



# Sample-based longitudinal discrete choice experiments: preferences for electric vehicles over time

Katharina Keller<sup>1</sup> · Christian Schlereth<sup>2</sup> · Oliver Hinz<sup>1</sup>

Received: 4 August 2019 / Accepted: 15 November 2020 / Published online: 19 February 2021  
© The Author(s) 2021, corrected publication 2021

## Abstract

Discrete choice experiments have emerged as the state-of-the-art method for measuring preferences, but they are mostly used in cross-sectional studies. In seeking to make them applicable for longitudinal studies, our study addresses two common challenges: working with different respondents and handling altering attributes. We propose a sample-based longitudinal discrete choice experiment in combination with a covariate-extended hierarchical Bayes logit estimator that allows one to test the statistical significance of changes. We showcase this method's use in studies about preferences for electric vehicles over six years and empirically observe that preferences develop in an unpredictable, non-monotonous way. We also find that inspecting only the absolute differences in preferences between samples may result in misleading inferences. Moreover, surveying a new sample produced similar results as asking the same sample of respondents over time. Finally, we experimentally test how adding or removing an attribute affects preferences for the other attributes.

**Keywords** Adoption · Electric vehicles · Complementary mobility services · Discrete choice experiment · Dual response · Sample-based longitudinal study

## Introduction

At present in the year 2020, the fictive manufacturer ACME is working hard to ensure that one of its recent innovations will become a mainstream technology in the foreseeable future. As the management at ACME know, innovations do not break through overnight. Instead, they pass through a long process until they become widely adopted (Rogers 1962). For example, it took 40 years for the TV to reach a market penetration of 50 million owners. The e-bike, which was developed before

1900, entered the mass market with almost 1 million units sold in 2018 (Statista), enduring peaks and troughs in between. Even the iPhone needed more than three years after the market launch to sell more than 50 million units.

One decisive factor for an innovation's success is whether manufacturers have carefully aligned these innovations with consumer preferences (Reinders et al. 2010). To measure preferences, practitioners have come to rely on discrete choice experiments (e.g., Gensler et al. 2012; Louviere et al. 2000; Papies et al. 2011; Schlereth and Skiera 2017). However, such tests have mostly been used cross-sectionally (i.e., at just one point in time), which makes it impossible to derive conclusions about changes in preferences over time. Only a few discrete choice experiments have tackled preference measurement in a longitudinal setting (Ambos et al. 2019; Jensen et al. 2014; Meeran et al. 2017). However, all of them retained the same set of attributes across the duration of their studies and, with the exception of Ambos et al. (2019), they also surveyed the same samples of respondents, which might explain the low number of longitudinal studies.

Discrete choice experiments face two challenges in longitudinal studies. First, it is often impossible to survey the same sample respondents over time, particularly over several years: some respondents may have changed their (email) address,

---

V. Kumar served as Area Editor for this article.

✉ Katharina Keller  
kakeller@wiwi.uni-frankfurt.de

Christian Schlereth  
christian.schlereth@whu.edu

Oliver Hinz  
ohinz@wiwi.unifrankfurt.de

<sup>1</sup> Goethe University Frankfurt, Theodor-W.-Adorno Platz 4, 60323 Frankfurt am Main, Germany

<sup>2</sup> WHU – Otto Beisheim School of Management, Burgplatz 2, 56179 Vallendar, Germany

lost their willingness to participate, classified the emails as spam, or just entered a different life situation. Meanwhile, a representative sample at one point in time might not be representative after a few years: all subjects in the sample become older and some may no longer be available. Second, innovative technologies evolve, with some attributes (e.g., price or certain capabilities) changing between two or more consecutive studies. For example, the iPhone originally had a screen size of 3.5" in its first four years. It has since then nearly doubled in size—a development that Apple considered impracticable at the beginning. Firms might also introduce new features in later stages of the innovation process, such as in the case of NFC technology, which has become a prerequisite for mobile payment. Naturally, researchers may want to adapt later studies to evolving market needs without losing the ability to compare respective preferences.

To overcome these challenges, this paper proposes a sample-based longitudinal discrete choice experiment. To this end, we developed a covariate-extended hierarchical Bayes logit estimator that combines the estimation of multiple discrete choice experiments within a single estimation. Thereby, we make the following main contributions to the literature: First, despite surveying different respondents, our approach enables researchers to estimate not only the parameter values that represent individual respondents' preferences, but also which preferences have changed over time and to what extent. Second, we formally outline how to interpret the results when continuous attribute levels, such as price, have changed over time, new attributes have been added, or attributes from earlier studies have been removed. These two main contributions provide a new level of flexibility for researchers and practitioners to employ discrete choice experiments. Many companies might have ready access to this kind of data from their past applications of discrete choice experiments. For them, our approach reduces the requirements of linking together already collected data to identify trends.

We showcase the utility of sample-based longitudinal discrete choice experiments by eliciting preferences for electric vehicles over more than six years. Electric vehicles are an interesting focal point because they have yet to achieve mainstream adoption. Besides highlighting the use of our proposed method, we examine a range of associated questions: First, we experimentally test how adding an attribute affects preferences for all other attributes. We find that only a subset of attributes proportionally lose importance in the face of a newly added attribute, and some attributes are statistically unaffected. Second, we experimentally compare the results of surveying a new sample versus asking the same respondents over time. The general purchase intention remained mostly the same. Lastly, we demonstrate how to handle changes in continuous attribute levels. We also compare our model's internal and predictive validity against a model that separately estimates preferences for each year and a model that assumes a single normal distribution for the population of all respondents.

From a theoretical perspective, we demonstrate that an innovation process does not necessarily follow a monotonously increasing pattern; it might fluctuate through peaks and troughs, thereby underscoring the benefits of using a sample-based longitudinal study. From a managerial perspective, we conclude by performing a suite of counterfactual exercises that provide managerial guidance on an essential pricing decision problem.

In the next section, we provide a literature overview of the related streams of research in discrete choice experiments and models describing the evolution of innovation processes. Then, we formally describe how to reflect changes to the list of attributes and attribute levels over time. Finally, we outline the study context on electric vehicles, the setup, and the results. We conclude with managerial implications, conclusions, and limitations.

## Related research

In the following, we establish a theoretical framework for embedding our new modeling approach. We start by summarizing two streams of research related to longitudinal studies in the context of innovations: One stream on the behavioral and analytical aspects of discrete choice experiments, and the other on product innovation growth.

### Related research on discrete choice experiments

Discrete choice experiments ask respondents to make hypothetical choices between multiple sets of alternatives. They are backed up by a long-standing theory (e.g., Louviere et al. 2000) and parallel real-world purchase decisions. These experiments can satisfactorily explain actual purchasing behavior, even when the product or service of interest does not exist or was recently launched. Table 1 summarizes a non-exhaustive list of related research, mostly on discrete choice experiments that inspired our study; we will cover these in more detail below. For a broader overview of recent accomplishments, we refer to, e.g., Agarwal et al. (2015).

