that WS-NECML is nested within Choquet-NECML. This is because the simpler model can be obtained from the more complex one by setting all Möbius parameters that pertain to more than one attribute equal to zero. A likelihood ratio test leads us to reject the restrictions imposed by the simpler model and to select Choquet-NECML over WS-NECML ($\tilde{\chi}^2 = 216.452$, df $= 120$, $p > 0.999$).

	WS-NECML	Choquet-NECML
No. of parameters	20	140
Log-likelihood	-7851.8	-7743.6

Table 3: Model fit

5.3 Parameter estimates

In Table 4, we report the parameter estimates for the two models. We omit the estimates of the Möbius parameters, as we will interpret these parameters in terms of their Shapley importance and interaction representations (see Section 5.4). Our first observation is that the estimated signs of the parameters pertaining to alternative-specific attributes in WS-NECML are consistent with our normalisation assumptions. As expected, the estimates of parameters pertaining to "desirable" attributes (i.e. effectiveness and protection period) are positive in WS-NECML, while the estimates of the parameters pertaining to the remaining "undesirable" attributes are negative in WS-NECML. Both models indicate that

[1]The constrained maximum simulated likelihood estimator for the Choquet-NECML model as defined in (30)–(34) includes 141 parameters. However, due to the boundary constraint (32), one of the Möbius parameters is identified given the remaining Möbius parameters. Therefore, we consider 140 parameters as unknown in Choquet-NECML.

respondents have a positive preference for vaccines that originate from the US. The estimates of the parameters entering the utility of the opt-out alternative have the same signs in both models. Both models suggest that individuals who identify as male, have obtained a bachelor's degree or a higher level of education, have a high household income, support the democratic party, had COVID-19, got vaccinated against the flu in winter 2020/21, and have an underlying health condition are significantly less likely to opt out from vaccination. Also, higher income significantly increase the propensity to select the opt-out option. By contrast, individuals who belong to the Baby Boomer generation or an older generation, and are black or African-American are significantly more likely to opt out. The scale of the normal error component entering the utility of the opt-out alternative is estimated to be statistically significantly different from zero in both models. The estimate of the Choquet precision parameter λ does not carry a substantive meaning.

	WS-NECML				Choquet-NECML			
Variable	Est.	SE	z-val.	p-val.	Est.	SE	z-val.	p-val.
Out-of-pocket cost [USD] $\times 10^{-2}$	-0.316	0.023	-13.910	0.000				
Effectiveness [%]	0.026	0.001	24.878	0.000				
Protection period [months]	0.042	0.005	8.808	0.000				
Incubation period [days]	-0.012	0.003	-4.761	0.000				
Severe side effects [out of 10^6] $\times 10^{-2}$	-0.195	0.036	-5.482	0.000				
Mild side effects [out of 10]	-0.039	0.009	-4.281	0.000				
No. required doses	-0.086	0.037	-2.330	0.020				
Origin is USA	1.128	0.033	34.101	0.000	1.424	0.065	21.954	0.000
Opt-out	3.216	0.300	10.729	0.000	2.781	0.316	8.810	0.000
Opt-out $\times$ male	-0.910	0.193	-4.710	0.000	-0.938	0.190	-4.932	0.000
Opt-out $\times$ cohort is Baby Boomer or older	1.167	0.215	5.416	0.000	1.209	0.190	6.379	0.000
Opt-out $\times$ education bachelor	-0.832	0.222	-3.752	0.000	-0.861	0.225	-3.818	0.000
Opt-out $\times$ education postgraduate	-1.444	0.253	-5.705	0.000	-1.484	0.285	-5.197	0.000
Opt-out $\times$ household income [10k USD]	-0.053	0.019	-2.754	0.006	-0.054	0.019	-2.912	0.004
Opt-out $\times$ black or African-American	0.657	0.261	2.516	0.012	0.715	0.299	2.392	0.017
Opt-out $\times$ democrat	-1.538	0.193	-7.962	0.000	-1.599	0.228	-7.015	0.000
Opt-out $\times$ had COVID-19	-0.616	0.280	-2.198	0.028	-0.600	0.290	-2.067	0.039
Opt-out $\times$ received flu shot	-1.318	0.201	-6.575	0.000	-1.371	0.207	-6.628	0.000
Opt-out $\times$ has underlying condition	-0.361	0.195	-1.857	0.063	-0.388	0.209	-1.857	0.063
Opt-out std. dev.	2.808	0.111	25.375	0.000	2.948	0.124	23.768	0.000
Choquet precision λ					7.513	1.210	6.209	0.000

Table 4: Parameter estimates

5.4 Shapley importance and interaction indices

Figures 2 and 3 visualise the estimates of the Shapley importance and the interaction indices, respectively. Table 5 provides a more detailed tabulation of the estimates of the interaction indices.

