



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

Rico Krueger <sup>a,\*</sup>, Ricardo A. Daziano <sup>b</sup>

<sup>a</sup> Department of Technology, Management and Economics, Technical University of Denmark (DTU), Denmark

<sup>b</sup> School of Civil and Environmental Engineering, Cornell University, United States



### ARTICLE INFO

**Keywords:**

COVID-19  
Vaccines  
Patient preferences  
Stated choice  
Discrete choice

### 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.

### 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., 2020a). 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

\* Corresponding author.

E-mail addresses: [rickr@dtu.dk](mailto:rickr@dtu.dk) (R. Krueger), [daziano@cornell.edu](mailto:daziano@cornell.edu) (R.A. Daziano).

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 individual preferences for multi-attribute products and services due to their behavioural foundations in microeconomic theory (see e.g. Ben-Akiva et al., 2019). In that vein, these methods are increasingly common in health economics for investigating individual preferences for health-related products, treatments and services (see de Bekker-Grob et al., 2012; Clark et al., 2014; Soekhai et al., 2019). Especially discrete choice experiments (DCEs) are a meaningful instrument for eliciting preferences for multi-attribute alternatives because they reveal information about the relative importance of attributes (Ben-Akiva et al., 2019). DCEs aim to emulate real-life choice situations by asking respondents to select the most preferred option from a set of alternatives with strategically varied attributes in multiple scenarios (Ben-Akiva et al., 2019).

Given the advantages of stated choice methods and DCEs, several studies have employed these methods to investigate preferences for COVID-19 vaccines (Borriello et al., 2021; Dong et al., 2020b; Eshun-Wilson et al., 2021; Leng et al., 2021; McPhedran and Toombs, 2021). The DCEs 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, the out-of-pocket cost 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 using a weighted sum aggregation of the attributes.

Discrete choice models in which the systematic utility is specified using a weighted sum are limited in their ability to explain preferences for multi-attribute alternatives. This is because a weighted sum aggregation makes it difficult to represent attribute interdependencies while also maintaining interpretability and monotonicity (Dubey et al., 2022; Tehrani et al., 2012). Interpretability of preferences is a key desideratum in discrete 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 utility specification based on a weighted sum aggregation is found wanting in these two regards. 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. However, the weighted sum aggregation lacks expressiveness due to its inability to capture dependencies between attributes. To overcome this limitation, analysts may include two-way and higher-order interaction effects in a weighted sum 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 DCE 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 cost, 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. 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 doses 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 among individuals aged 55 years old or older.

[Dong et al. \(2020b\)](#) collected data via a stated choice survey in China in June and July 2020. The authors' mixed logit analysis finds that 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 DCE to investigate preferences for vaccine attributes in a hypothetical pandemic outbreak. The DCE considered several vaccine attributes, including effectiveness, safety, advice, media coverage and out-of-pocket cost. 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 cost 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. \(2022\)](#) 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