**Longitudinal use of discrete choice experiments** We are aware of few longitudinal studies that use discrete choice experiments. One is Jensen et al.'s (2014) study, which asked respondents to decide between electric vehicles and conventional cars. In the first round, the researchers surveyed respondents without any experience; then, the respondents had the opportunity to use an electric vehicle for three months and were surveyed again afterward with the same questionnaire. The researchers compared the choice proportions of choosing an electric vehicle before and after respondents' experience in real life and found that after the trial phase, respondents chose the electric vehicle only half as often as before. Yet, their

**Table 1** Non-exhaustive overview of related research

Study	Study size	Longitudinal study	Modelling issues	Key findings
<b>Longitudinal Use of Discrete Choice Experiments</b>				
Jensen et al. (2014)	196 respondents in Denmark	Yes, Time frame: 3 months same attributes & levels	Preferences for electric vehicles and conventional vehicles for respondents without experience and after a trial period, CBC	Respondents chose electric vehicles after trial phase only half as often as before
Meeran et al. (2017)	161 respondents	Yes Time frame: 6 months same attributes & levels	Change in consumer preferences in four different product categories, CBC	For products with rapid technological change and short life cycles, consumer preferences are more likely to shift
Ambos et al. (2019)	238 managers from Germany, Austria, and Switzerland	Yes Time frame: 4 years same attributes & levels	Management decisions regarding international locations over time, CBC	Managerial preferences vary over time, but heuristics for decision-making remain stable
<b>Theory of Unseen Attributes</b>				
Bradlow et al. (2004)	130 undergraduate students	No	Examine how respondents assume missing attribute levels, missing attributes were shown to participants beforehand, rating based conjoint analysis	Consumers deduce missing attribute levels from missing and non-missing attribute information
Yang et al. (2015)	70 respondents	No	Joint modeling of information processing (eye movements) and conjoint analysis	Improves predictions outside the sample and differentiation between characteristics, allows shorter questionnaires
Gilbride et al. (2016)	895 shoppers	No	Modeling of information processing of an information display board at the point-of-purchase in actual purchase situations, choice model	Inferring ex-ante value of attribute level dominates
<b>Covariate-extended Hierarchical Bayes Modelling and Combining Multiple Data Sources</b>				
McCullough and Best (1979)	100 undergraduate students	Yes Time frame: 3 days One attribute exchanged	Test/retest, altering attribute configuration between studies, CBC	High temporal stability and structural reliability, despite changes in the list of attributes
Teas (1985)	148 respondents	Yes Time frame: 1 week One attribute deleted	Within-subject test/retest, 1 control – 4 test groups, containing interaction terms, excluding attributes, OLS regression analysis using effects coding, CBC	Repetition of earlier findings: high temporal stability and structural reliability, negligible interaction effects
Lenk et al. (1996)	Study 1: 100 synthetic respondents, study 2: 179 students	No	Investigating shorter questionnaires for complex products/services, conjoint analysis	Response quality is unchanged by questionnaire length
Swait and Andrews (2003)	Scanner Panel Data: 526 panelists, choice experiment: 384 Chicago residents	No	Combining real data (scanner panel choice data) and experimental data (choice experiments)	Joint model shows better prediction performance on holdout data set compared to pure scanner panel model
Orme and Howell (2009)	326 respondents	No	Investigation of the influence of covariates in CBC/HB estimates	Usage of too many covariates in the model should be avoided
Kurz and Binmer (2010)	10 studies, 30,000+interviews	No	Investigating the added value of covariates in HB Modeling based on CBC data sets	Mostly, covariates did not improve results
Sentis and Geller (2010)	5 studies, 8,445 interviews	No	Investigating the impact of covariates in HB estimates within CBC data sets	Fit and predictive accuracy unchanged
Liakhovitski and Shmulyan (2011)	160 synthetic respondents	No	Investigating improvements of CBC HB estimate by using covariates	No significant improvement in the hit rate by using covariates, but partly better distributed utilities +aggregated metrics
Ellickson et al. (2019)	510 survey respondents, actual purchase data from 4,288 customers	No	Hybrid approach: fusion of SP (CBC) and RP (actual purchase) data	The approach increases the accuracy of forecasts for subsequent actual purchases and overcome selection and contextual bias

perceptions about the electric vehicles' driving performance had improved.

Meeran et al. (2017) measured consumers' changing preferences across four different product categories over six months. They conducted the same survey three times, questioning the same respondents and estimating the separate and joint aggregate parameters using Maximum Likelihood Techniques. Like in Jensen et al. (2014), the authors used the same attributes and their levels between studies. Meeran et al. (2017) found that consumer preferences are more likely to shift for products that experience rapid technological change and short lifecycles.

In another longitudinal discrete choice experiment, Ambos et al. (2019) studied management decisions regarding international locations at two points in time. The authors used hierarchical Bayes to separately estimate the preferences and found that they varied over time, although the decision-making heuristics remained stable. Like in Jensen et al. (2014) and Meeran et al. (2017), the authors used the same attributes and levels across studies.

In sum, all these studies were conducted within a relatively short time (three months, six months, four years), so preferences may not have changed that much. Observing greater changes may require a longer study period. This short time-span may be due to their inherent challenges: ensuring answers from the same respondents and keeping the focal product the same. Concerning the former, it is unclear whether a study benefits from surveying the same sample of respondents. Certain biases might arise, such as the mere exposure effect (Zajonc 1968), whereby respondents assess an originally neutral product or service more positively when repeatedly asked about it. Concerning the second challenge, the shifting nature of innovation processes makes it difficult to measure preferences for the same set of attributes over a long period. With our proposed sample-based longitudinal discrete choice experiment, we aim to overcome these challenges.

**Capturing the longitudinal nature of the data** Comparing the preferences of different samples of respondents over time requires the joint estimation of multiple discrete choice experiments. We propose a covariate-extension of the generic hierarchical Bayes logit estimator, which explicitly allows the sample preferences to differ. We base our implementation on Lenk et al.'s (1996) work concerning a more flexible definition of the upper-layer model in hierarchical Bayes. Whereas the generic hierarchical Bayes model assumes a single multivariate normal distribution for all respondents, we explicitly allow for multiple distributions that depend on when the data collection occurred.

Often, we think of covariates such as common demographics like gender or age; however, according to Orme and Howell (2009), these variables have a low correlation with preferences within choice contexts. The use of covariates

is more meaningful when researchers are concerned that the traditional shrinkage of individual parameters toward the population mean goes against their expectation of different, distinguishable segments. Another reason for their use is when a segment of respondents is oversampled: Without including covariates, the shrinkage toward the population means can bias the estimates of the undersampled group.