Figure 2 shows the relative importance of the attributes. The dashed vertical line in the plot indicates average importance. The error bars represent the 95% confidence intervals. Effectiveness is the most important attribute, followed by severe side effects, and protection period. Mild side effects is the least important attribute, followed by out-of-pocket costs, and incubation period. Effectiveness and severe side effects are significantly more important than the average, whereas mild side effects and incubation period are significantly less important than the average. These findings suggest that improving the availability of highly effective vaccines with minimal severe side effects is the comparatively most effective way to improve vaccine uptake.

Figure 3 shows the estimated interaction indices. The estimated values range from -0.01 to 0.19. Table 5 indicates that none of the estimated interaction indices assume a statistically significant value below zero. Hence, the attributes are either synergistic or mutually independent.

A careful examination of the estimated interaction indices reveals that the non-pecuniary vaccine attributes should be well satisfied together, as they are synergistic. Effectiveness, which is the most important attribute according to Figure 2, interacts strongly with other attributes. The interactions of the effectiveness attribute are largest with severe side effects, incubation period, and protection period. The values of the respective interaction indices are 0.19, 0.17 and 0.16. Thus, to enhance vaccine attractiveness in the most effective way, efforts to improve vaccine effectiveness should be combined with efforts to extend the protection period and to reduce the incubation period and the risk of severe side effects. Also, the risk of severe side effects, the second most important attribute according to Figure 2, has pronounced synergies with other attributes. The attribute risk of severe side effects interacts most strongly with effectiveness, protection period, and the number of required doses. The values of the respective interaction indices are 0.19, 0.17 and 0.15. Protection period, the third most important attribute according to Figure 2, also exhibits strong positive interactions with other attributes, in particular with severe side effects, effectiveness and the number of required doses. Consequently, efforts to extend the protection period should be combined with efforts to reduce the risk of severe side effects, improve effectiveness, and lower the number of required doses. Also, the attribute risk of mild side effects has moderate synergistic interactions with other attributes, which again underlines that the non-pecuniary vaccine features should be well satisfied together.

By contrast, the attribute out-of-pocket costs interacts comparatively weakly with other attributes. Out-of-pocket costs are independent of effectiveness, incubation period, and mild side effects. Consequently, the attractiveness of a vaccine can be effectively increased by lowering out-of-pocket costs in isolation of these three attributes. However, out-of-pocket costs exhibit moderate synergies with the remaining attributes. For example, the synergistic interaction of out-of-pocket costs and protection period suggests that the two attributes should be well satisfied together.

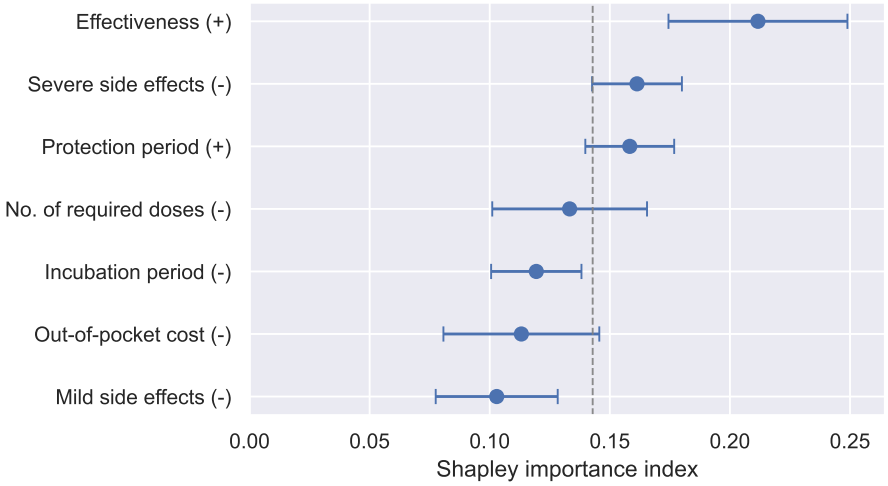

Figure 2: Estimated Shapley importance indices. The filled circles represent the point estimates. The error bars represent the 95% confidence intervals. The dashed vertical line indicates average importance. (+) indicates a desirable attribute, and (-) indicates an undesirable attribute.