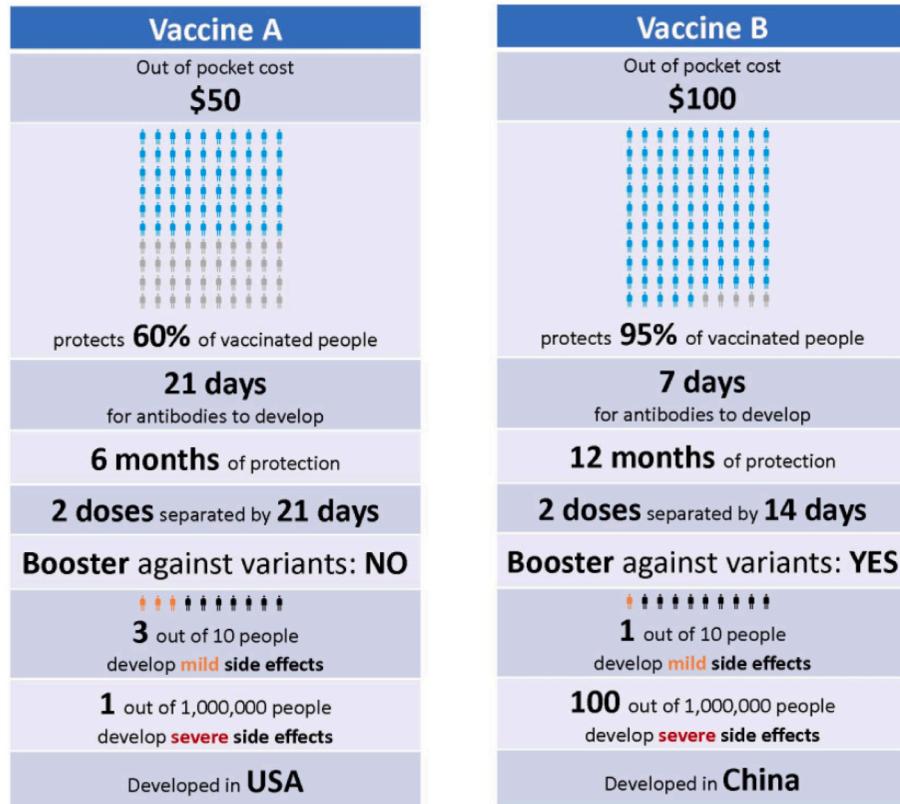
We conducted a nationwide stated choice survey in the US from 4 to 10 March 2021 to investigate preferences for COVID-19 vaccines. The survey included a DCE which involved a choice between two hypothetical COVID-19 vaccines and an opt-out alternative. In total, we collected 1421 valid responses. Each respondent completed seven choice scenarios. Details about the experiment, which uses a Bayesian efficient design with two-way interactions, are given in [Daziano \(2022\)](#).

The vaccines in the DCE were described by nine attributes, namely out-of-pocket cost, effectiveness, protection period, incubation period, risk of severe side effects, risk of mild side effects, number of required doses, whether the vaccine has a booster against variants, and origin of the vaccine. An example of a choice task is shown in [Fig. 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 among the general public. 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 cost was included to facilitate eventual welfare calculations. We also included the number of required doses with levels one and two as an attribute, 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.

If COVID-19 vaccines A and B as presented below were your only options, you would prefer:



Vaccine A

I would not get vaccinated

Vaccine B

Fig. 1. Example of a choice task.

**Table 1**  
Attributes and levels.

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)

#### 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  $n = 1, \dots, N$ . Every individual is observed to choose an alternative  $y_{nq}$  from the set  $\mathcal{C} = \{1, \dots, J\}$  in  $Q$  choice situations indexed by  $q = 1, \dots, Q$ . Each

**Table 2**  
Sample description ( $N = 1421$ ).

Variable	Sample proportion [%]
Gender: male	49.8
Cohort: Generation Z	4.2
Cohort: Millennial	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

alternative is described by a set  $X_{nqj} = \{x_{nqj1}, \dots, x_{nqjK}\}$  of  $K$  attributes. Random utility theory (McFadden et al., 1973) posits that an individual selects the alternative with the highest random utility, i.e.

$$y_{nq} = j \text{ iff } U_{nqj} > U_{nqj'} \forall j' \in \mathcal{C} \setminus j, \quad (1)$$

where

$$U_{nqj} = V_{nqj}(X_{nqj}; \theta) + \epsilon_{nqj} \quad (2)$$

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

$$V_{nqj}(X_{nqj}; \theta) = \mathcal{H}(X_{nqj}; \theta). \quad (3)$$

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

$$V_{nqj}(X_{nqj}; \theta = \beta) = W_\beta(X_{nqj}) = \sum_{k=1}^K \beta_k x_{nqjk} \quad (4)$$

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

#### 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  $q$  in the subsequent exposition.

#### 4.2.1. Definition

A fuzzy measure on a set of continuous attributes  $X = \{x_1, \dots, x_K\}$  of cardinality  $K$  is a set function  $\mu : 2^K \rightarrow [0, 1]$ , which satisfies the following two conditions:

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

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

For any  $S \subseteq X$ ,  $\mu(S)$  represents the weight or importance of the subset  $S$  of attributes in  $X$ .