Researchers who apply covariate modeling emphasize that improving predictions is not the main motivation. Several studies have independently found that improvements in predictive validity are rather modest (if any) in comparison to the generic hierarchical Bayes modeling with the single distribution assumption (Kurz and Binner 2010; Liakhovitski and Shmulyian 2011; Orme and Howell 2009; Sentis and Geller 2010). Nevertheless, covariates are meaningful because they “allow us to test more formally the differences between segments and the part-worth” (Orme and Howell 2009). Thus, understanding and explicitly capturing developments in preferences in longitudinal settings is the primary benefit of covariates for our study.

Besides using a model that assumes homogeneity on the population level, previous studies applied two other techniques in the case of multiple data sources. The first is to estimate each study separately (like in, e.g., Ambos et al. (2019)); however, changes in parameter values might not always be apparent, especially for altered lists of attributes over time.

The stream on data enrichment, which jointly estimates scanner data with discrete choice experiments, provides a second technique (c.f., Louviere et al. 2000, chapter 8). This stream combines two data sources with only a partial overlap in the attributes (e.g., Swait and Andrews 2003; Ellickson et al. 2019) by distinguishing two sets of attributes: a common set across both data sources and an uncommon one. Ellickson et al. (2019) incorporated a mean additive preference shifter and a multiplicative variance shifter for common attributes, assuming that both data sources share the same covariance matrix. For the uncommon attributes, they separately captured preferences with separate covariance matrices. Nevertheless, their study context differs from ours because they always fused exactly two data sources; generalizing this technique to more than two can quickly become challenging, as the number of distinguishable sets of attributes rises exponentially. Some elements, such as the mean additive shifter, served as inspiration for our sample-based longitudinal discrete choice experiment.

In the empirical study, we will compare the performance of the covariate-extended hierarchical Bayes logit estimator against the performance of the generic and the separate estimation. More importantly, we will outline how the covariate-extended model's estimation output (i.e., the posterior) directly provides a simple, but effective way of testing whether changes in parameters across longitudinal samples are significant or just noise.

**Theory of unseen attributes** Besides asking different samples of respondents, longitudinal studies for innovations require solutions to address changes in market-relevant attributes. In this respect, we rely on the theory of unseen attributes to account for situations where attributes are added or removed, such that they are “unseen” in some of the studies. A prerequisite is that preferences for attributes are structurally reliable, even though the set of attributes does not overlap completely across studies (see also McCullough and Best 1979; Teas 1985). This assumption complies with common theory in design generation techniques (Street and Burgess 2007), i.e., that a rigorously constructed design improves the quality of the parameter estimates (i.e., the standard deviations of the beta parameters), but it does not affect the location of the beta parameters (i.e., their means).

Based on the behavioral literature on unseen attributes, we identified three lines of arguments on how respondents process them (c.f., Bradlow et al. 2004; Gilbride et al. 2016): The first argument assumes a “null effect” for missing attributes, i.e., that respondents exclusively focus on the information that is directly available in the choice tasks. This line of argument builds on bounded rationality theory and assumes that respondents tend to ignore information that is not directly available (Bettman et al. 1998).

The second argument is that respondents substitute “market means” for missing attributes in the choice sets. Used in cases like Yang et al. (2015), it builds on experimental findings wherein respondents treated unseen attributes as if they had the average perceived utility in the market. Branco et al. (2012) broadened the assumption of one commonly known market average and proposed that each respondent inferred an “expected individual value” that can differ across respondents, but they did not specify the process of setting the value.

The third argument fills the lack of an explicitly specified process by assuming that respondents impute expected individual values through a “pattern-matching learning model” (Bradlow et al. 2004). This means that respondents infer values based on seen information; for example, they conclude that a certain, yet unseen, attribute is available in case of a higher price. Using partial profile rating-based conjoint, Bradlow et al. (2004) showed that their model featured better internal and predictive validity than the “null effect” model. However, unlike in their study, where respondents were aware that some of the information was missing, ours does not employ partial profiles. For this reason, we will concentrate our modeling on the first two assumptions and show that they eventually result in the same model.

## Related research on product innovation growth

**Growth models** Another stream of research that complements the analysis of preference evolution over time deals with product innovation growth. We distinguish between two growth

processes: The first one is the Gartner hype cycle, an annually updated model by the American research and advisory firm Gartner, Inc (Linden and Fenn 2003), which assesses the commercial viability of an emerging technology. It distinguishes five phases, starting with the *innovation trigger*. The subsequent phase, the *peak of inflated expectations*, is followed by a negative hype caused by decreasing interest (the *trough of disillusionment*), as consumers come to understand that the innovation will not fulfill all expectations. Companies can overcome this low-point phase by developing second-generation products that make the innovation’s utility more visible and understandable. These efforts lead to the final phases: *slope of enlightenment* and *plateau of productivity*. The innovation begins to achieve wide acceptance from the masses, and mainstream adoption begins. The theoretical basis of this non-monotonous concept is an over-expectation of the technology’s potential, coupled with excessive confidence in its functionality on the part of developers, stakeholders, and early adopters. Several factors cause these initial perceptions, such as the sheer optimism of stakeholders about promising, but not yet proven, abilities, as well as the intended (exaggerated) publicity as a means of facilitating rapid adoption (Lucker et al. 2018). The Gartner hype cycle’s third phase, *trough of disillusionment*, aligns with the *expectation disconfirmation theory*, according to which individual expectations interact with objective criteria to determine people’s product satisfaction (Bhattacharjee 2001; Oliver 1980).

Market diffusion models provide the second representation of growth processes. Fourt and Woodlock (1960), for example, assumed that the cumulative sales curve of an innovation follows an exponential, i.e., monotonously increasing shape. Rogers (1962) and other researchers predicted a bell-shaped frequency curve, with a steep rise in the early adopter segment and a decline after reaching a peak due to market saturation. The corresponding cumulative curve is S-shaped, again, monotonously increasing, but with a low gradient at the beginning as well as the end and a higher gradient in the middle.

Market diffusion models and models like the Gartner hype cycle can complement each other and are not necessarily in conflict. In contrast to the market diffusion models, the Gartner hype cycle captures the evolution of expectations rather than sales. Thus, its conceptual foundation aligns more with preferences: Preferences translate into sales when the innovation reaches the slope of enlightenment or the plateau of productivity. Hence, the Gartner hype cycle’s two final stages reflect the beginning of the market diffusion models.

**Reasons for changes in preferences over time** Product assessments—and the accompanying consumer acceptance of an innovation—are largely influenced by consumer preferences. Thus, it is important to understand why these preferences may change over time. One reason is that consumers face a trade-off between the monetary costs of a new

technology and its perceived benefits (Dodds et al. 1991). If the ratio between value and price is high, customers’ usage intention and purchase intention will increase (Venkatesh et al. 2012). This ratio improves with decreasing prices, as is common for technological innovations. The technology also becomes more attractive if an innovation accumulates more benefits over time. If both occur, the cost–benefit ratio greatly improves.