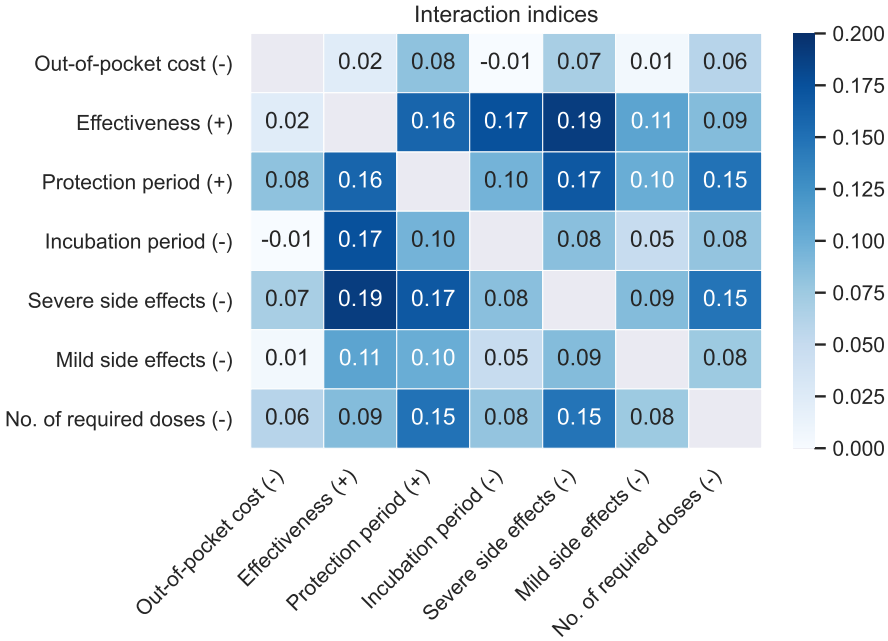

Figure 3: Heatmap of estimated interaction indices. The reported values are the point estimates. (+) indicates a desirable attribute, and (-) indicates an undesirable attribute.

Attribute 1	Attribute 2	Est.	SE	[2.5%	97.5%]
Out-of-pocket cost (-)	Effectiveness (+)	0.023	0.045	-0.066	0.112
Out-of-pocket cost (-)	Protection period (+)	0.083	0.021	0.042	0.123
Out-of-pocket cost (-)	Incubation period (-)	-0.006	0.024	-0.054	0.041
Out-of-pocket cost (-)	Severe side effects (-)	0.069	0.021	0.028	0.110
Out-of-pocket cost (-)	Mild side effects (-)	0.005	0.023	-0.040	0.051
Out-of-pocket cost (-)	No. of required doses (-)	0.060	0.023	0.016	0.104
Effectiveness (+)	Protection period (+)	0.159	0.027	0.107	0.212
Effectiveness (+)	Incubation period (-)	0.175	0.025	0.125	0.224
Effectiveness (+)	Severe side effects (-)	0.189	0.019	0.152	0.226
Effectiveness (+)	Mild side effects (-)	0.109	0.020	0.071	0.148
Effectiveness (+)	No. of required doses (-)	0.088	0.046	-0.001	0.178
Protection period (+)	Incubation period (-)	0.096	0.023	0.051	0.140
Protection period (+)	Severe side effects (-)	0.169	0.021	0.128	0.210
Protection period (+)	Mild side effects (-)	0.101	0.017	0.068	0.135
Protection period (+)	No. of required doses (-)	0.149	0.025	0.100	0.198
Incubation period (-)	Severe side effects (-)	0.085	0.027	0.032	0.138
Incubation period (-)	Mild side effects (-)	0.049	0.019	0.012	0.086
Incubation period (-)	No. of required doses (-)	0.080	0.026	0.030	0.130
Severe side effects (-)	Mild side effects (-)	0.088	0.022	0.044	0.132
Severe side effects (-)	No. of required doses (-)	0.147	0.025	0.097	0.197
Mild side effects (-)	No. of required doses (-)	0.076	0.025	0.026	0.126

