The Patient - Patient-Centered Outcomes Research (2019) 12:1–14
https://doi.org/10.1007/s40271-018-0324-6

PRACTICAL APPLICATION

Including Opt-Out Options in Discrete Choice Experiments: Issues to Consider

Danny Campbell[1] · Seda Erdem[1]

Published online: 2 August 2018

Abstract

Providing an opt-out alternative in discrete choice experiments can often be considered to be important for presenting real-life choice situations in different contexts, including health. However, insufficient attention has been given to how best to address choice behaviours relating to this opt-out alternative when modelling discrete choice experiments, particularly in health studies. The objective of this paper is to demonstrate how to account for different opt-out effects in choice models. We aim to contribute to a better understanding of how to model opt-out choices and show the consequences of addressing the effects in an incorrect fashion. We present our code written in the R statistical language so that others can explore these issues in their own data. In this practical guideline, we generate synthetic data on medication choice and use Monte Carlo simulation. We consider three different definitions for the opt-out alternative and four candidate models for each definition. We apply a frequentist-based multimodel inference approach and use performance indicators to assess the relative suitability of each candidate model in a range of settings. We show that misspecifying the opt-out effect has repercussions for marginal willingness to pay estimation and the forecasting of market shares. Our findings also suggest a number of key recommendations for DCE practitioners interested in exploring these issues. There is no unique best way to analyse data collected from discrete choice experiments. Researchers should consider several models so that the relative support for different hypotheses of opt-out effects can be explored.

Key Points for Decision Makers

Overlooking opt-out effects in discrete choice experiments can lead to erroneous policy recommendations.

Opt-out effects are context specific, and there are a myraid of potential reasons that may explain why participants choose the opt-out alternative, meaning that there is no unique best way to analyse opt-out choices.

Practitioners should consider many models and subsequently apply a multimodel inference procedure so that the relative support for each model can be assessed.

1 Introduction

Discrete choice experiments (DCEs) are now an established method for preference elicitation and non-market valuation in health and other areas of applied economics. In health economics, they have been applied to elicit preferences for a broad range of health service interventions, treatments,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40271-018-0324-6) contains supplementary material, which is available to authorized users.

Seda Erdem
seda.erdem@stir.ac.uk

Danny Campbell
danny.campbell@stir.ac.uk

[1] Economics Division, University of Stirling Management School, Stirling, UK

devices and medications. In these applications, participants (e.g. patients, health professionals or carers) are typically asked to choose between two or more product or service alternatives based on their preferences and the attributes that describe these alternatives (e.g. see Craig et al. [1] for a recent overview). In many DCE applications, participants are also provided with an alternative that is not designed by the experimenter, but represents an 'opt-out' option [2, 3]. This opt-out alternative—also referred to as the 'status quo'—is the participant's reference point or current situation [4].

The inclusion of the opt-out alternative in DCEs depends on the research question [5]. For example, if the research seeks to predict the likely adoption of a new intervention, service, treatment or medication, it is necessary to include an opt-out option. It creates realism in the sense that participants are not forced to choose between the experimentally designed alternatives and can, instead, opt out. Ensuring participants choose in a way that is consistent with how they would do in a real-life situation is important for welfare-consistent estimation of DCEs [6]. Indeed, in such cases, restricting the choice to be between two or more potentially unappealing alternatives raises concerns of external validity [2]. If alternatives are unlikely to be chosen in practice, any interpretations of the estimated marginal utilities and choice share predictions may well be inappropriate. For these reasons, the inclusion of an opt-out option in DCEs is generally recommended (e.g. see Lancsar and Louviere [3], Louviere and Lancsar [6] and Bridges et al. [7] for justification in health applications, and Johnston et al. [8] for contemporary guidance on the opt-out alternative for stated preference practitioners in general). On the other hand, if the objective of the study is primarily to estimate marginal rates of substitution among attributes and to compare levels and attributes or alternatives of the choice experiments, an opt-out option may be unnecessary and so forced choice tasks could be applied [9, 10].

While the inclusion of an opt-out alternative is widespread practice, it is much less clear how the opt-out alternative should be defined and presented to participants. Researchers designing DCEs have some latitude in the manner in which this alternative is defined (e.g. as a 'none of these', or as an 'actual status quo' described by the baseline attribute levels or a participant's current levels). Most importantly, however, what is meant by the opt-out alternative should be clear to participants. It should be understood and viewed as credible and in a manner that allows participants to anticipate the likely effects on their welfare [8]. Since some opt-out definitions are considered more real or plausible than others, researchers are faced with the challenge of how best to present the hypothetical market for health goods or services in question so that it resembles what the real-life choice situation might look like. There is also the need to be mindful of the fact that the opt-out alternative can draw disproportionately from the other available alternatives, such that the inclusion of the opt-out alternative may affect the relative choice shares observed for the other alternatives [5, 11–13]. While the opt-out alternative is a genuine choice in cases where a participant feels that it is most aligned with their preferences, it is well known and documented within the DCE literature that the propensity of participants choosing this alternative is often explained by more than just its attributes. One of the leading explanations for this is the endowment effect [14–17], whereby participants have a preference for retaining their current situation and thus a tendency to choose what they already have (even if the other alternatives are clearly superior). Relatedly, the opt-out alternative, which is often experienced by the participant, is in many instances perceived differently from the other alternatives. As a result, the potential losses or gains associated with the experimentally designed alternatives are considered relative to the opt-out alternative [18]. A further reason for the selection of opt-out alternatives is to avoid making difficult trade-offs [19]. This reluctance to choose is further subdivided by Boxallet al. [4] into a preference for inaction (omission) or a statement of non-participation ('choose none'). Similarly, a failure of the participant to understand the choice context may also give rise to opt-out choices. Also, when two or more of the non-opt-out alternatives have significant advantages and disadvantages on the basis of some (or all) of the attributes (thus making the choice difficult), or when the choice can be delayed, participants often revert back to their default or status quo [20]. Baron and Ritov [21] also find that people prefer bearing the consequences of inaction by sticking to their status quo rather than those of a wrong action by choosing an alternative that is not their usual option. Participants who are attempting to be equivocal, provide strategic or protest responses, or who do not have a strong opinion or preference may also be more inclined to choose the opt-out alternative. This would also hold for situations where participants are indecisive or indifferent between presented choice alternatives [22]. In such cases, the opt-out choices would not provide information about the attractiveness of non-opt-out alternatives in choice tasks [23].

Whatever the reason might be, it is important to account for these opt-out effects when modelling DCEs. Indeed, overlooking these effects could lead to erroneous policy recommendations and inaccurate measurement of welfare [4], since participants' preferences and decision-making behaviour are not appropriately reflected. This is a particular concern in health-focused DCEs, given that opt-out effects may be more prominent in the domain of public policy outcomes [24]. However, the myriad of different reasons that may explain why participants choose the opt-out alternative can make it difficult to know what model specification should be used. Ideally, the model should enable the researcher to distinguish between situations where participants have made genuine opt-out choices and where they exhibit behaviour beyond the theory of rational choice. This paper demonstrates how to account for some of the possible different opt-out effects in choice models using simulated datasets with different choice behaviours. While, admittedly, this paper does not provide the last word in the opt-out issues (since this would require a conceptual framework and testing across many different empirical datasets in different contexts), it does provide a primer on how to model opt-out choices. In doing so, we hope that it contributes to a better understanding of the issues and shows some of the potential consequences of addressing the effects in an incorrect fashion. Our intention is to provide a resource for practitioners who are currently using or considering using DCEs and are keen to explore potential opt-out effects present in their own data. To facilitate this, and for the purpose of replication, we provide our codes written in the R statistics program [25] for our practical demonstration.

The remainder of the paper is structured as follows: in Sect. 2 we outline some of the modelling approaches for dealing with opt-out effects in DCEs; Sect. 3 presents a practical demonstration of how to econometrically deal with opt-out effects in DCEs using simulated data; Sect. 4 reports the main findings; and Sect. 5 concludes and provide advice for other practitioners exploring opt-out effects.