The Choquet integral  $C$  is defined as

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

where  $(\cdot)$  is a permutation operator such that the attributes are arranged in non-decreasing order  $x_{(1)} \leq \dots \leq x_{(K)}$ . Furthermore,  $A_{(k)} = \{x_k, \dots, x_K\}$  and  $A_{(K+1)} = \emptyset$  is the subset of ordered attributes from the smallest to the largest starting with the  $k$ th smallest attribute. For example, if  $X = \{x_1, x_2, x_3\}$  and  $x_3 \leq x_1 \leq x_2$ , then

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

#### 4.2.2. Möbius representation

Fuzzy measures have an equivalent representation in terms of the Möbius transform, which simplifies the calculation of the Choquet integral and other quantities derived from the fuzzy measure, namely Shapley importance and interaction indices (see Sections 4.2.3 and 4.2.4).

The Möbius representation of a fuzzy measure  $\mu$  is a set function  $m_\mu : 2^K \rightarrow \mathbb{R}$  with

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

The corresponding inverse transform is

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

For  $\mu$  to be a valid fuzzy measure,  $m_\mu$  must satisfy the following conditions:

$$m(\emptyset) = 0, \quad \sum_{T \subseteq X} m(T) = 1, \quad (11)$$

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

In terms of the Möbius representation  $m_\mu$ , the Choquet integral can be equivalently expressed as:

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

where  $\min_{i \in T} x_i$  selects the smallest value in the subset  $T$  of attributes. In Möbius representation, the Choquet integral is thus simply a weighted sum of the minima of all subsets of attributes, whereby the weight of each minimum is the importance  $m_\mu(T)$  of the subset from which the minimum is selected.

Furthermore, the Möbius representation of the Choquet integral shows that the Choquet integral nests the weighted sum aggregation which is usually considered in the specification of random utility models. The usual weighted sum aggregation is obtained when the weights pertaining to subsets with more than one attribute are set to zero, i.e. when only main effects and no interaction effects are captured.

#### 4.2.3. Shapley importance index

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

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

The Shapley importance index can also be calculated more simply in terms of the Möbius representation  $m_\mu$ :

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

The Shapley importance index of  $x_k$  can be viewed as the weighted average of the marginal contribution of attribute  $x_k$  to all subsets of attributes  $T$  that exclude  $x_k$ .  $T \cup x_k$  is the union of  $T$  and  $x_k$ .  $|T| + 1$ , i.e. the cardinality of the union, is the inverse weight of the union used in the calculation of the weighted average in (15). The summation in (15) is over all subsets of attributes 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.

#### 4.2.4. Interaction index

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

$$\kappa_\mu(\{x_k, x_{k'}\}) = \sum_{T \subseteq X \setminus \{x_k, x_{k'}\}} \frac{(K - |T| - 2)! |T|!}{(K - 1)!} [\mu(T \cup \{x_k, x_{k'}\}) - \mu(T \cup x_k) - \mu(T \cup x_{k'}) + \mu(T)]. \quad (16)$$

The interaction index can also be calculated in terms of the Möbius representation  $m_\mu$ :

$$\kappa_{m_\mu}(\{x_k, x_{k'}\}) = \sum_{T \subseteq X \setminus \{x_k, x_{k'}\}} \frac{m_\mu(T \cup \{x_k, x_{k'}\})}{|T| + 1}. \quad (17)$$

The interaction index of a set of two attributes  $\{x_k, x_{k'}\}$  can be viewed as the average marginal interaction between  $x_k$  and  $x_{k'}$ . The interaction index is contained within  $[-1, 1]$ . A positive interaction index implies that two attributes  $x_k$  and  $x_{k'}$  are *synergistic*, i.e. improving  $x_k$  and  $x_{k'}$  jointly gives strictly more than improving either  $x_k$  or  $x_{k'}$ . A negative interaction index implies that two attributes  $x_k$  and  $x_{k'}$  are *redundant*, i.e. it is not necessary to improve  $x_k$  and  $x_{k'}$  jointly. An interaction index of zero implies that two attributes  $x_k$  and  $x_{k'}$  are *independent*.