Another reason is that unfamiliarity with new products, lacking experience, and limited information may influence consumer preferences over time (Coupey et al. 1998). According to the bounded rationality theory, which acknowledges that people have limited capacity for processing information when making decisions (Simon 1955), unfamiliarity with certain attributes may result in situations where consumers only consider a subset of attributes. The more that customers learn about those products, the more they enhance their experience; over time, they may change the subset of attributes used for decision-making, and assign different levels of importance to the attributes (Meeran et al. 2017). Also, the certainty of these decisions may change when more information becomes available. The bounded rationality theory aligns with decision researchers’ growing opinion that consumer preferences are often not well-defined ex-ante; rather, preferences are constructed during decision-making (Bettman et al. 1998). Taking these arguments together, we will focus on capturing the shift in preferences. Later, in the Web Appendix, we will test whether the model benefits from explicitly accounting for differences in choice consistency.

### Summary of potential changes

In sum, innovation processes can evolve in manifold ways, and it is of great importance to analyze this evolution, for which sample-based longitudinal studies are particularly well suited. Taking into account the argumentation in the aforementioned literature streams, Table 2 presents the potential changes that need to be considered in sample-based longitudinal studies. Some of these changes are induced by the respondents (cases #1 and #2) and some of these changes are induced by the researchers (cases #3–#8). Table 2 also summarizes the related assumptions in handling these changes and reveals our core findings, which we explain in more detail in the following sections.

## Model and estimation

### Covariate-extended hierarchical Bayes logit estimator

We assume a utility-maximizing respondent and base our estimation on the random utility theory (Thurstone 1927). We decompose the utility  $u_{h,i}$  of respondent  $h$  and alternative  $i$  into

a deterministic part  $v_{h,i}$  that contains observable, experimentally manipulated elements and a stochastic, i.e., unobservable part  $\varepsilon_{h,i}$ :  $u_{h,i} = v_{h,i} + \varepsilon_{h,i}$ . The stochastic part accounts for the Thurstone (1927) realization that respondents make errors in their choices that cannot be explained by the attributes and levels in the deterministic part alone. Let  $v_{h,0} = 0$  be the deterministic utility of the reference alternative, i.e., an unobserved outside option. We assume an extreme value distribution for  $\varepsilon_{h,i}$  and obtain the choice probability  $\Pr_{h,i}$  of respondent  $h$  for alternative  $i$  in choice set  $a$ , with  $I_a$  being the set of alternatives in choice set  $a$  as follows (c.f., Train 2009, p. 34f):

$$\Pr_{h,i} = \frac{e^{v_{h,i}}}{1 + \sum_{i \in I_a} e^{v_{h,i}}} \quad (h \in H; i \in I_a) \tag{1}$$

We estimate the parameters for the preferences at two layers. The upper layer captures respondents’ behavior aggregated over the population, i.e., it assumes that a multivariate normal distribution describes the behavior of all respondents. The lower layer captures respondents’ individual behavior. We assume that, given a respondent’s parameter values, his/her probability of choosing an alternative  $i$  is governed by the multinomial logit model.

Without sacrificing generalizability, we consider two discrete choice experiments at time  $t1$  and  $t2$  (in the empirical study, we expand the model to more points in time). On the upper layer, we account for the change in preferences across multiple studies through a covariate matrix  $Z_h$ , which is of size  $|H| \times 2$  and consists of  $|H|$  (the number of respondents) rows  $z_h$ , each with two columns. The first column is 1 (i.e., it captures preferences at time  $t1$ ) for all respondents, while the second column is 1 for responses in  $t2$  and 0 for the reference study in  $t1$ .

On the lower layer, we use an additive model for the individual deterministic utility  $v_{h,i}$  consisting of a  $|P|$ -dimensional vector of parameter values,  $\beta_h$ , times the design vector,  $x_{i,l}$  of product  $i$ , where  $|P|$  is for the number of parameters to estimate. Following Lenk et al. (1996), we link the lower and upper layers through the following regression model:

$$\beta_h = \theta \cdot z'_h + \varsigma_h \quad (h \in H) \tag{2}$$

with  $\varsigma_h$  following a normal distribution  $N(0, \Sigma)$  and the matrix  $\theta$  including the upper layer regression parameters to express the sensitivity of the parameter values  $\beta_h$  to the study-related covariates  $z_h$ . Equation (2) does not include a time index because the individual index  $h$  already contains the information that a respondent participated at time  $t$ .  $\theta$  is of size  $|P| \times 2$ , i.e., one row and two columns for each estimated parameter. The vector  $\beta_h$  follows a multivariate normal distribution  $N(\theta \cdot z'_h, \Sigma)$  with means  $\theta \cdot z'_h$  and covariance matrix  $\Sigma$ . The Technical Appendix (see Web Appendix) provides an outline of the estimation procedure.

**Table 2** Summary of potential changes in longitudinal studies

Case	Type of change	Induced by	Main assumptions	What we show	Addressed in empirical study
#1	Changes in respondents' preferences for attributes	Respondent	Changes in preferences can be captured through additive shifter on population level, Samples are comparable	Section Model and Estimation: How to capture and test for the significance of such shifts through a covariate-extended model	Yes
#2	Changes in respondents' consistency in decision making through better understanding of attributes	Respondent	Changes in consistency can be captured through a variance shifter on population level, Samples are comparable	Robustness section: Adding a variance shifter for each year or for each sample did improve neither internal nor predictive validity. We conclude that changes in consistency played a minor role when combining multiple discrete choice experiments	Yes
#3	Changes in experimental design	Researcher	Using the same design generation process ensures structural reliability of measured preferences	Not in the study focus, but c.f., Street and Burgess (2007)	Yes
#4	Adding attributes, which existed at times of earlier studies, but were not included	Researcher	"Null effect" or "expected individual value" assumption for missing attributes	Section Model and Estimation: For discrete choice experiments: "Null effect" or "expected individual value" assumption for missing attributes eventually leads to the same model for the estimation	Not in the focus
#5	Adding attributes, which did not exist at times of earlier studies, but are now included	Researcher	"Null effect" assumption for missing attributes		Yes (experimentally tested in 2019)
#6	Removing attributes, which remain as part of the product but are no longer in the focus of the study	Researcher	"Null effect" or "expected individual value" assumption for missing attributes		Not in the focus
#7	Removing attributes, which are no longer part of the product	Researcher	"Null effect" assumption for missing attributes		Yes (experimentally tested in 2019)
#8	Changing values of continuous attributes levels (e.g. price)	Researcher	Linearity in preferences for a continuous attribute, Changes by an additive constant shift of level values to avoid biases referred to as attribute range effects	Section Model and Estimation: Shift does not affect interpretation of the beta parameter that is associated with the continuous attribute, but rather it impacts constant. Alternative solution on handling this change in robustness section: by dummy coding continuous values	Yes (demonstrated between 2013 and 2017)

To sum it up,  $\beta_h$  captures the individual preferences of respondent  $h$  on the lower layer. On the upper layer, the first column of  $\theta$  captures the preferences in  $t1$  as a  $p$ -dimensional vector  $\alpha_{t1}$  and the second column captures the change in preferences  $\Delta\alpha$  in  $t2$  in comparison to the reference study (case #1 in Table 2), such that  $\alpha_{t2} = \alpha_{t1} + \Delta\alpha$ .