2 Modelling Discrete Choice Data with Opt-Out Alternatives

2.1 Background Notation

Starting with the conventional random utility maximisation framework,[1] we specify the utility, U, where participants are indexed by n, chosen alternatives by i, the set of available alternatives by J, and choice occasions by z, and attributes of the chosen alternative are represented by the column vector $\mathbf{x}_{niz}$. We have

$$U_{niz} = \boldsymbol{\beta}\mathbf{x}_{niz} + \epsilon_{niz}, \tag{1}$$

where $\boldsymbol{\beta}$ is the row vector of marginal utility parameters for the attributes, ϵ_{niz} is an error term from an independent and identically distributed type I extreme value distribution with variance $\pi^2/6\lambda^2$, and λ is a scale parameter (that, for identification purposes, is normally set to 1). Given these assumptions, the probability of the sequence of choices made by participant n can be represented by the multinomial logit model

$$\Pr(\mathbf{y}_n|\mathbf{X}_n, \boldsymbol{\beta}, \lambda = 1) = \prod_{z=1}^{Z_n} \frac{\exp\left(\boldsymbol{\beta}\mathbf{x}_{niz}\right)}{\sum_{j=1}^{J} \exp\left(\boldsymbol{\beta}\mathbf{x}_{njz}\right)}, \tag{2}$$

where $\mathbf{y}_n$ gives the sequence of choices over the Z_n choice occasions for participant n, $\mathbf{y}_n = \left[i_{n1}\ i_{n2}\ \dots\ i_{nZ_n}\right]$.

The choice probability retrieved from Eq. 2 assumes that the likelihood of choice depends only on the attribute levels of each alternative, and that the error terms are uncorrelated over alternatives and have the same variance. However, it is important to recognise that the probability of choice may depend not only on the utilities associated with the attributes, but also on opt-out effects. This is because there can be systematic differences in preferences, substitution patterns or decision rules for the opt-out alternative. In such cases, the multinomial logit model can be inappropriate, meaning that different model specifications may be warranted. In this practical demonstration, we compare four alternative model specifications: (i) a multinomial logit model with an opt-out alternative-specific constant to accommodate the average influence of factors that are not explained by the attributes on opt-out choices, (ii) a nested logit model that permits the random error terms for the non-opt-out alternatives to share a common component, (iii) an independent availability logit model to allow for elimination-by-aspect-like behaviour, and (iv) a combination of (i)–(iii).

2.2 Modelling Opt-Out Effects

2.2.1 Multinomial Logit Model with an Opt-Out Alternative-Specific Constant

The most straightforward approach for addressing opt-out effects is to introduce an alternative-specific constant, γ, into the utility function for the opt-out alternative:

$$\Pr\left(\mathbf{y}_n|\mathbf{X}_n, \boldsymbol{\beta}, \lambda = 1, \gamma\right) = \prod_{z=1}^{Z_n} \frac{\exp\left(\boldsymbol{\beta}\mathbf{x}_{niz} + \mathbb{I}_i\gamma\right)}{\sum_{j=1}^{J} \exp\left(\boldsymbol{\beta}\mathbf{x}_{njz} + \mathbb{I}_j\gamma\right)}, \tag{3}$$

where $\mathbb{I}$ is an indicator variable equal to 1 when the alternative is the opt-out alternative and zero otherwise, so that the constant is only added to the utility expression for the opt-out alternative. In cases where there are no systematic differences between the average effect of factors not included in the utility expressions for the non-opt-out alternatives and the opt-out alternatives, we would be unable to reject the null hypothesis that $\gamma = 0$. However, in situations where $\gamma \neq 0$ (either negative or positive), the systematic differences have a bearing on choice probabilities, meaning that there

[1] We note that random utility maximisation is not the only framework for modelling choices. Indeed, for certain decisions, other choice axioms may be better suited, such as regret minimisation. In this paper, we utilise the most widely used framework to analyse opt-out effects.

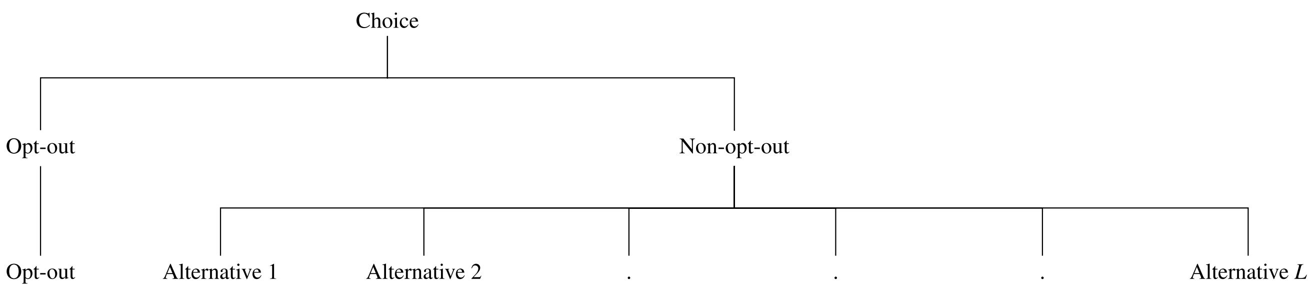

Fig. 1 Tree diagram showing the hierarchy of opt-out and L non-opt-out choices

are opt-out effects. While this additional parameter captures the average effect of all factors that influence opt-out choices that are not included in the utility specification, it should be noted that this parameter includes various components (e.g. status-quo bias, unobserved attributes and the impacts of complexity). This means that its interpretation as a utility parameter can be unclear.

2.2.2 Nested Logit Model

The opt-out alternative is often experienced by participants, while the experimentally designed alternatives can only be imagined. For this reason, participants may consider the non-opt-out alternatives as substitutes, meaning that the utilities of the non-opt-out alternatives may be more correlated among themselves than with the opt-out alternative (e.g. see Scarpa et al. [26] for a discussion). A nested logit specification may, therefore, be an appropriate approach for exploring opt-out effects since it can accommodate this sort of substitution pattern.

To illustrate this substitution pattern, consider the tree diagram in Fig. 1. This depicts an upper-level choice between opting out and not opting out and a lower-level (conditional) choice between L alternatives in the non-opt-out 'nest'.[2] In this case, the error terms of the non-opt-out alternatives are correlated. This violates the multinomial logit model assumption that the error terms are independently distributed. A nested logit model can be specified to address this opt-out effect that assumes the random terms for the non-opt-out alternatives can be partitioned into a distinct (i.e. alternative-specific) random component and a common random component that is shared across all non-opt-out alternatives. This common random component leads to covariance between the overall errors (both distinct and common errors) for the non-opt-out alternatives. However, the errors for the non-opt-out alternatives remain uncorrelated with those of the opt-out alternative (see Train [27] for further details).

The overall error terms for every alternative are assumed to be type I extreme value distributed with variance $\pi^2/6\lambda^2$, as in the multinomial logit model (where $\lambda = 1$). The distinct and common error components for non-opt-out alternatives are also assumed to be type I extreme value distributed, but with variances $\pi^2\mu^2/6$ and $\pi^2\left(1-\mu^2\right)/6$ respectively, where $0 \le \mu \le 1$. The value $1-\mu$ can be used as a measure of correlation or substitution: $\mu = 0$ leads to perfect correlation between pairs of non-opt-out alternatives, meaning that the choice between the non-opt-out alternatives is deterministic; $\mu = 1$ signifies zero correlation among non-opt-out alternatives, which is equivalent to the multinomial logit model; and $0 < \mu < 1$ implies non-zero correlation among non-opt-out alternatives, with increased substitution as $\mu \to 0$. Therefore, in cases where there is substitution among the non-opt-out alternatives, we can reject the null hypothesis that $\mu = 1$.