#### 4.2.5. Practicalities

The Choquet integral is only applicable to continuous attributes. It is not applicable to ordinal and categorical attributes. An application of the Choquet integral requires that the continuous attributes in question 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

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

where  $\tilde{x}_{nqjk}$  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

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

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.

#### 4.2.6. Embedding the Choquet integral in a random utility model

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

$$V_{nqj} = \lambda C_{m_\mu}(X_{nqj}), \quad (20)$$

we have

$$U_{nqj} = \lambda C_{m_\mu}(X_{nqj}) + \epsilon_{nqj}, \quad (21)$$

whereby  $\lambda > 0$  is an unknown precision parameter. We include  $\lambda$  because the output of the Choquet integral is constrained between 0 and 1 by design.  $\lambda$  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.<sup>1</sup>

<sup>1</sup> One anonymous reviewer raised the question whether willingness to pay (WTP) indicators can be calculated using the Choquet integral representation of the systematic utility. Deriving WTP indicators using the Choquet integral is not always possible. This is because the Choquet integral is not differentiable with respect to an attribute  $x_k$  at any point where another attribute  $x_{k'}$  has the same value as  $x_k$ . This can be seen in (13), where the partial derivative of the minimum operator with respect to  $x_k$  does not exist if  $x_k = x_{k'}$  for any  $k, k' \in T | k \neq k'$ . Thus, deriving WTP indicators under the Choquet integral requires a specialised experimental design in which at any level of an attribute  $k$ , there is no other attribute  $k'$  that has the same value. In the current application, this requirement is not satisfied.

#### 4.3. Extensions

As explained in Section 4.2.5, the Choquet integral only accommodates continuous attributes. However, in many applications, analysts may wish to include ordinal and categorical attributes in the utility specification. Also, analysts may choose to not include certain continuous attributes in the Choquet integral. To accommodate such applications, the deterministic utility can be constructed using a combination of different aggregation operators. Let  $Z_{nqj} = \{z_{nqj1}, \dots, z_{nqjL}\}$  denote a second set of  $L$  attributes that the analyst does not wish to include in the Choquet integral. For example, we could specify

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

such that

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

where  $m_\mu$  and  $\beta$  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  $b = 1, \dots, B$ . Then, the random utility becomes

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

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

#### 4.4. Final model

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

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

with

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

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

$$P(y_n|X_{nqj}, Z_{nqj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n) = \prod_{q=1}^Q P(y_{nq}|X_{nqj}, Z_{nqj}, d_{jb}; \lambda, m_\mu, \beta, \sigma, \xi_n). \quad (27)$$

The unconditional probability is

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

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

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

Consequently, the simulated log-likelihood is given by

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

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

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

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

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

**Table 3**  
Model fit.

	WS-NECML	Choquet-NECML
No. of parameters	20	140
Log-likelihood	-7853.5	-7743.6

$$\lambda, \sigma \geq 0.$$

(34)

We implement the constrained maximum simulated likelihood estimation problem defined in (30)–(34) in Python. The unconditional choice probabilities are simulated using 200 randomised and shuffled Halton draws (Hess et al., 2006) per individual. The constrained maximisation of the simulated likelihood is performed using the sequential least squares programming method 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 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 cost, 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 Eqs. (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.<sup>2</sup> 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.