For more than two studies, let  $b$  quantify the number of parameters in the vector  $\beta_h$  and  $n$  be the number of longitudinal studies. Then, we estimate  $n \cdot b + b \cdot (b + 1) / 2$  parameters on the upper layer, i.e., one set of parameters for each longitudinal study and one covariance matrix. In contrast, the generic estimation with the single, normal population assumption estimates  $b + b \cdot (b + 1) / 2$  parameters (one set of parameters and one covariance matrix), and the separate estimator uses  $n \cdot b + n \cdot b \cdot (b + 1) / 2$  parameters ( $n$  sets of parameters and  $n$  covariance matrices). Because the covariate-extended model explicitly accounts for heterogeneity across studies but assumes one common covariance matrix, it makes the estimation more parsimonious than separating the samples (Orme and Howell

2009, p. 3). Next, we explain how changes in the attributes and levels affect the interpretability of the results.

**Adding and removing attributes over time**

We first consider the cases #4, #5, #6, and #7 from Table 2, i.e., that researchers added or removed some attributes in later studies. For illustration purposes, we subsequently consider a product class with two attributes in time  $t1$  and  $t2$  as well as a third attribute, present only in  $t2$ . Let all attributes have two levels, such that only one parameter per attribute is estimated. Respondent  $h$ 's deterministic utility for that product  $i$  is then:

$$v_{h,i} = \beta_{h,0} + \beta_{h,1} \cdot x_{i,1} + \beta_{h,2} \cdot x_{i,2} + \beta_{h,3} \cdot x_3 (h \in H; i \in I_a). \tag{3}$$

Assuming an expected individual value for the missing attribute (cases #4 and #6) implies that  $x_3 = \tilde{x}_{h,3}$  and that  $\beta_{h,3} \cdot \tilde{x}_{h,3}$  is constant across alternatives at time  $t1$ . Hence, the

unseen attribute does not affect the choice decision when using the logit model. Yet, it affects the decisions to buy or not buy the product, which is captured in the new parametrization of the constant at time  $t1$ , i.e.,  $\beta_{h,0}^* = \beta_{h,0} + \beta_{h,3} \cdot \tilde{x}_{h,3}$ , resulting into the deterministic utility function for:  $v_{h,i} = \beta_{h,0}^* + \beta_{h,1} \cdot x_{i,1} + \beta_{h,2} \cdot x_{i,2}$ . Thus, we do not need to make explicit assumptions about the level of  $x_3$  because the new constant implicitly captures preferences for the unseen information.

Assuming a null effect for missing attributes (cases #4, #5, #6, and #7) implies that  $\beta_{h,3} = 0$  at time  $t1$ . The main difference to the previous assumption is the interpretation of  $\beta_{h,0}$ , i.e., whether it contains unobserved information or not. The estimation is the same for both lines of arguments.

To illustrate the estimation, let Table 3 contain the exemplary design matrices  $X_t$  for the studies in  $t1$  and  $t2$  for the example in the previous section. Each row represents an alternative shown to respondents, and the columns contain the effects-coded attribute levels. Since only the study at time  $t2$  showed the third attribute, the column representing the unseen attribute is zero at time  $t1$ .

To account for unseen attributes in sample-based longitudinal studies, we modified the standard covariate-extended hierarchical Bayes logit estimator that turned out to be effective. Assuming a normal distribution for parameters that are related to zero-columns in the design matrix (i.e., in our example for  $\beta_{h,3}$  in  $t1$ ) would add noise to the estimation. Instead, we set this parameter in the Metropolis Hastings-step to zero. We also adjusted the draw of the  $\theta$ -matrix (see the Technical Appendix in the Web Appendix for details). We applied the two modifications solely to parameters that relate to a zero column in the design matrix. In simulations, we found that the ability to recover parameter values substantially increased with this change.

### Adjusting continuous levels of an attribute over time

In some circumstances, researchers wish to change continuous levels of attributes over time (case #8 in Table 2). Take the attribute price, which rarely stays constant over the years. Adjusting it in later studies based on market prices makes the experiment more realistic. One solution for handling changes in the continuous levels is to treat the attribute as two separate attributes: as one newly added attribute (the new prices) and as one removed attribute (the previous prices). A more elegant solution may be possible if the researcher can reasonably assume a linear relationship of the preferences for the continuous levels. Scholars have often made this linearity assumption when

dealing with the price attribute (e.g., Papiés et al. 2011; Meyer et al. 2018; Völckner 2008).

Subsequently, for illustration purposes, we consider a product class with two attributes in  $t1$  and  $t2$ . The first attribute is the same in both studies and contains two (effects-coded) levels  $x_{i,t}$ . The second attribute contains (e.g., four) continuous levels, for which we assume linearity in the corresponding preferences. Furthermore, we assume that all level values in  $t2$  have been shifted by  $\Delta x_2$ , i.e.,  $x_{i,2,t2} = x_{i,2,t1} + \Delta x_2$ . This shift by a constant factor ensures that we do not modify the attribute range, such that we can confidently rule out biases, which are commonly associated with the attribute range effect (Liu et al. 2009; Ohler et al. 2000; Verlegh et al. 2002).

On the individual layer, the deterministic utility function  $v_{h,i}$  for alternative  $i$  can be written as:

$$v_{h,i} = \beta_{h,0} + \beta_{h,1} \cdot x_{i,1} + \beta_{h,2} \cdot (x_{i,2,t1} + w \cdot \Delta x_2) \quad (h \in H; i \in I_a) \tag{4}$$

where  $\beta_{h,2}$  refers to the individual preferences for the continuous attribute, and  $w$  is a binary indicator that is 1 for observations at time  $t2$  and 0 otherwise. On the upper layer, we write the deterministic utility  $v_{i,t}$  at time  $t$  as:

$$v_{i,t} = (\alpha_{0,t1} + w \cdot \Delta \alpha_0) + (\alpha_{1,t1} + w \cdot \Delta \alpha_1) \cdot x_{i,1} + (\alpha_{2,t1} + w \cdot \Delta \alpha_2) \cdot (x_{i,2,t1} + w \cdot \Delta x_2) \quad (i \in I_a; t \in \{t1, t2\}) \tag{5}$$

For  $t1$ , (i.e.,  $w = 0$ ), Eq. (5) shrinks to  $v_{i,t1} = \alpha_{0,t1} + \alpha_{1,t1} \cdot x_{i,1} + \alpha_{2,t1} \cdot x_{i,2,t1}$ . For  $t2$ , we substitute  $\alpha_{j,t2} = \alpha_{j,t1} + \Delta \alpha_j$  for  $j \in \{1,2\}$  and obtain

$$v_{i,t2} = (\alpha_{0,t1} + \Delta \alpha_0 + \alpha_{2,t2} \cdot \Delta x_2) + \alpha_{1,t2} \cdot x_{i,1} + \alpha_{2,t2} \cdot x_{i,2,t1} \quad (i \in I_a) \tag{6}$$