The upper-level marginal choice probability, $\Pr\left(m_{niz}|\cdot\right)$, in Fig. 1 between opting out and not opting for participant n in choice occasion z is given by

$$\Pr\left(m_{niz}|\mathbf{X}_{nz}, \boldsymbol{\beta}, \lambda = 1, \mu\right) = \frac{\exp\left[\mathbb{I}_i\left(\boldsymbol{\beta}\mathbf{x}_{niz}\right) + \left(1-\mathbb{I}_i\right)\mu\Upsilon\right]}{\exp(\boldsymbol{\beta}\mathbf{x}_{n\mathbb{I}_j z}) + \exp\left(\mu\Upsilon\right)}. \tag{4a}$$

The respective lower-level conditional choice probability, $\Pr\left(c_{niz}|\cdot\right)$, in Fig. 1 is expressed using

$$\Pr\left(c_{niz}|\mathbf{X}_{nz}, \boldsymbol{\beta}, \lambda = 1, \mu\right) = \begin{cases} 1 & \text{if the opt-out alternative is chosen; and,} \\ \exp\left(\dfrac{\boldsymbol{\beta}\mathbf{x}_{niz}}{\mu} - \Upsilon\right) & \text{for alternatives within the non-opt-out nest.} \end{cases} \tag{4b}$$

The marginal and conditional probabilities are linked by the term Υ, which can be interpreted as the expected utility that

[2] Note, however, that the derivation of the nested logit model does not necessarily imply that participants make choices in this hierarchical manner.

a participant derives from the choice among the non-opt-out alternatives:

$$\Upsilon = \ln\left[\sum_{l=1}^{L}\exp\left(\frac{\boldsymbol{\beta}\mathbf{x}_{nlz}}{\mu}\right)\right]. \tag{4c}$$

The nested logit choice probability can, therefore, be expressed as the product of the upper-level marginal choice probability and the lower-level conditional probability, meaning that the overall probability for the sequence of choices made by participant n is given by

$$\begin{aligned}&\Pr\left(\mathbf{y}_n|\mathbf{X}_n,\boldsymbol{\beta},\lambda=1,\mu\right)\\&=\prod_{z=1}^{Z_n}\left[\Pr\left(m_{niz}|\mathbf{X}_{nz},\boldsymbol{\beta},\lambda=1,\mu\right)\Pr\left(c_{niz}|\mathbf{X}_{nz},\boldsymbol{\beta},\lambda=1,\mu\right)\right].\end{aligned} \tag{4d}$$

2.2.3 Independent Availability Logit Model

The models introduced thus far assume that all participants consider all offered alternatives, including those that are unacceptable to them. However, suppose some participants, for whatever reason, have an overwhelming preference for the current state of affairs and that any change from this opt-out baseline is perceived as a loss. In an extreme case, these participants may exclude the non-opt-out alternatives from their actual consideration set and therefore consistently choose the opt-out alternative, a phenomenon often referred to as serial non-participation (e.g. see von Haefen et al. [28], Meyerhoff and Liebe [16] and Boxall et al. [4] for a discussion). Conversely, other participants, for whatever reason, may have a strong dislike of the opt-out, in which case they adopt a semi-compensatory choice process with the non-opt-out alternatives constituting their actual consideration set. These participants make their choice among the alternatives within this consideration set following a utility maximisation compensatory rule. The standard consideration set assumption may therefore be inappropriate. Following Manski [29], a probabilistic model can be formulated to account for this type of choice behaviour to help distinguish between the experimentally designed choice task that is presented to participants and the participant's actual consideration set. In order to achieve this, we consider the independent availability logit model (e.g. see Frejinger et al. [30], Campbell et al. [31], Kaplan et al. [32] and Campbell and Erdem [33] for some examples). Under this specification, the choice probability is given by

$$\Pr\left(\mathbf{y}_n|\mathbf{X}_n,\boldsymbol{\beta},\lambda=1,\boldsymbol{\phi}\right)=\sum_{s=1}^{S}\phi_s\Pr\left(\mathbf{y}_n|C_s,\mathbf{X}_n,\boldsymbol{\beta},\lambda=1\right), \tag{5a}$$

where $\Pr\left(\mathbf{y}_n|C_s,\mathbf{X}_n,\boldsymbol{\beta},\lambda=1\right)$ is the conditional probability of the sequence of choices given that the consideration set is $C_s\subseteq S$, where S is the set of all subsets and ϕ_s is the unconditional probability that C_s is the 'true' consideration set. Specifically, S is the set of all non-empty subsets of C_s (i.e. all the potential choice subsets, which we describe below in the context of this practical demonstration). Since a participant's true consideration set cannot be known with certainty, the model assumes that actual choice tasks are latent and vary across the S classes. While conditional on the consideration set (and hence the class) the choice probability is multinomial logit:

$$\Pr\left(\mathbf{y}_n|C_s,\mathbf{X}_n,\boldsymbol{\beta},\lambda=1\right)=\prod_{z=1}^{Z_n}\frac{\exp\left(\boldsymbol{\beta}'\mathbf{x}_{niz}\right)}{\sum_{j\in C_s}\exp\left(\boldsymbol{\beta}'\mathbf{x}_{njz}\right)}. \tag{5b}$$

Typically, in an independent availability logit model, the number of classes, S, is determined as a function of the number of alternatives (e.g. for a universal set with J alternatives, there are 2^J-1 possible consideration sets). Here, however, we are interested in exploring whether some participant's choices are governed by an elimination-by-aspects decision rule (see Erdem et al. [34]), whereby they restrict their choice task on the basis of the opt-out alternative. Based on this, three types of behaviour can be identified: (i) a subset ($C_{s=1}$) who always only consider (and choose) the opt-out alternative (perhaps for genuine reasons or due to serial non-participation), (ii) a subset ($C_{s=2}$) who restrict their actual choice task to only the non-opt-out alternatives, and (iii) a subset ($C_{s=3}$) whose actual choice task consists of all alternatives offered in the choice task (i.e. who consider both the opt-out and non-opt-out alternatives). These three patterns (i.e. $S=\left\{C_{s=1},C_{s=2},C_{s=3}\right\}$) can be dealt with using an independent availability logit model with three latent classes, where each class describes a unique consideration set. As noted above, the alternatives considered by a participant cannot be known with certainty and therefore remain latent. However, their observed choice behaviour helps us to make probabilistic statements about the likelihood of competing consideration sets being their true choice task, with the full probability per participant allocated across all S classes (i.e. $\sum_{s=1}^{S}\phi_s=1$). Therefore, ϕ_s can be considered the unconditional probability associated with observing the elimination-by-aspects behavioural rule characterised by class s (i.e. the prior likelihood of competing behavioural rules being their actual behaviour).

2.2.4 Combined Model

It is also possible combine the above three specifications that allow for opt-out effects using the following specification:

$$\Pr\left(\mathbf{y}_n|\mathbf{X}_n, \boldsymbol{\beta}, \lambda = 1, \gamma, \mu, \boldsymbol{\phi}\right) = \phi_{s=1} \prod_{z=1}^{Z_n} \left(\mathbb{I}_{niz} i_{nz}\right) + \phi_{s=2} \prod_{z=1}^{Z_n} \frac{\exp\left(\boldsymbol{\beta}\mathbf{x}_{niz}\right)}{\sum_{l=1}^{L} \exp\left(\boldsymbol{\beta}\mathbf{x}_{nlz}\right)} + \phi_{s=3} \prod_{z=1}^{Z_n} \left[\frac{\exp\left[\mathbb{I}_i\left(\boldsymbol{\beta}\mathbf{x}_{niz} + \gamma\right) + \left(1 - \mathbb{I}_i\right)\mu\Upsilon\right]}{\exp(\boldsymbol{\beta}\mathbf{x}_{n\mathbb{I}_j z} + \gamma) + \exp\left(\mu\Upsilon\right)} \Pr\left(c_{niz}|\mathbf{X}_{nz}, \boldsymbol{\beta}, \lambda = 1, \mu\right) \right]. \tag{6}$$

This specification includes an alternative-specific constant for the opt-out alternative, a hierarchical decision-making process and the elimination-by-aspects decision rules. We note that the model in Eq. (6) essentially nests the other models in the sense that (i) constraining $\gamma = 0$, $\mu = 1$ and $\phi_{s=3} = 1$ reduces the model to the basic multinomial logit expression in Eq. (2), (ii) relaxing only the restriction on γ is consistent with the multinomial logit model with an opt-out alternative-specific constant in Eq. (3), (iii) removing only the constraint on μ so that $0 \leq \mu < 1$ produces the nested logit model in Eq. (4d), and (iv) allowing $\forall \phi_s : 0 \leq \phi_s \leq 1$ leads to the independent availability logit model in Eq. (5a). Therefore, in the absence of information on which of the four specifications (i.e. Eqs. 2–5) is the true specification, the specification in Eq. (6) should be able to expediently explain the nature of the opt-out effects, albeit with the disadvantage of being less parsimonious. Moreover, starting with the full model that encompasses all opt-out effects enables backward selection through the sequential removal of the opt-out effect that is least supported by the data until only significant opt-out effects remain.