The model labelled WS-NECML is equivalent to a restricted Choquet-NECML model in which the Möbius parameters that pertain to more than one attribute are fixed to zero. The Möbius parameters of the restricted model are then equivalent to Shapley importance indices. To facilitate a comparison of WS-NECML and Choquet-NECML, we thus estimate a restricted Choquet-NECML model instead of a standard WS-NECML model. However, before we estimated the restricted Choquet-NECML model, we estimated a standard WS-NECML model to convince ourselves that the alternative-specific attributes in WS-NECML are consistent with our normalisation assumptions. As expected, we found that the estimates of parameters pertaining to “desirable” attributes (i.e. effectiveness and protection period) are positive, while the estimates of the parameters pertaining to the remaining “undesirable” attributes are negative.

### 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.<sup>3</sup> Note that WS-NECML is nested within Choquet-NECML 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 ( $\chi^2 = 219.836$ ,  $df = 120$ ,  $p < 0.001$ ).

<sup>2</sup> In the subsequent presentation of the estimation results, desirable attributes are denoted by ‘(+)’ and undesirable attributes are denoted by ‘(−)’.

<sup>3</sup> 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.

**Table 4**  
Parameter estimates.

Parameter	WS-NECML			Choquet-NECML		
	Est.	SE	z-val.	Est.	SE	z-val.
Out-of-pocket cost (-)	-0.238***	0.016	14.702			
Effectiveness (+)	-0.393***	0.019	20.471			
Protection period (+)	-0.109***	0.012	8.999			
Incubation period (+)	-0.072***	0.013	5.709			
Severe side effects (-)	-0.083***	0.014	5.781			
Mild side effects (-)	-0.068***	0.015	4.657			
No. required doses (-)	-0.037**	0.015	2.408			
Origin is USA	-1.128***	0.046	24.726	-1.424***	0.065	21.954
Opt-out	-2.778***	0.282	9.844	-2.781***	0.316	8.810
Opt-out $\times$ male	-0.907***	0.183	-4.948	-0.938***	0.190	-4.932
Opt-out $\times$ cohort is Baby Boomer or older	-1.148***	0.187	6.139	-1.209***	0.190	6.379
Opt-out $\times$ education bachelor	-0.829***	0.215	-3.851	-0.861***	0.225	-3.818
Opt-out $\times$ education bachelor	-1.466***	0.275	-5.331	-1.484***	0.285	-5.197
Opt-out $\times$ household income [10k USD]	-0.054***	0.018	-3.040	-0.054***	0.019	-2.912
Opt-out $\times$ black or African-American	-0.669**	0.288	2.323	-0.715**	0.299	2.392
Opt-out $\times$ democrat	-1.551***	0.218	-7.105	-1.599***	0.228	-7.015
Opt-out $\times$ had COVID-19	-0.593***	0.276	-2.144	-0.600**	0.290	-2.067
Opt-out $\times$ received flu shot	-1.294***	0.197	-6.578	-1.371***	0.207	-6.628
Opt-out $\times$ has underlying condition	-0.366*	0.201	-1.824	-0.388*	0.209	-1.857
Opt-out std. dev.	-2.807***	0.120	23.323	-2.948***	0.124	23.768
Choquet precision $\lambda$	-2.324***	0.135	17.227	-7.513***	1.210	6.209

\*\*\*Significance levels:  $p < 0.01$ .

\*\*Significance levels:  $p < 0.05$ .

\*Significance levels:  $p < 0.1$ .

(+) indicates a desirable attribute; (-) indicates an undesirable attribute.