Table 5: Estimated interaction indices

6 Conclusion

Mass immunisations with COVID-19 vaccines are viewed as the most effective way to end the global COVID-19 pandemic and the associated public health crisis. The success of mass vaccination campaigns depends critically on the decisions of individuals to get vaccinated. In this paper, we analyse individual preferences for COVID-19 vaccines using data from a nationwide stated choice survey (N=1421). The survey featured a discrete choice experiment consisting of a choice between two hypothetical COVID-19 vaccines and an opt-out alternative. Several attributes, including effectiveness, protection period, incubation period, risk of severe side effects, risk of mild side effects, the number of required doses, and the origin of the vaccine described the vaccine options. For the analysis of the stated choice data, we formulate and apply a new normal error components mixed logit (NECML) model in which the Choquet integral replaces the standard weighted sum operator to represent a component of the systematic utility. The Choquet integral is a flexible aggregation operation which captures interactions between attributes while ensuring interpretability and monotonicity of preferences. In our analysis, the new proposed model provides a significantly better goodness-of-fit than a conventional NECML model relying on a weighted sum aggregation.

Our empirical findings indicate that effectiveness is the most important vaccine attribute, followed by risk of severe side effects, and protection period. Even though these results are somewhat expected, our use of the Choquet integral and associated interaction anal-

ysis reveal that on the one hand the non-pecuniary vaccine attributes are synergistic and should thus be well satisfied together in order to maximise vaccine attractiveness. On the other hand, out-of-pocket costs are independent of effectiveness, incubation period, and mild side effects but exhibit moderate synergies with the remaining attributes. Also, we estimate that respondents prefer vaccines from the US. Our analysis of preferences for the opt-out alternative in the discrete choice experiment offers insights into the factors that are likely associated with vaccine (non-)adoption. We estimate that vaccine adoption is significantly more likely among individuals who identify as male, have obtained a bachelor's degree or a higher level of education, have a high household income, support the democratic party, had COVID-19, got vaccinated against the flu in winter 2020/21, and have an underlying health condition. By contrast, individuals who belong to the Baby Boomer generation or an older generation, and are black or African-American are significantly more likely to select the opt-out alternative.

Our analysis suggests that people's preferences should be considered in the design of information campaigns, vaccine procurement and the development of new vaccines. For example, information campaigns aimed at improving vaccine acceptance should emphasise vaccine attributes that are perceived as most important by respondents (i.e., effectiveness, risk of severe side effects, and protection period as elicited in our work by the estimated Shapley importance indices). Information campaigns should also explicitly target sociodemographic groups with a lower likelihood of vaccine adoption. In addition, our findings suggest that the likelihood of widespread vaccine adoption can be increased by improving the availability of vaccines that satisfy important attributes. Due to the synergistic interactions between vaccine attributes unveiled by the Choquet integral, the most effective way to maximise vaccine adoption is to improve the availability of vaccines that perform well across all non-pecuniary vaccine attributes. These insights should be exploited in the procurement of vaccines and the development of new vaccines.

This research is not devoid of limitations. First, our analysis does not account for systematic heterogeneity in preferences for attributes that enter the Choquet integral. As a remedy to this issue, (Dubey et al., 2021) parameterise the normalisation of the attributes as a function of individual-specific characteristics. Second, growing evidence suggests that stated choice methods possess a high external validity for explaining and predicting health-related behaviours (de Bekker-Grob et al., 2020). Nonetheless, stated choice data may still exhibit a hypothetical bias. One way to circumvent this limitation is to combine stated preference data with revealed preference data, a technique that is exercised in other application areas of discrete choice analysis (Ben-Akiva et al., 1994). To collect revealed preference data on vaccine preferences, clinical studies in which patients are given a choice between multiple COVID-19 vaccines could be conducted.

Acknowledgments

This research was supported by the National Science Foundation Award No. SES-2031841.

CRediT author statement

Rico Krueger: Conceptualization, Methodology, Software, Formal analysis, Resources, Writing - Original Draft, Visualization **Ricardo A. Daziano:** Conceptualization, Methodology, Software, Investigation, Resources, Data Curation, Writing - Review & Editing

References

Aggarwal, M. (2020). Logit choice models for interactive attributes. *Information Sciences*, 507:298–312.

Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., et al. (2021). Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *New England Journal of Medicine*, 384(5):403–416.