3 A Practical Demonstration

To demonstrate the above modelling approaches to accommodate some of the potential opt-out effects found in DCEs and the consequences of misspecifying them in the model on welfare and scenario analysis, we provide a practical demonstration using the R statistical software [25]. Full details and the codes necessary to replicate our results are given in Appendix A in the Electronic supplementary material (ESM). We use synthetic datasets generated using Monte Carlo experiments, which are especially useful because the true parameters underlying the data-generating process are known. This will allow us to judge model performance in terms of how close the model estimates are to the true values. For this demonstration, we construct a medication DCE characterised by three non-cost attributes (each of which has two levels) and a cost attribute (with four levels). The three non-cost attributes that describe the medication are efficacy, side effects and monitoring. Table 1 presents the attributes and the levels used in this example.

3.1 Data

For this practical demonstration, we generate three DCE datasets meeting different assumptions. The first DCE dataset assumes a choice behaviour that assumes everyone considered all alternatives. The model specification used in the data-generating process is based on the multinomial logit model with an alternative-specific constant that differentiates the opt-out alternative from non-opt-out alternatives. The second DCE dataset assumes a choice behaviour aligned with the nested logit specification. It assumes that the unobserved parts of the utility functions for non-opt-out alternatives are correlated within the same nest but uncorrelated with the opt-out nest. This specification also assumes that everyone considered all alternatives. In the third DCE

Table 1 Attributes and levels

Attribute	Levels (coding)
Efficacy	Worst level (0)[a]
	Best level (1)
Side effects	None (0)
	Some (1)[a]
Monitoring	No (0)[a]
	Yes (1)
Cost	€1 (1)[a]
	€2 (2)
	€3 (3)
	€4 (4)

[a]Baseline level

Table 2 Data-generating parameters for each treatment

DCE dataset	Opt-out alternative-specific constant (γ)	Degree of independence among non-opt-out alternatives (μ)	Consider opt-out alternative only ($\phi_{s=1}$)	Consider non-opt-out alternatives only ($\phi_{s=2}$)	Consider all alternatives ($\phi_{s=3}$)
1	0.3	1.0	0.0	0.0	1.0
2	0.0	0.5	0.0	0.0	1.0
3	0.0	1.0	0.3	0.2	0.5

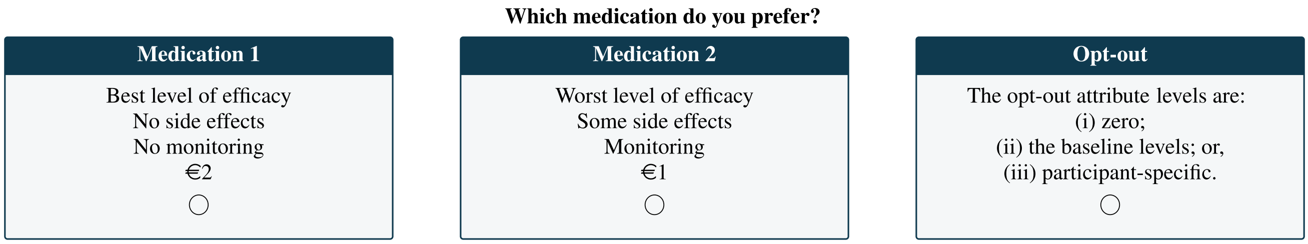

Fig. 2 Illustrative choice task

dataset, we assume that some participants ignored certain alternative(s) (which may be due to genuine preferences, serial non-participation or some other decision-making heuristic). More specifically, we assume that there were three groups of participants: (i) those who restricted their choice to the opt-out alternative, (ii) those who restricted their choice to the non-opt-out alternatives, and (iii) those who did not restrict their choice and considered all alternatives.

For the purposes of this application, for all DCE datasets, the true vector of marginal utilities for the four attributes is specified as $\boldsymbol{\beta} = \left[1.5\ -0.9\ 1.1\ -0.5\right]$. With the aim of generating the three DCE datasets, we choose different values for γ, μ and $\boldsymbol{\phi}$, as presented in Table 2. For our first DCE dataset, an alternative-specific constant is included, specifically $\gamma = 0.3$, while the values of μ and $\phi_{s=3}$ are both fixed at 1. The second DCE dataset assumes $\mu = 0.5$, so that there is covariance between the error terms of the non-opt-out alternatives, whereas $\gamma = 0$ and $\phi_{s=3} = 1$. The third DCE dataset is generated on the basis that some simulated participants reduced the number of alternatives they contemplated. We specify that 30% of participants serially non-participated by choosing the opt-out alternative, 20% chose only between the non-opt-out alternatives, and 50% considered all alternatives. The data generation process for this third dataset assumed $\gamma = 0$ and $\mu = 1$.

To demonstrate the effects of different opt-out representations, we consider three different definitions for attribute levels in the opt-out alternative: (i) all levels are set to zero, (ii) the attributes are set to their baseline levels, and (iii) a participant-specific pivot design is used, where, for the sake of illustration, the opt-out levels are randomly chosen for each participant. As a result, we have a total of nine cases that accommodate for different choice behaviour via three generated datasets for each of the three opt-out definitions. We refer to these nine cases as ‘treatments’ in the sense that each treatment is based on a different data-generating process.

For this demonstration, we make use of orthogonal main-effects experimental designs, and define the DCE as having two non-opt-out alternatives and an opt-out alternative itself.[3] An illustrative choice task is presented in Fig. 2. We use a sample of 350 participants who each completed a panel of eight choice tasks, producing 2800 choice observations for model estimation. Since idiosyncratic results can arise from a single sample of participants for each treatment, we generate multiple replications of the experimental design. In total, we generate $r = 1, 2, \ldots, R = 1000$ replications for the nine treatments.[4] The syntax to generate all datasets is given in Box A1 in Appendix A of the ESM.

3.2 Analysis

For every dataset generated under the three opt-out definitions, we estimate all four models described in Sect. 2.2. Doing so allows us to compare the opt-out effects under correctly specified and misspecified cases and to make inferences regarding the consequences for welfare analysis and choice prediction. The syntax for all candidate models is given in Box A2 in Appendix A of the ESM. All models were estimated using the R package maxLik [35]. The syntax for this process is given in Box A3, which is also provided in Appendix A of the ESM.

The four candidate models each produce a different insight into and interpretation of the opt-out effects. For this reason, a multimodel inference procedure is recommended so that judgements can be made regarding the relative suitability of each model accommodating the opt-out effects. Consequently, by regarding all four candidate models, we are in a better position to identify appropriate assumptions for addressing opt-out effects that are conditional on the data and the set of considered models. For further details

[3] While this design ensures that all attribute levels can be estimated independently of each other, we recognise that a more efficient experimental design could have been used to minimise the variance of the parameters. However, in a Monte Carlo experiment with specified parameters it may be more appropriate to show that the results stand up in cases where the experimental design is not tailored too closely to the data-generating parameters. Indeed, this would be the case in a real-life empirical application.

[4] This is sufficient for the purpose at hand since idiosyncratic simulation errors are not found to be large, as will be shown in Tables 3 and 4.