### 5.3. Parameter estimates

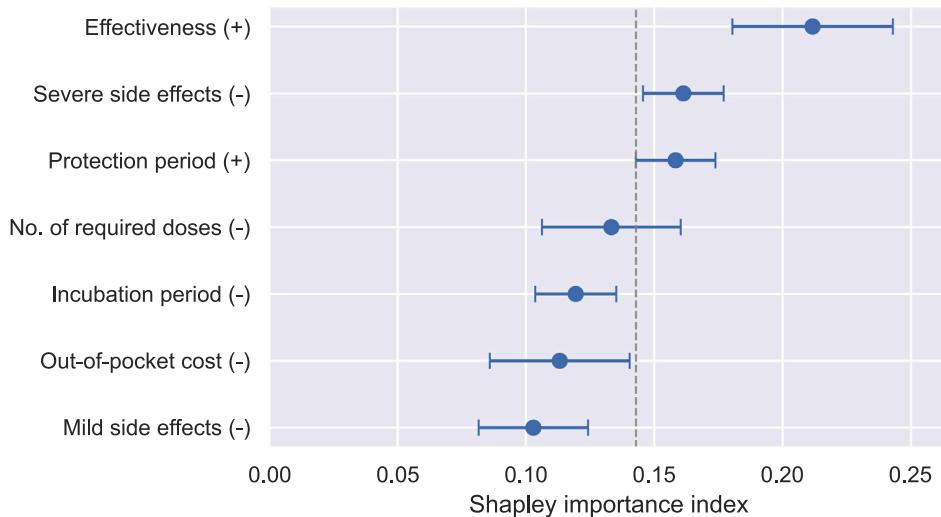
In Table 4, we report the parameter estimates for the two models. We omit the estimates of the Möbius parameters, as we will interpret these parameters in terms of their Shapley importance and interaction representations (see Sections 5.4 and 5.5). Both models indicate that respondents have a positive 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 decreases 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  $\lambda$  does not carry a substantive meaning. The estimate of  $\lambda$  is smaller in the model labelled WS-NECML than in the model labelled Choquet-NECML because WS-NECML does not account for interaction effects.

### 5.4. Shapley importance indices

Figs. 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.

Fig. 2 shows the relative importance of the attributes. The dashed vertical line in the plot indicates average importance. The error bars represent the 90% 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 cost, 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. Protection period is not significantly more important than the average because the 90% confidence interval of the respective Shapley importance index includes  $1/K$ . 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.

The values reported in Fig. 2 can be compared with the values reported in Table 4 because the model labelled WS-NECML is a restricted Choquet-NECML model in which Möbius parameters pertaining to more than one attribute are fixed to zero. Thus, the Möbius parameters in WS-NECML are equivalent to Shapley importance indices. We observe that WS-NECML offers quite different insights about the relative importance of attributes than the model labelled Choquet-NECML because WS-NECML does not account for interaction effects. WS-NECML suggest that effectiveness and out-of-pocket cost are more important than the average, whereas the remaining attributes are less important than the average.



**Fig. 2.** Estimated Shapley importance indices. The filled circles represent the point estimates. The error bars represent the 90% confidence intervals. The dashed vertical line indicates average importance. (+) indicates a desirable attribute, and (−) indicates an undesirable attribute.

### 5.5. Interaction indices

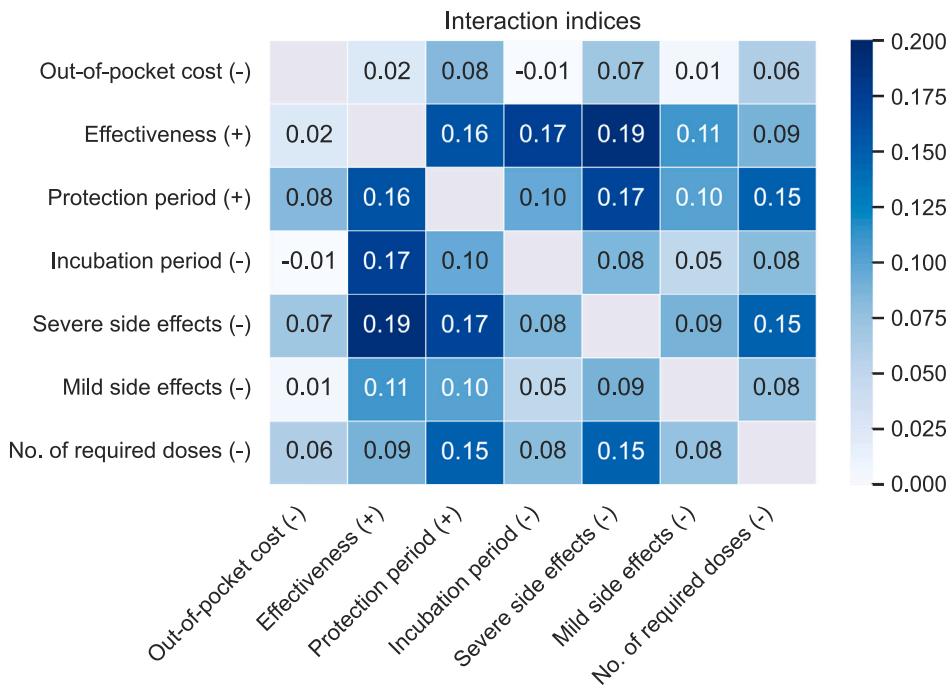
Fig. 3 shows the estimated interaction indices. The estimated values range from  $-0.01$  to  $0.19$ . Table 5 indicates that none of the estimated interaction indices assume a statistically significant value below zero. Hence, the attributes are either synergistic or mutually independent.