Equation (6) outlines two implications on how shifts by the factor of  $\Delta x_j$  affect the results: First, despite the shift in the level values, the parameter  $\Delta \alpha_2$  still reflects the change in preferences for the updated continuous attribute over time. Consequently, its interpretation of the parameter value is not affected by the shift. Second, the changes in the level values affect the constant of the deterministic utility function, such that it must be readjusted. For the study at  $t1$ , the constant is  $\alpha_{0,t1}$  on the upper layer. For the study at  $t2$ , we must account for the shift by readjusting the constant



**Table 3** X-matrices in case of an unseen attribute in  $t1$

Respondents in $t1$				Respondents in $t2$			
Constant	Attribute 1	Attribute 2	Attribute 3	Constant	Attribute 1	Attribute 2	Attribute 3
1	1	1	0	1	-1	1	-1
1	1	-1	0	1	1	1	1
1	...	...	...	1	...	...	...
1	-1	1	0	1	-1	-1	1

according  $\alpha_{0,t2} = \alpha_{0,t1} + \Delta\alpha_0 + \alpha_{2,t2} \cdot \Delta x_2$ , but we can interpret  $\alpha_{2,t2}$  as if the original values have been used. Hence,  $\beta_{h,2}$  on the lower layer is also comparable over time.

## Empirical study

### Background

We chose electric vehicles for the study context because they represent an innovative technology that has yet to penetrate the market. Electric vehicles are an opportunity to reduce CO2 emissions and offer a promising alternative to vehicles running on gasoline or diesel in the face of limited oil resources and climate change. The push for electric vehicle diffusion has been one of the leading environmental topics for several decades (Graham-Rowe et al. 2012).

In Germany, which constitutes the focal region of our studies, the electric vehicle market was only 1.8% of the total market share in 2019, according to the German Federal Automotive Office. Like in many other countries, Germany’s federal government has set an official target of at least one million registered electric vehicles by 2020. In 2016, the government started a financial incentive program that offered 4,000€ for each bought electric vehicle with a net-list price below 40,000€. In February 2020, the government increased the financial incentive to 6,000€ and, in June 2020, to 9,000€. Of the 9,000€, the government contributes 6,000€, and the manufacturer pays the remaining 3,000€. Nevertheless, given that there were only 179,473 newly registered electric vehicles as of February 2020 (according to the German Federal Office of Economics and Export Control), the official target will be missed by the end of 2020. Hence, it is important to observe and analyze preference development over time to understand consumers’ perceptions of electric vehicles and generate reliable estimates of the market potential. After we discuss the market study results, we will revisit this case and clarify how the findings support managerial decision-making.

### Attribute selection

We started with a literature search (see Table A1 in the Web Appendix), focusing on consumer reports on electric vehicles as external sources, which resulted in a list of the ten most frequently used attributes. To narrow down the list, we employed the dual questioning approach proposed by Myers and Alpert (1968). Thereby, we surveyed 251 respondents in advance of the empirical study and asked for the perceived importance of and difference in the proposed attribute levels (c.f. Hinz et al. 2015). Based on this, we included the attributes *purchase price*, *electricity cost per 100 km*, *range per charge*, *charging time*, and *motor power*. For further information, we refer to the paper of Hinz et al. (2015), which was published in a prestigious journal.

When discussing electric vehicles with industry experts, we found that IT-enabled complementary mobility services can substantially foster user acceptance by offering a unique driving experience. Few studies have included the availability of electric charging stations or permissions to use bus lanes with an electric vehicle (Hackbarth and Madlener 2016). Still, these attributes are exogenous to manufacturers. Together with a consultancy agency, we created a list of nine exemplary complementary services that could provide a unique driving experience. Using the method of best–worst scaling, case 1 (Louviere et al. 2013), which we included after the dual questioning task, we ranked these services according to their importance in stimulating purchases.

The top four services, in order from most to least important, are: “IT-based parking space and payment”, “intelligent charging station”, “augmented reality services via head-up displays”, and “remote diagnostics and update supply”. IT-based parking systems guide drivers to parking spaces and enable automatic payment. Intelligent charging stations simplify charging the battery by automatically identifying drivers and billing the consumed energy. Augmented reality services project relevant information on the windshield via head-up display, such as navigation, or nearby charging stations’ prices and locations. The remote diagnostics and update supply enable the remote detection of defects during operation to initiate the right measures immediately. Moreover, a decentralized update supply for the software provides a fundamentally different user experience compared to conventional cars, for which updates are often carried out locally in workshops.

Even though some of these services might also be available with conventional cars, they embody the exclusive experience of using electric vehicles and thus have the potential to raise their attractiveness.

## Study setup

We conducted the study in April 2013, June 2017, and November 2019.<sup>1</sup> The questionnaire was the same with the following exceptions: In 2017, we adjusted the level values of two continuous attributes to reflect changes in the market offering. In 2019, we randomly assigned respondents to one of three experimental conditions: (1) either they saw the same survey from 2017, (2) the study contained an additional attribute, or (3) the study featured both an additional and removed attribute. In a fourth experimental condition, we separated all respondents who participated in 2017 and re-invited them to take the same survey. Re-inviting respondents from 2013 was not possible because changes in the database system prevented the market research firm from matching respondent IDs. According to their expert opinion, the overlap was at most 10%. We implemented and executed the questionnaire with the dual response task (for illustration, see Figure A1 in the Web Appendix) using the online survey platform DISE (Schlereth and Skiera 2012). Table 4 summarizes the setup of the studies.

In 2013 and 2017, we used only the top three of the four IT-enabled complementary mobility services. In 2017, we accounted for the fact that, due to technological improvements, range per charge improved substantially compared to 2013, while electric vehicles became slightly more expensive (case #8 in Table 2). Accordingly, to create realistic products in our discrete choice experiment, we increased all levels of range per charge by 150 km and the purchase prices by 5,000 €, thereby keeping the attribute ranges constant. In 2019, we kept the levels the same, but added the attribute “remote diagnostics and update supply” and removed “augmented reality services via head-up displays” in some versions of the study (case #5 and #7 in Table 2).

We constructed 14 choice sets using a D-optimal (4·2·2·4·4·2·2·2) fractional factorial design that was specialized for measuring main effects (Street and Burgess 2007). We used this design in all studies, except for the one version with the additional attribute, which required a different D-optimal (4·2·2·4·4·2·2·2·2) fractional factorial design (case #3 in Table 2). Each choice set (illustrated in Figure A1 in the Web Appendix) presented three electric vehicles. After choosing the preferred electric vehicle, respondents had to answer whether they would buy the chosen electric vehicle or not.

<sup>1</sup> For the year 2013, we use the data from Hinz et al. (2015). The studies in 2017 and 2019 are replications of the one in 2013 to demonstrate our approach and to study the aspects of sample-based longitudinal discrete choice experiments.