Table 3 Observed shares for each treatment (averaged over 1000 replications; standard deviations given in brackets)

	Breakdown by choices (2800 total choices)			Breakdown by consideration set alternatives (350 total participants)		
Treatment	Alternative 1	Alternative 2	Opt-out alternative	Only opt-out	Only non-opt-out	All
(a) None (all attributes in the opt-out alternative set to zero)						
T1	0.273 (0.008)	0.273 (0.007)	0.454 (0.009)	0.001 (0.001)	0.006 (0.004)	0.994 (0.004)
T2	0.299 (0.007)	0.299 (0.007)	0.402 (0.008)	0.000 (0.001)	0.013 (0.006)	0.987 (0.006)
T3	0.251 (0.011)	0.251 (0.012)	0.498 (0.020)	0.300 (0.025)	0.208 (0.021)	0.492 (0.027)
(b) Baseline (all attributes in the opt-out alternative set to the baseline level)						
T4	0.397 (0.008)	0.398 (0.008)	0.205 (0.007)	0.000 (0.000)	0.155 (0.019)	0.845 (0.019)
T5	0.413 (0.007)	0.413 (0.007)	0.174 (0.007)	0.000 (0.000)	0.214 (0.022)	0.786 (0.022)
T6	0.309 (0.013)	0.309 (0.013)	0.382 (0.022)	0.300 (0.025)	0.317 (0.025)	0.383 (0.026)
(c) Participant-specific (a pivot design where the opt-out alternative varies by participant)						
T7	0.304 (0.009)	0.304 (0.009)	0.392 (0.013)	0.011 (0.006)	0.092 (0.015)	0.897 (0.016)
T8	0.325 (0.009)	0.325 (0.009)	0.350 (0.013)	0.006 (0.004)	0.119 (0.018)	0.875 (0.018)
T9	0.264 (0.012)	0.264 (0.012)	0.472 (0.021)	0.303 (0.024)	0.264 (0.023)	0.433 (0.026)

The breakdown by choices is a summary of the proportion of choices for each alternatives, whereas the breakdown by consideration set alternatives is a summary of the simulated share of participants' choices that comply with each processing rule

Table 4 Weight of evidence (averaged over 1000 replications; standard deviations given in brackets)

Treatment	MNL with ASC	NL	IAL	Combined
(a) None (all attributes in the opt-out alternative set to zero)				
T1	0.822 (0.213)	0.175 (0.211)	0.003 (0.014)	0.000 (0.000)
T2	0.000 (0.000)	1.000 (0.001)	0.000 (0.000)	0.000 (0.001)
T3	0.000 (0.000)	0.000 (0.000)	0.998 (0.008)	0.002 (0.008)
(b) Baseline (all attributes in the opt-out alternative set to the baseline level)				
T4	0.960 (0.102)	0.039 (0.101)	0.000 (0.001)	0.000 (0.000)
T5	0.007 (0.058)	0.187 (0.322)	0.000 (0.000)	0.806 (0.324)
T6	0.000 (0.000)	0.000 (0.000)	0.998 (0.009)	0.002 (0.009)
(c) Participant-specific (a pivot design where the opt-out alternative varies by participant)				
T7	1.000 (0.002)	0.000 (0.002)	0.000 (0.000)	0.000 (0.000)
T8	0.000 (0.004)	0.998 (0.023)	0.000 (0.000)	0.002 (0.023)
T9	0.000 (0.000)	0.000 (0.000)	0.997 (0.028)	0.003 (0.028)

This table is a summary of the weight of evidence obtained for each candidate model by treatment, which can be interpreted as the probability that a given model is the best-approximating model (given the data and set of candidate models) in the respective treatment

on multimodel inference, see for example Buckland et al. [36] and Symonds and Moussalli [37], as well as Layton and Lee [38] and Campbell et al. [39] for their use in stated preference contexts.

As part of the multimodel inference procedure, we derive weights of evidence that each model correctly captures the choice behaviour assumed in the data-generation process of each dataset used. This can be accomplished by calculating the difference ($\Delta_{m_{tr}}$) between a penalised-likelihood information criterion (IC) value of the best model for treatment t and replication r and the equivalent value for the other models estimated in this treatment and replication:

$$\Delta_{m_{tr}} = \mathrm{IC}_{m_{tr}} - \mathrm{IC}_{\mathrm{min}_{tr}}, \tag{7a}$$

where $m = 1, 2, \ldots, M$, with M being the number of models (i.e. $M = 4$ in our case) and $\mathrm{IC}_{\mathrm{min}_{tr}}$ is the smallest value of $\mathrm{IC}_{m_{tr}}$ in the model set.[5] The term $\Delta_{m_{tr}}$ is a calibration of model fit, using the best-fitting model as the baseline. The best-fitting model has $\Delta_{m_{tr}} = 0$, and all other models have $\Delta_{m_{tr}} > 0$. Importantly, $\Delta_{m_{tr}}$ can be used to calculate a measure to assess the relative strength of each candidate model.

[5] In this paper, we use the Bayesian information criterion . We derive this for each estimated model m in treatment t and replication r as follows: $\mathrm{IC}_{m_{tr}} = \ln(N)K_{m_{tr}} - 2\ln\left(\hat{\mathcal{L}}_{m_{tr}}\right)$, where N is the number of choice observations, $\hat{\mathcal{L}}_{m_{tr}}$ is the maximised value of the likelihood function for model m in treatment t and replication r, and $K_{m_{tr}}$ is the number of estimated parameters associated with this model.

Specifically, a weight of evidence measure, $\omega_{m_{tr}}$, which is a probability scaling of $\Delta_{m_{tr}}$, can be derived using the following widely used expression:

$$\omega_{m_{tr}} = \frac{\exp\left(-0.5\Delta_{m_{tr}}\right)}{\sum_{m=1}^{M} \exp\left(-0.5\Delta_{m_{tr}}\right)}, \tag{7b}$$

where the sum is over all models in the set. The scaling is convenient as $0 < \omega_{m_{tr}} < 1$ and $\sum_{m=1}^{M} \omega_{m_{tr}} = 1$, meaning that they can be considered analogous to the probability that a given model in a given treatment and replication is the best-approximating model, given the data and set of candidate models.

In addition to the weight of evidence measure, we also consider the root-mean-square errors as indicators of estimation performance for the four candidate models per treatment. The root-mean-square error is a measure of the magnitude of the differences between the estimated parameters and the true parameters used in the data-generating process. It represents the standard deviation of the difference between predicted and actual values over the 1000 replications, thus giving a single measure of the predictive power for a parameter of interest for all candidate models. It is given by

$$\mathrm{RMSE}_{m_{t}} = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\hat{\tau}_{m_{tr}} - \tau_{t}\right)^{2}}, \tag{8}$$

where $\hat{\tau}_{m_{tr}}$ denotes the estimated value of a parameter of interest retrieved using model m in treatment t and replication r, and τ_t represents its true value in treatment t.

Our parameters of interest are the marginal willingness to pay estimates and choice predictions. Marginal willingness to pay estimates are calculated by dividing the parameters of the non-cost attributes by the negative of the parameter of the cost attribute. In addition to willingness to pay estimates, due to the different opt-out effects and the variation in the ability of each of our candidate models to explain these effects, there may be consequences for forecasting the demand for different medications. We therefore use the model estimates to simulate uptake for alternative medications. For this analysis, we consider the choice shares for the three medication profiles portrayed in Fig. 2.

4 Results

4.1 Observed Choice Shares and Consideration Set

Before continuing with the estimation results, we first report the observed choices and consideration sets for each dataset under each opt-out definition. As we generated three datasets under each opt-out definition, we have a total of nine treatments. The treatments T1, T4 and T7 are the three DCE datasets with different opt-out definitions generated from a multinomial logit model with an opt-out alternative-specific constant; T2, T5 and T8 are the respective treatments generated based on a nested logit specification; and T3, T6 and T9 are the treatments based on an independent availability logit specification that accommodates for alternative processing strategies for the three opt-out definitions. For each of the nine treatments, the mean and standard deviations of the observed share of choices for each alternative across the 1000 replications are reported in Table 3. We first remark on the low standard deviations, indicating that the observed proportions of choices tend to be close to their respective mean and so any idiosyncratic simulation errors are likely to be relatively small. As expected, we see no notable distinction between the proportions of alternative 1 versus alternative 2 choices regardless of the treatment. However, the shares for the opt-out alternative differ depending on the treatment. As can be observed, the setting that includes an opt-out alternative-specific constant is found to have a higher share of opt-out choices than the respective dataset based on the same opt-out definition but generated on the basis of a nested logit. This is expected, given that the opt-out constant is specified as being positive (recall $\gamma = 0.3$). Nevertheless, the observed share of optout choices is highest in the setting with explicitly defined consideration sets. This is due to the assumptions we use in our data-generating process: a higher proportion of participants are specified to consider only the opt-out alternative compared to those who only contemplate the non-opt-out alternatives (recall the respective shares of 30 and 20% for these behaviours).