Basta, N. E., Moodie, E. E. M., and McGill University COVID19 Vaccine Tracker Team (2020). COVID-19 Vaccine Development and Approvals Tracker. `covid19.trackvaccines.org`.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., and Rao, V. (1994). Combining revealed and stated preferences data. *Marketing Letters*, 5(4):335–349.

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological*, 35(7):677–693.

Borriello, A., Master, D., Pellegrini, A., and Rose, J. M. (2021). Preferences for a COVID-19 vaccine in Australia. *Vaccine*, 39(3):473–479.

Centers for Disease Control and Prevention (2020). COVID Data Tracker. `https://covid.cdc.gov/covid-data-tracker`.

Choquet, G. (1954). Theory of capacities. In *Annales de l'institut Fourier*, volume 5, pages 131–295.

de Bekker-Grob, E., Donkers, B., Bliemer, M., Veldwijk, J., and Swait, J. (2020). Can healthcare choice be predicted using stated preference data? *Social Science & Medicine*, 246:112736.

de Bekker-Grob, E. W., Veldwijk, J., Jonker, M., Donkers, B., Huisman, J., Buis, S., Swait, J., Lancsar, E., Witteman, C. L., Bonsel, G., et al. (2018). The impact of vaccination and patient characteristics on influenza vaccination uptake of elderly people: A discrete choice experiment. *Vaccine*, 36(11):1467–1476.

Determann, D., Korfage, I. J., Lambooij, M. S., Bliemer, M., Richardus, J. H., Steyerberg, E. W., and de Bekker-Grob, E. W. (2014). Acceptance of vaccinations in pandemic outbreaks: a discrete choice experiment. *PloS one*, 9(7):e102505.

Dong, D., Xu, R. H., Wong, E. L.-y., Hung, C.-T., Feng, D., Feng, Z., Yeoh, E.-k., and Wong, S. Y.-s. (2020a). Public preference for COVID-19 vaccines in China: A discrete choice experiment. *Health Expectations*, 23(6):1543–1578.

Dong, E., Du, H., and Gardner, L. (2020b). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*, 20(5):533–534.

Dubey, S., Cats, O., Hoogendoorn, S., and Bansal, P. (2021). A Multinomial Probit Model with Choquet Integral and Attribute Cut-offs. *arXiv preprint arXiv:2105.03275*.

Eshun-Wilson, I., Mody, A., Tram, K. H., Bradley, C., Sheve, A., Fox, B., Thompson, V., and Geng, E. H. (2021). Preferences for COVID-19 vaccine distribution strategies in the US: A discrete choice survey. *PloS one*, 16(8):e0256394.

Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European journal of operational research*, 89(3):445–456.

Grabisch, M., Kojadinovic, I., and Meyer, P. (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European journal of operational research*, 186(2):766–785.

Grabisch, M. and Labreuche, C. (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175(1):247–286.

Leng, A., Maitland, E., Wang, S., Nicholas, S., Liu, R., and Wang, J. (2021). Individual preferences for COVID-19 vaccination in China. *vaccine*, 39(2):247–254.

Marichal, J.-L. (2002). Aggregation of interacting criteria by means of the discrete Choquet integral. In *Aggregation operators*, pages 224–244. Springer.

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., and Rodés-Guirao, L. (2021). A global database of COVID-19 vaccinations. *Nature human behaviour*, pages 1–7.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

McPhedran, R. and Toombs, B. (2021). Efficacy or delivery? An online Discrete Choice Experiment to explore preferences for COVID-19 vaccines in the UK. *Economics letters*, 200:109747.

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Marc, G. P., Moreira, E. D., Zerbini, C., et al. (2020). Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *New England Journal of Medicine*.

Tehrani, A. F., Cheng, W., Dembczyński, K., and Hüllermeier, E. (2012). Learning monotone nonlinear models using the Choquet integral. *Machine Learning*, 89(1):183–211.

Tenforde, M. W., Self, W. H., Adams, K., Gaglani, M., Ginde, A. A., McNeal, T., Ghamande, S., Douin, D. J., Talbot, H. K., Casey, J. D., et al. (2021). Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. *JAMA*.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020).

SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

Walker, J. L., Ben-Akiva, M., and Bolduc, D. (2007). Identification of parameters in normal error component logit-mixture (NECLM) models. *Journal of Applied Econometrics*, 22(6):1095–1125.

Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y., and Jit, M. (2021). Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *The Lancet*.