The two types of questions are known as dual response (Brazell et al. 2006; Schlereth et al. 2018). This variant overcomes a common disadvantage of including the no-purchase option in every choice set because we now also observe respondents’ trade-offs between the attributes even in cases where respondents would rather refrain from buying one of the presented products.

The questionnaire showed some general information about electric vehicles and asked respondents about their gender, age, and income. Afterward, they were asked: “Can you imagine purchasing an electric vehicle?” Only respondents with a general purchase intention for electric vehicles entered the discrete choice experiment. If they claimed that they could not imagine purchasing an electric vehicle or chose instead to use one as part of a car-sharing service, their survey directly ended. We assumed that their willingness-to-pay would be lower than the lowest purchase price in the experiment. We hired the same market research firm for all studies to select respondents with the same demographic criteria to create comparable samples.

## Results

### General purchase intention

Table 4 summarizes respondents’ general purchase intention in each study. In 2013, we collected 327 completed questionnaires, and 51.38% of the respondents claimed to have a general purchase intention. In 2017, of the 553 respondents, 58.95% entered the discrete choice experiment. In 2019, we randomly assigned 2,184 respondents to one of the three study versions, whereby only 40.96% indicated their general purchase intention. This strong decline is consistent across all the 2019 questionnaires (39.40%, 41.32%, and 42.16%).

We can only speculate about the reasons, but the observed decrease is backed up by the Gartner hype cycle for connected vehicles from July 2019, in which experts placed electric vehicles at the *trough of disillusionment* (Ramsey 2019). In 2019, electric vehicles still had a low market share of 1.8% in the focal country. As more information about the experience of using electric vehicles became available, their practical problems also became more apparent. As noted by Jensen et al. (2014), the purchase intention for electric vehicles decreased after the individual trial periods.

We also note the extent to which we succeeded in re-inviting the 2017 respondents. In 2019, 387 of the 553 respondents from 2017 were still active in the panel, and 305 respondents accepted the invitation (55.15%). Their general purchase intention substantially declined to 40.33%—a share that aligns with the percentage of the new respondents in 2019. The percentage is 63.37% among those respondents who also indicated a general purchase intention in 2017, and 10.53% among the others. We conclude that, in a time when electric

**Table 4** Description of our studies, number of respondents, attributes, and attribute levels

	2013	2017	2019_1_same	2019_2_add	2019_3_add_remove	2019_retakers
Description	Original study	Changes in the continuous levels range per charge and purchase price	Same study as the one in 2017; new respondents	One attribute has been added in comparison to study 2017; new respondents	One attribute has been added and one removed in comparison to study 2017; new respondents	Same study as the one in 2017; sent out to the same respondents
Number of Respondents	327	553	764	680	740	305
Number of Respondents with General Purchase Intention (%)	168 (51.38%)	326 (58.95%)	301 (39.40%)	281 (41.32%)	312 (42.16%)	No intention in 2017: 172 Intention in 2017: 133
Attributes	Unit Levels	Levels	Levels	Levels	Levels	Levels
Range per charge	km 100; 175; 250; 325	250; 325; 400; 475	250; 325; 400; 475	250; 325; 400; 475	250; 325; 400; 475	250; 325; 400; 475
Charging time	h 1; 4	1; 4	1; 4	1; 4	1; 4	1; 4
Motor power	kW 40; 80	40; 80	40; 80	40; 80	40; 80	40; 80
Purchase price	€ 15,000; 20,000; 25,000; 30,000	20,000; 25,000; 30,000; 35,000	20,000; 25,000; 30,000; 35,000	20,000; 25,000; 30,000; 35,000	20,000; 25,000; 30,000; 35,000	20,000; 25,000; 30,000; 35,000
Electricity cost per 100 km	€ 1; 3; 5; 7	1; 3; 5; 7	1; 3; 5; 7	1; 3; 5; 7	1; 3; 5; 7	1; 3; 5; 7
IT-based parking space and payment	supported; not supported	supported; not supported	supported; not supported	supported; not supported	supported; not supported	supported; not supported
Intelligent charging station	supported; not supported	supported; not supported	supported; not supported	supported; not supported	supported; not supported	supported; not supported
Augmented reality services via head-up displays	supported; not supported	supported; not supported	supported; not supported	supported; not supported	-	supported; not supported
Remote diagnostics and update supply	-	-	-	supported; not supported	supported; not supported	-

vehicles became more and more available, respondents' tendency to purchase them decreased more than increased.

We next tested for differences in gender, age, and income. Using multiple t-tests with a Bonferroni correction, we found no significant differences for gender in any comparisons (all  $p > 0.1$ ). Concerning age, the 2019 group of retakers was significantly older than they were in 2017 ( $p < 0.01$ ); while we did expect this result, we found no significant differences in age when compared to the three other groups of new respondents in 2019. Further, for income, we did not observe any significant difference between the 2017 and 2019 retakers ( $p > 0.1$ ), as well as between the three groups of new respondents in 2019 (all  $p > 0.1$ ). Yet the 2019 retakers indicated that they earned significantly more ( $p < 0.01$ ) than the other respondents in 2019, such that among all the tests, the only one that could indicate a problem of the sample's representativeness is the significant income difference between 2017 and 2019. It could also indicate some changes over time in the general population. In the Web Appendix, we provide robustness tests, which rule out that the differences in income systematically biased our results.

## Model comparison

We assumed a linear relationship for the purchase price and range per charge (c.f., Table A1 in the Web Appendix). We estimated purchase price per 10,000€ and range per charge per 100 km to ensure that all X-matrix values vary with a similar magnitude and that prior distributions are consistent across parameters. In total, we ran 7,000 iterations, of which we used the last 2,000 iterations to estimate the posterior. We visually assessed the convergence of the posterior in the burn-in phase using trace plots of the likelihood function.

We compared the log-marginal density (LMD) of the covariate-extended hierarchical Bayes model against the separate estimation and the generic model. The estimation that used most of the parameters on the upper layer, i.e., the separate estimation with 475 parameters (=  $5 \cdot 11 + 1 \cdot 12$  parameters plus  $5 \cdot (11 \cdot (11 + 1)/2) + 12 \cdot (12 + 1)/2$  for the six covariance matrices), performed best concerning internal validity (LMD of -16,297, see Table A2 in the Web Appendix). It is followed by the model with the second most parameters, i.e., the covariate-extended model (145 parameters; LMD of -16,410) and the model with the least number of parameters, i.e., the generic one (90 parameters; LMD of -16,676) performs worst.

Yet an improvement in fitting the choices that were used for the estimation does *not* necessarily equal a better model. Overfitting can occur if these parameters do not improve the fit to the choices held out for validation (Ellickson et al. 2019; Orme and Howell 2009). For the two holdouts, the covariate-extended model (-4765) edges out the separate (-5953) and generic (-4940) models. The separate estimation performs worst, such that the improved internal model fit does not translate into better predictions. To assess the size of the

differences, we adapted a test from Ellickson et al. (2019), which entails leaving out 100 respondents and predicting the mean absolute differences of their observed choice shares. Even though the covariate-extended model is best here, the differences are rather small (between 6.73% and 6.87%).