Parameter settings aside, we find larger proportions of opt-out choices for the treatments where the attribute levels of the opt-out alternative are set to zero (i.e. treatments T1–T3). This is followed by the participant-specific (pivot) opt-out treatments (i.e. treatments T7–T9), and the treatments where the attribute levels are set to their baseline (i.e. treatments T4–T6). However, we, once more, emphasise that these differences are context specific and are driven by the data-generating assumptions.

Table 3 also reports the breakdown of participants with respect to the shares of alternatives included in their consideration set.[6] Irrespective of the definitions used for the opt-out alternative, settings in which the deterministic consideration set is assumed (i.e. all treatments aside from T3, T6 and T9) are, as expected, observed to have the highest share of participants who considered all three alternatives.

[6] As noted when describing the independent availability logit model in Sect. 2.2.3, the alternatives taken into account by a (real or simulated) participant cannot be established with certainty. For the sake of comparison, we assume an alternative is deemed to be not in a participant's consideration set if they never choose it in any of their eight choices.

Nonetheless, there are notable differences when we compare the shares across the opt-out definitions: a higher share of participants are identified as having the deterministic consideration set that includes all three alternatives when the opt-out attribute levels are set to the baseline (i.e. treatments T1 and T2) compared to where the levels are set to zero (i.e. treatments T4 and T5) and participant-specific values (i.e. treatments T7 and T8). We, again, accentuate that this is a consequence of the assumptions we use in data generation.

All else being equal, the most obvious difference in observed elimination-by-aspect behaviour and serial non-participation is in settings where these behaviours form part of the data-generating process. The dispersion of the observed serial non-participation behaviour is also low and very closely reflects the proportion used in generating the data. The importance of this result should not be understated, since it signals that a straightforward non-parametric comparison of the chosen alternatives can give a clear insight into the choices that make up the consideration set (e.g. whether there is any systematic choice behaviour relating to the opt-out alternative). Crucially, this can be used to inform the decision on the models to be included within the candidate set. We also find that the shares of participants identified as having exclusively considered the non-opt-out alternatives are reasonably consistent with the data-generation process. This finding is strongest for the first opt-out definition, which is, once more, an artefact of the data-generating process. For the second opt-out definition, where the non-cost attributes for the opt-out alternative are set at their inferior level, it is not surprising to find fewer simulated choices for the opt-out alternative. But, more importantly, this has given rise to the upwardly biased share of participants identified as having considered only between the non-opt-out alternatives in treatments T4 and T5. Therefore, while a simple comparison of chosen alternatives may be a useful starting point to garner some rudimentary intuition about the processing strategies, this highlights that its ability to do so depends on the true data-generating process and the empirical setting. Model estimation should give more definitive insight, which we now turn to.

4.2 Estimation Results

The repercussions of misspecified opt-out effects are best understood by assessing the ability of different modelling approaches to explain the true opt-out effects. The weights of evidence and the root-mean-square performance indicators will help in this regard.

4.2.1 Multimodel Inference

In Table 4, we report the mean of the weight of evidence measure, $\omega_{m_{\mathrm{tr}}}$, over the 1000 replications for each model specification and each definition. Recall that these values sum to 1 over the four models in the candidate set, and they can be interpreted as the probability of each model having the most appropriate specification to account for the opt-out effects (conditional on the set of models included in the candidate set). The larger the value, the more confidence we have that the model is the best-approximating model.

Our first observation is that, as expected, the most probable model in each treatment complies with the data-generating process. For instance, treatments generated on the basis of a multinomial logit with an opt-out alternative-specific constant are found to have the highest average weight of evidence for this model (i.e. 82, 96 and 100% for treatments T1, T4 and T7, respectively). Treatments T2 and T8, which are generated based on a nested logit, are also found to have the highest weight of evidence for the nested logit model. However, this is not the case for the opt-out definition using the baseline levels of attributes in treatment T5, which, related to an earlier observation, stems from the inferior levels representing the opt-out alternative (i.e. all else being equal, the tendency to choose a non-opt-out alternative is therefore higher). A further interesting result in Table 4 is that neither the multinomial logit with an opt-out alternative-specific constant specification nor the nested logit model perform well when participants do not consider all alternatives. In fact, the weights of evidence for these two models in their corresponding treatments are effectively zero. This would suggest that we could tentatively omit these models in cases where there is a strong belief (perhaps informed by a non-parametric comparison of chosen alternatives and/or by follow-up statements of information processing strategies) of serial non-participation or semi-compensatory choice behaviour.

A critical insight afforded by this multimodel inference is the fact that there are instances where the misspecified models also form part of the confidence set of models (i.e. the models that represent the majority of the evidence). This reinforces the need to evaluate a range of opt-out effects and to assess the quality of the predicted opt-out effect under each model relative to each of the other models. Therefore, using only a single model to base inferences on opt-out effects may not be recommended. This is especially the case in treatment T5, where the combined model is the best-approximating model. Nevertheless, it is important to be mindful that an incorrect model specification can lead to erroneous interpretation. However, as it is the most inclusive model for all choice behaviours assumed in the generated datasets, it offers the potential to backwardly eliminate the opt-out effects that are not supported by the data.

4.2.2 Estimation Performance

As we know the true parameters, we can assess how closely each of these candidate models predicts the true values. We use the root-mean-square error to measure the performance

Table 5 Performance indicator of estimation performance for marginal willingness to pay (root-mean-square error)

Treatment	Model	Efficacy	Side effects	Monitoring
(a) None (all attributes in the opt-out alternative set to zero)				
T1	MNL with ASC	0.225	0.189	0.212
	NL	0.371	0.163	0.412
	IAL	0.396	0.163	0.397
	Combined	0.224	0.188	0.220
T2	MNL with ASC	0.406	0.322	0.822
	NL	0.209	0.157	0.232
	IAL	0.197	0.269	0.228
	Combined	0.191	0.148	0.230
T3	MNL with ASC	0.530	0.315	0.850
	NL	0.706	0.168	0.858
	IAL	0.174	0.201	0.159
	Combined	0.266	0.202	0.271
(b) Baseline (all attributes in the opt-out alternative set to the baseline level)				
T4	MNL with ASC	0.217	0.167	0.210
	NL	0.441	0.279	0.444
	IAL	0.438	0.293	0.418
	Combined	0.218	0.165	0.223
T5	MNL with ASC	0.497	0.269	0.909
	NL	0.298	0.376	0.253
	IAL	0.309	0.596	0.149
	Combined	0.195	0.134	0.243
T6	MNL with ASC	0.644	0.470	1.097
	NL	0.328	0.231	0.341
	IAL	0.229	0.180	0.243
	Combined	0.263	0.196	0.275
(c) Participant-specific (a pivot design where the opt-out alternative varies by participant)				
T7	MNL with ASC	0.178	0.142	0.165
	NL	0.204	0.160	0.178
	IAL	0.179	0.149	0.174
	Combined	0.177	0.142	0.167
T8	MNL with ASC	0.284	0.166	0.581
	NL	0.178	0.128	0.314
	IAL	0.282	0.162	0.575
	Combined	0.217	0.136	0.364
T9	MNL with ASC	0.784	0.471	0.992
	NL	1.413	0.508	1.467
	IAL	0.234	0.179	0.224
	Combined	0.232	0.178	0.224

This table gives a comparison of a measure of the magnitude of the differences between the estimated values of marginal willingness to pay and the true values of marginal willingness to pay used in the data-generating process by candidate model and treatment

Table 6 Performance indicator of estimation performance for choice prediction (root-mean-square error)