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 Fig. 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 Fig. 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 Fig. 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 cost interacts comparatively weakly with other attributes. Out-of-pocket cost is 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 cost exhibits moderate synergies with the remaining attributes. For example, the synergistic interaction of out-of-pocket cost and protection period suggests that the two attributes should be well satisfied together.

## 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 (DCE) consisting of a choice between two hypothetical COVID-19 vaccines and an opt-out alternative. Several attributes, including out-of-pocket cost, effectiveness, protection period, incubation period, risk of severe side effects, risk of mild side effects, number of required doses, whether the vaccine has a booster against variants, and 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.



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

**Table 5**  
Estimated interaction indices.

Attribute 1	Attribute 2	Est.	SE	[5%]	95%
Out-of-pocket cost (-)	Effectiveness (+)	0.023	0.045	-0.052	0.097
Out-of-pocket cost (-)	Protection period (+)	0.083	0.021	0.049	0.117
Out-of-pocket cost (-)	Incubation period (-)	-0.006	0.024	-0.046	0.034
Out-of-pocket cost (-)	Severe side effects (-)	0.069	0.021	0.034	0.103
Out-of-pocket cost (-)	Mild side effects (-)	0.005	0.023	-0.033	0.044
Out-of-pocket cost (-)	No. of required doses (-)	0.060	0.023	0.023	0.097
Effectiveness (+)	Protection period (+)	0.159	0.027	0.115	0.203
Effectiveness (+)	Incubation period (-)	0.175	0.025	0.133	0.216
Effectiveness (+)	Severe side effects (-)	0.189	0.019	0.158	0.221
Effectiveness (+)	Mild side effects (-)	0.109	0.020	0.077	0.142
Effectiveness (+)	No. of required doses (-)	0.088	0.046	0.013	0.163
Protection period (+)	Incubation period (-)	0.096	0.023	0.058	0.133
Protection period (+)	Severe side effects (-)	0.169	0.021	0.134	0.204
Protection period (+)	Mild side effects (-)	0.101	0.017	0.073	0.129
Protection period (+)	No. of required doses (-)	0.149	0.025	0.107	0.190
Incubation period (-)	Severe side effects (-)	0.085	0.027	0.040	0.129
Incubation period (-)	Mild side effects (-)	0.049	0.019	0.018	0.080
Incubation period (-)	No. of required doses (-)	0.080	0.026	0.038	0.122
Severe side effects (-)	Mild side effects (-)	0.088	0.022	0.051	0.125
Severe side effects (-)	No. of required doses (-)	0.147	0.025	0.105	0.189
Mild side effects (-)	No. of required doses (-)	0.076	0.025	0.034	0.118

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 analysis reveal that on 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 DCE 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 socio-demographic 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. \(2022\)](#) parameterise the normalisation of the attributes as a function of individual-specific characteristics. Second, the considered DCE is based on a Bayesian efficient design with two-way interactions. Future research could provide insights about the influence of discrete choice experimental design on models exploiting the Choquet integral. Third, 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.