Treatment	Model	Medication 1	Medication 2	Opt-out
(a) None (all attributes in the opt-out alternative set to zero)				
T1	MNL with ASC	0.024	0.017	0.014
	NL	0.045	0.026	0.025
	IAL	0.038	0.018	0.029
	Combined	0.026	0.019	0.014
T2	MNL with ASC	0.082	0.160	0.082
	NL	0.057	0.035	0.079
	IAL	0.036	0.148	0.118
	Combined	0.083	0.029	0.089
T3	MNL with ASC	0.071	0.064	0.024
	NL	0.070	0.048	0.029
	IAL	0.023	0.017	0.021
	Combined	0.027	0.020	0.022
(b) Baseline (all attributes in the opt-out alternative set to the baseline level)				
T4	MNL with ASC	0.026	0.021	0.008
	NL	0.028	0.026	0.015
	IAL	0.027	0.029	0.017
	Combined	0.028	0.025	0.008
T5	MNL with ASC	0.136	0.184	0.050
	NL	0.044	0.080	0.046
	IAL	0.149	0.217	0.069
	Combined	0.048	0.034	0.039
T6	MNL with ASC	0.095	0.101	0.025
	NL	0.036	0.027	0.032
	IAL	0.027	0.021	0.023
	Combined	0.029	0.023	0.023
(c) Participant-specific (a pivot design where the opt-out alternative varies by participant)				
T7	MNL with ASC	0.018	0.015	0.009
	NL	0.050	0.018	0.044
	IAL	0.046	0.018	0.051
	Combined	0.019	0.016	0.009
T8	MNL with ASC	0.078	0.134	0.058
	NL	0.036	0.069	0.038
	IAL	0.073	0.134	0.063
	Combined	0.031	0.068	0.046
T9	MNL with ASC	0.072	0.071	0.023
	NL	0.033	0.033	0.020
	IAL	0.022	0.016	0.021
	Combined	0.023	0.018	0.022

This gives a comparison of a measure of the magnitude of the differences between the predicted choice shares and the true choice shares used in the data-generating process by candidate model and treatment of the models estimating marginal willingness to pay estimates and choice predictions, which are, respectively, given in Tables 5 and 6.

Focusing first on the magnitudes of the errors in predictions of marginal willingness to pay in Table 5, we find some key differences between the four models in the candidate set. As anticipated, we observe that, on average, the model which aligns with the data-generating process provides the most precise estimates of marginal willingness to pay. Of especial importance, however, is the finding that the root-mean-square errors relating to the combined model are qualitatively similar to those produced by the true model specification. In fact, in many instances, the combined model produces lower root-mean-square errors than those of other models. Crucially, this indicates that the average performance of the combined model in terms of predicting marginal willingness to pay may even be superior to the average performance of the true model. This gives rise to a dilemma relating to parsimony: while (from Table 4) the combined model is generally less parsimonious, it produces relatively consistent (and, at times, more consistent) estimates of marginal willingness to pay. Non-parsimonious models obviously have the potential to overfit the data and lead to misguided judgements, but in this case, since the combined model effectively nests the other models, any loss of parsimony may be offset by the increased predictive performance.

A common goal of DCEs in health and other areas of applied economics is to predict demand and market share. For this reason, we consider the three medication scenarios presented in Fig. 2, and for every estimated model we retrieve the predicted choice share for each alternative. In Table 6, we present the root-mean-square errors of these predictions by treatment and model specification. The results show that the systematic errors in prediction vary by data-generating process, but, more importantly, also by model specification. Like what was observed for the marginal willingness to pay predictions, we find that, on average, misspecified models (with the exception of the combined model) produce less accurate predictions of the choice shares. We, again, find that the models which are consistent with the data-generating process generally produce the most precise choice forecasts on average. However, on average, those produced by the combined model specification are not found to be materially different. As a matter of fact, in many instances the combined model, on average, leads to more accurate predictions of the true choice shares. This further highlights the possible advantages of more flexible specifications, as they may offer potentially superior predictive power. However, it again raises the trade-off between the desire for parsimony and prediction accuracy as well as the distinction between explaining versus predicting.

5 Concluding Remarks

While the necessity to include an opt-out option in discrete choice experiments (DCEs) is contingent on the objectives of the study and is context dependent [5], its inclusion is, nevertheless, widespread practice, and is also recommended [3, 6–8]. It is well known and documented within the DCE literature that the propensity of participants to choose the opt-out alternative is often explained by many factors [11–13]. However, it is not always obvious how to accommodate it. This stems from the myriad of potential explanations for this phenomenon, and the inability to discern which (if any) of these potential explanations is at play. In this paper, we provide practical guidance on how to accommodate a range of opt-out effects in various formats and show the consequences of addressing these effects for marginal willingness to pay estimation and scenario predictions using simulated datasets with correctly specified and misspecified discrete choice models.

We focus on three common definitions for the opt-out alternative used in DCE studies. The first definition is the one where the attribute levels are set to zero (i.e. "none of them"). The second definition uses attributes that are set to their baseline level. The final definition uses participant-specific baseline levels (i.e. a pivot design). To account for the range of opt-out effects that might be present in DCE data, for each opt-out definition we generate three DCE datasets based on different assumptions using Monte Carlo simulation. The first dataset assumes an opt-out effect which is explained by an alternative-specific constant. The second dataset assumes correlation between the non-opt-out alternatives, which can be accommodated via a nested logit specification. The third dataset assumes a share of participants eliminate the opt-out alternative or non-opt-out alternatives from their consideration set. We estimate three models corresponding to these datasets for each opt-out definition, and a combined model that essentially nests the specifications considered in the first three models. We subsequently conduct a frequentist-based multimodel inference approach using information criteria to derive weights of evidence for each model. To help judge the performance of estimations in terms of predicting the true values, we calculate root-mean-square errors for both welfare estimates and choice predictions. We present our code written in the R statistics program to encourage others to explore these opt-out effects in their own data.

Our findings in this practical demonstration suggest a number of key recommendations for DCE practitioners interested in exploring opt-out effects. Before any model estimation, we suggest a straightforward exploration of the choice shares to determine if a subset of participants consistently (or predominately) chose the opt-out alternative or the non-opt-out alternatives. Based on what is observed, it may be possible to refine the set of candidate models. Indeed, models should be linked with very specific and testable hypotheses and/or be driven by prior beliefs of what is driving the opt-out effect(s) in the context studied. In the context of opt-out effects, discrete choice models should, as far as possible, distinguish between situations where participants have made genuine opt-out choices and where they exhibit behaviour beyond the theory of rational choice. Indeed, there may be potential confounding between different opt-out effects. For this reason, there may be advantages in choosing econometric specifications that can simultaneously explain multiple opt-out effects, and, if necessary, backwardly eliminate the opt-out effects that are not supported by the data. Focusing solely on one opt-out effect may explain only part of the story and, crucially, could lead to biased inferences regarding the opt-out effects. Yet, there is the need to be mindful of the proliferation of parameters and, therefore, loss of parsimony. While a more comprehensive model will ensure the DCE data is fitted well, it comes at the risk of being tailored too closely to the sample data, which compromises the ability to generalise the model beyond the existing dataset. This is especially important when the aim is to use the DCE results to derive some aggregation measure, such as the average marginal willingness to pay within a population or the average change in demand in response to a change in a medication attribute. While parsimony is an important consideration, it may not always result in the best predictive performance (for both marginal willingness to pay and choice prediction). For this reason, there may be other considerations aside from parsimony. Given the range of possible opt-out effects and the fact that these are likely to be heterogeneous across the population, it might be unrealistic to expect accurate predictions from simpler (parsimonious) models. There is, of course, a distinction between estimation and prediction, and there may be a need to consider different modelling assumptions for each. But additional factors to consider when deciding on the model specification include plausibility, consistency with established opt-out effects and behavioral phenomena, and in most practical settings there is a need to ensure that the model results are understandable to a non-technical entity. Nevertheless, due to the inability to know why participants chose, or did not choose, opt-out alternatives in DCEs, it is difficult to know which model specification(s) to use. Not surprisingly, many different model specifications have been offered in the DCE literature. However, we suggest that there is no unique best way to analyse DCE data. We do not wish to give the impression that any of the models used in this paper will offer the best solution for every DCE dataset. Indeed, a familiar aphorism among econometricians is that "all models are wrong". For this reason, we encourage researchers to consider several models and to subsequently apply multimodel inference and embrace model averaging. By doing so, different hypotheses of opt-out effects can be explained by several models, which can then be ranked and weighted to provide a quantitative measure of the relative support for each competing hypothesis.

Regarding the modelling frameworks presented here, for illustrative purposes, we have considered only four models. Of course, there is scope for further specifications and formats, such as the dual response format, where participants are first presented with a forced choice and then asked a follow-up question on whether or not they would choose the option or opt out (e.g. see Brazell et al. [23] and Schlereth and Skiera [13]). We also acknowledge that our results are based on Monte Carlo experiments and, due to the assumptions made when generating the datasets, they may not apply in all contexts. However, we find qualitatively similar results in a range of different settings. Indeed, we encourage others to utilise the R code to investigate the sensitivity of our results to sample size and experimental design properties as well as the number of alternatives, attributes and choice tasks. But, more importantly, we hope that others make use of the code to explore opt-out effects present in their own data, which will contribute towards a better understanding of opt-out effects in DCEs.

Acknowledgements We thank the editor Christopher Carswell for his invitation to write this paper. We also thank four anonymous reviewers for their helpful comments and suggestions on previous versions of this paper. Any remaining errors or misinterpretations are solely the authors' responsibility.

Author contributions DC and SE contributed equally to all aspects of this paper, including the conceptualisation, data generation, analysis and drafting of the manuscript.

Data availability statement For this paper, the data have been synthetically generated. Full details on the data-generating process and the code required to replicate our analysis are given in Appendix A of the ESM.

Compliance with ethical standards

Funding The study was not supported by any external sources or funds.

Ethical approval The study did not involve the collection of primary data or the use of secondary data sources, thus negating the need for ethical approval.

Informed consent Participants have been artifically generated as part of the Monte Carlo simulation, meaning that informed consent is not applicable.

Conflict of interest Danny Campbell and Seda Erdem declare no conflicts of interest relevant to the content of this manuscript.

References

1. Craig BM, Lancsar E, Mühlbacher AC, Brown DS, Ostermann J. Health preference research: an overview. Patient Patient Cent Outcomes Res. 2017;10(4):507–10.
2. Ryan M, Skåtun C. Modelling non-demanders in choice experiments. Health Econ. 2004;13(4):397–402.
3. Lancsar E, Louviere JJ. Conducting discrete choice experiments to inform healthcare decision making. Pharmacoeconomics. 2008;26(8):661–77.
4. Boxall P, Adamowicz WL, Moon A. Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. Aust J Agric Resour Econ. 2009;53(4):503–19.
5. Veldwijk J, Lambooij MS, de Bekker-Grob EW, Smit SA, de Wit DA. The effect of including an opt-out option in discrete choice experiments. PLoS ONE. 2014;9(11):e111805.
6. Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. Health Econ Policy Law. 2009;4(4):527–46.
7. Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, Johnson FR, Mauskopf J. Conjoint analysis applications in health—a checklist: a report of the ISPOR good research practices for conjoint analysis task force. Value Health. 2011;14(4):403–13.
8. Johnston RJ, Boyle KJ, Adamowicz W, Bennett J, Brouwer R, Cameron TA, Hanemann WM, Hanley N, Ryan M, Scarpa R, Tourangeau R, Vossler C. Contemporary guidance for stated preference studies. J Assoc Environ Resour Econ. 2017;4(2):319–405.
9. Niebor A, Xander K, Elly S. Preferences for long-term care services: willingness to pay estimates derived from a discrete choice experiment. Soc Sci Med. 2010;70(9):1317–25.
10. Milte R, Ratcliffe J, Miller M, Whitehead C, Cameron I, Crotty M. What are frail older people prepared to endure to achieve improved mobility following hip fracture? A discrete choice experiment. J Rehabil Med. 2013;45(1):81–6.
11. Dhar R, Simonson I. The effect of forced choice on choice. J Market Res. 2003;40(2):146–60.
12. Bahamonde-Birke FJ, Navarro I, de Dios Ortúzar J. If you choose not to decide, you still have made a choice. J Choice Model. 2017;22:13–23.
13. Schlereth C, Skiera B. Two new features in discrete choice experiments to improve willingness-to-pay estimation that result in SDR and SADR: separated (adaptive) dual response. Manag Sci. 2017;63(3):829–42.
14. Salkeld G, Ryan M, Short L. The veil of experience: do consumers prefer what they know best? Health Econ. 2000;9(3):267–70.
15. Ryan M, Ubach C. Testing for an experience endowment effect in health care. Appl Econ Lett. 2003;10(7):407–10.
16. Meyerhoff J, Liebe U. Status quo effect in choice experiments: empirical evidence on attitudes and choice task complexity. Land Econ. 2009;85(3):515–28.
17. Oehlmann M, Meyerhoff J, Mariel P, Weller P. Uncovering context-induced status quo effects in choice experiments. J Environ Econ Manag. 2017;81:59–73.
18. Kahneman D, Knetsch JL, Thaler RH. Anomalies: the endowment effect, loss aversion, and status quo bias. J Econ Perspect. 1991;5(1):193–206.
19. Krosnick JA, Holbrook AL, Berent MK, Carson RT, Hanemann WM, Kopp RJ, Mitchell RM, Presser S, Ruud PA, Smith VK, Moody WR, Green MC, Conaway M. The impact of "no opinion" response options on data quality: non-attitude reduction or an invitation to satisfice? Public Opinion Q. 2002;66(3):371–403.
20. Tversky A, Shafir E. Choice under conflict: the dynamics of deferred decision. Psychol Sci. 1992;3(6):358–61.
21. Baron J, Ritov I. Reference points and omission bias. Organ Behav Hum Decis Process. 1994;59(3):475–98.
22. Masatlioglu Y, Ok EA. Rational choice with status quo bias. J Econ Theory. 2005;121(1):1–29.
23. Brazell JD, Diener CG, Karniouchina E, Moore WL, Séverin V, Uldry PF. The no-choice option and dual response choice designs. Market Lett. 2006;17(4):255–68.
24. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertain. 1988;1(1):7–59.
25. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016. https://www.R-project.org/.
26. Scarpa R, Ferrini S, Willis K. Performance of error component models for status-quo effects in choice experiments. In: Scarpa R, Alberini A, editors. Applications of simulation methods in environmental and resource economics. Dordrecht: Springer; 2005. p. 243–73.
27. Train K. Discrete choice methods with simulation. 2nd ed. Cambridge: Cambridge University Press; 2009.
28. von Haefen RH, Massey RH, Adamowicz WL. Serial nonparticipation in repeated discrete choice models. Am J Agric Econ. 2005;87(4):1061–76.
29. Manski CF. The structure of random utility models. Theory Decis. 1977;8:229–54.
30. Frejinger E, Bierlaire M, Ben-Akiva M. Sampling of alternatives for route choice modeling. Transp Res Part B Methodol. 2009;43(10):984–94.
31. Campbell D, Hensher DA, Scarpa R. Cost thresholds, cut-offs and sensitivities in stated choice analysis: identification and implications. Resour Energy Econ. 2012;34(3):396–411.
32. Kaplan S, Shiftan Y, Bekhor S. Development and estimation of a semi-compensatory model with a flexible error structure. Transp Res Part B Methodol. 2012;46(2):291–304.
33. Campbell D, Erdem S. Position bias in best-worst scaling surveys: a case study on trust in institutions. Am J Agric Econ. 2015;97(2):526–45.
34. Erdem S, Campbell D, Thompson C. Elimination and selection by aspects in health choice experiments: prioritising health service innovations. J Health Econ. 2014;38:10–22.
35. Henningsen A, Toomet O. Maxlik: a package for maximum likelihood estimation in R. Comput Stat. 2011;26(3):443–58.
36. Buckland ST, Burnham KP, Augustin NH. Model selection: an integral part of inference. Biometrics. 1997;53(2):603–18.
37. Symonds MRE, Moussalli A. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol. 2011;65(1):13–21.
38. Layton DF, Lee ST. Embracing model uncertainty: strategies for response pooling and model averaging. Environ Resour Econ. 2006;34(1):51–85.
39. Campbell D, Mørkbak MR, Olsen SB. The link between response time and preference, variance and processing heterogeneity in stated choice experiments. J Environ Econ Manag. 2018;88(1):18–34.
40. Wuertz D et al. fExtremes: Rmetrics—extreme financial market data. 2013. R package version 3010.81. https://CRAN.R-project.org/package=fExtremes.
41. Aizaki H. Basic functions for supporting an implementation of choice experiments in R. J Stat Softw. 2012;50(2):1–24.