#### CRediT authorship contribution 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.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgment

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

#### References

Aggarwal, M., 2020. Logit choice models for interactive attributes. *Inform. Sci.* 507, 298–312.

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

Basta, N.E., Moodie, E.E.M., McGill University COVID19 Vaccine Tracker Team, 2020. COVID-19 vaccine development and approvals tracker. URL <https://covid19.trackvaccines.org>.

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

Ben-Akiva, M., McFadden, D., Train, K., et al., 2019. Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. *Found. Trends® Econom.* 10 (1–2), 1–144.

Borriello, A., Master, D., Pellegrini, A., 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. URL <https://covid.cdc.gov/covid-data-tracker>.

Choquet, G., 1954. Theory of capacities. In: *Annales de L'Institut Fourier*. 5, pp. 131–295.

Clark, M.D., Determann, D., Petrou, S., Moro, D., de Bekker-Grob, E.W., 2014. Discrete choice experiments in health economics: A review of the literature. *Pharmacoeconomics* 32 (9), 883–902.

Daziano, R.A., 2022. A choice experiment assessment of stated early response to COVID-19 vaccines in the USA. *Health Econ. Rev.* 12 (1), 23.

de Bekker-Grob, E., Donkers, B., Bliemeier, M., Veldwijk, J., Swait, J., 2020. Can healthcare choice be predicted using stated preference data? *Soc. Sci. Med.* 246, 112736.

de Bekker-Grob, E.W., Ryan, M., Gerard, K., 2012. Discrete choice experiments in health economics: A review of the literature. *Health Econ.* 21 (2), 145–172.

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., Bliemeier, M., Richardus, J.H., Steyerberg, E.W., de Bekker-Grob, E.W., 2014. Acceptance of vaccinations in pandemic outbreaks: A discrete choice experiment. *PLoS One* 9 (7), e102505.

Dong, E., Du, H., Gardner, L., 2020a. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* 20 (5), 533–534.

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

Dubey, S., Cats, O., Hoogendoorn, S., Bansal, P., 2022. A multinomial probit model with choquet integral and attribute cut-offs. *Transp. Res. B* 158, 140–163.

Eshun-Wilson, I., Mody, A., Tram, K.H., Bradley, C., Sheve, A., Fox, B., Thompson, V., 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 J. Oper. Res.* 89 (3), 445–456.

Grabisch, M., Kojadinovic, I., 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 J. Oper. Res.* 186 (2), 766–785.

Grabisch, M., Labreuche, C., 2010. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Ann. Oper. Res.* 175 (1), 247–286.

Hess, S., Train, K.E., Polak, J.W., 2006. On the use of a modified latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. *Transp. Res. B* 40 (2), 147–163.

Leng, A., Maitland, E., Wang, S., Nicholas, S., Liu, R., 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*. Springer, pp. 224–244.

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., Rodés-Guirao, L., 2021. A global database of COVID-19 vaccinations. *Nat. Hum. Behav.* 1–7.

McFadden, D., et al., 1973. Conditional logit analysis of qualitative choice behavior. Institute of Urban and Regional Development, University of California.

McPhedran, R., Toombs, B., 2021. Efficacy or delivery? An online discrete choice experiment to explore preferences for COVID-19 vaccines in the UK. *Econom. Lett.* 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. *N. Engl. J. Med.*

Soekhai, V., de Bekker-Grob, E.W., Ellis, A.R., Vass, C.M., 2019. Discrete choice experiments in health economics: Past, present and future. *Pharmacoeconomics* 37 (2), 201–226.

Tehrani, A.F., Cheng, W., Dembczyński, K., Hüllermeier, E., 2012. Learning monotone nonlinear models using the Choquet integral. *Mach. Learn.* 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., SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods* 17, 261–272. <http://dx.doi.org/10.1038/s41592-019-0686-2>.

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

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