Orme and Howell (2009, p. 19) wrote that predictions are not the main reason for using covariate models, but they “allow us to test more formally [...] the differences between segments and the part-worth.” And indeed, we can see the main advantage of using the covariate-extended model, as this modeling approach best aligns with the goals of a sample-based longitudinal study. These goals are to link multiple studies within one estimation (which the separate estimation cannot do) and to explicitly capture differences in preferences across samples (which the generic model cannot do). Besides, it enables researchers to formally test whether changes in preferences across samples are substantial or merely noise: We simply inspect the signs of the posterior draws for each parameter in the  $\theta$ -matrix (Orme and Howell, 2009). The matrix contains the upper layer's changes in preferences relative to a reference study (Eq. 2); hence, when  $> 95\%$  or  $< 5\%$  of draws are positive, these covariate weights are significantly different from zero, at or better than the 90% confidence level (two-sided test). We will subsequently demonstrate the value of such a test.

## Changes in preferences

In Table 5, we report the results of the covariate-extended model. All signs and magnitudes are consistent and reasonable, indicating high face validity. Table 5 also reports the percentage of positive and negative draws with the 2017 study as a reference, when the general purchase interest in electric vehicles peaked. For example, when comparing the 2013 study results to the ones of 2017, the preferences that largely changed related to the electric vehicle-specific attributes, but not to the complementary mobility services. Respondents in 2013 were more price-sensitive, with a greater emphasis on charging time (the “1 h” attribute level had a positive change and “4 h” a negative one) and less emphasis on motor power (the less-preferred 40KW had a positive change and the more-preferred 80KW had a negative one).

Table 5 demonstrates the benefit of inspecting the signs of the posterior draws of the  $\theta$ -matrix: A pragmatic approach would have been to compare the absolute or relative differences in parameter values between two samples. However, in Table 5, we observe that large differences do not automatically imply that the two samples differ in preferences. For instance, a rather large difference has the constant between 2017 and the 2019 retakers (i.e.,  $0.65 = -1.34 - (-1.99)$ ), and a rather small difference has the attribute intelligent charging station (i.e.,  $0.10 = 0.63 - 0.53$ ). Yet, the  $\theta$  draws only statistically support that the preferences for the complementary mobility service “intelligent charging stations” decreased; there was no

**Table 5** Average parameter values based on covariate-extended model

Attributes	Attribute level	2013		2017 (reference)		2019_1_same		2019_2_add		2019_3_add_remove		2019_4_retakers	
		Parameter values	Change	Parameter values	Change	Parameter values	Change	Parameter values	Change	Parameter values	Change	Parameter values	Change
Constant		-1.35		-1.99		-1.54		-1.20	+	-1.01	++	-1.34	
Range per charge	Per 100 km	0.71		0.70		0.60	-	0.63		0.62		0.69	
Purchase price	per 10,000 €	-1.18	-	-0.94		-1.10	-	-1.18	--	-1.15	--	-1.15	
Charging time	1 h	0.29	++	0.14		0.24	+	0.26	++	0.28	+++	0.28	++
	4 h	-0.29	--	-0.14		-0.24	-	-0.26	--	-0.28	---	-0.28	--
Electricity cost	1 €	1.07		0.99		0.96		0.93		1.04		0.90	
per 100 km	3 €	0.48	++	0.34		0.53	+++	0.44	+	0.42	+	0.45	
	5 €	-0.30		-0.32		-0.43	-	-0.27		-0.38		-0.33	
	7 €	-1.25	--	-1.02		-1.06		-1.10		-1.08		-1.02	
Motor power	40 kW	-0.41	+	-0.52		-0.43	+	-0.25	+++	-0.38	+++	-0.42	
	80 kW	0.41	-	0.52		0.43	-	0.25	---	0.38	---	0.42	
IT-based parking space and payment	Supported	0.42		0.50		0.40	--	0.29	---	0.36	---	0.44	
	Not supported	-0.42		-0.50		-0.40	++	-0.29	+++	-0.36	+++	-0.44	
Intelligent charging station	Supported	0.58		0.53		0.53		0.32	---	0.47		0.63	+
	Not supported	-0.58		-0.53		-0.53		-0.32	+++	-0.47		-0.63	-
Augmented reality services via head-up displays	Supported	0.20		0.22		0.22		0.24				0.15	-
	Not supported	-0.20		-0.22		-0.22		-0.24				-0.15	+
Remote diagnostics and update supply	Supported							0.21	+++	0.26	+++		
	Not supported							-0.21	---	-0.26	---		

Posterior assessment of change in parameters in comparison to 2017 study: + : > 90%; + + : > 95%; + + + : > 99%; - : < 10%; - - : < 5%; - - - : < 1% positive  $\Delta$  draws

statistical support for a change of the constant. Hence, looking only at the value of parameters (as with the separate estimation) might provide misleading insights Table 5.

Fig. 1 compares and graphically displays the average attribute importance weights (Louviere and Islam 2008) over time. Economic considerations (i.e., electricity costs and purchase price) dominated choice decisions. This is particularly interesting because electricity costs, which were the most important attribute throughout the three studies, are rather neglected in communication about electric vehicles. Range per charge came in third: It gained importance in 2017, but lost some in 2019. Obviously, it was initially important to be able to overcome a certain range, which was still too small in 2013. Considering that we have adjusted the range per charge between 2013 and 2017, the benefit of the longer range decreases after 2017 as the market-observed ranges increase further. Overall, the importance weights for the attributes remained mostly stable for over six years. There were some position changes, but rarely by more than one rank.

### Adding and removing an attribute

Next, we examine how the inclusion of the ninth attribute affected preferences for the other attributes. Thereby, we used only the experimental conditions of the 2019 study and reran

the estimation, using *2019\_1\_same* as a reference. Table A3 in the Web Appendix contains the posterior mean for each study, together with a test on the percentage of positive and negative  $\theta$  draws indicating whether differences are substantial or mere noise.

We observe that the fourth complementary mobility service (*2019\_2\_add* vs. *2019\_1\_same*) proportionally drew its importance weight mostly from the four lower-ranked attributes, i.e., the other three complementary mobility services and motor power. All other attributes are unaffected. This may be because the new variable itself is one of the lower-ranked attributes. When the fourth complementary mobility service replaced the head-up display attribute (*2019\_3\_add\_remove* vs. *2019\_1\_same*), the other importance weights did not change on a significant level.

### Preferences of retakers vs. new sample of respondents

Using Table 5 and Table A3 (in the Web Appendix), we compare the preferences of the retakers (*2019\_4\_retakers*) against their preferences in 2017 and the new sample of respondents (*2019\_1\_same*). In comparison to 2017, the attribute “charging time” is the one that changed on a substantial level by gaining importance. There were also some changes

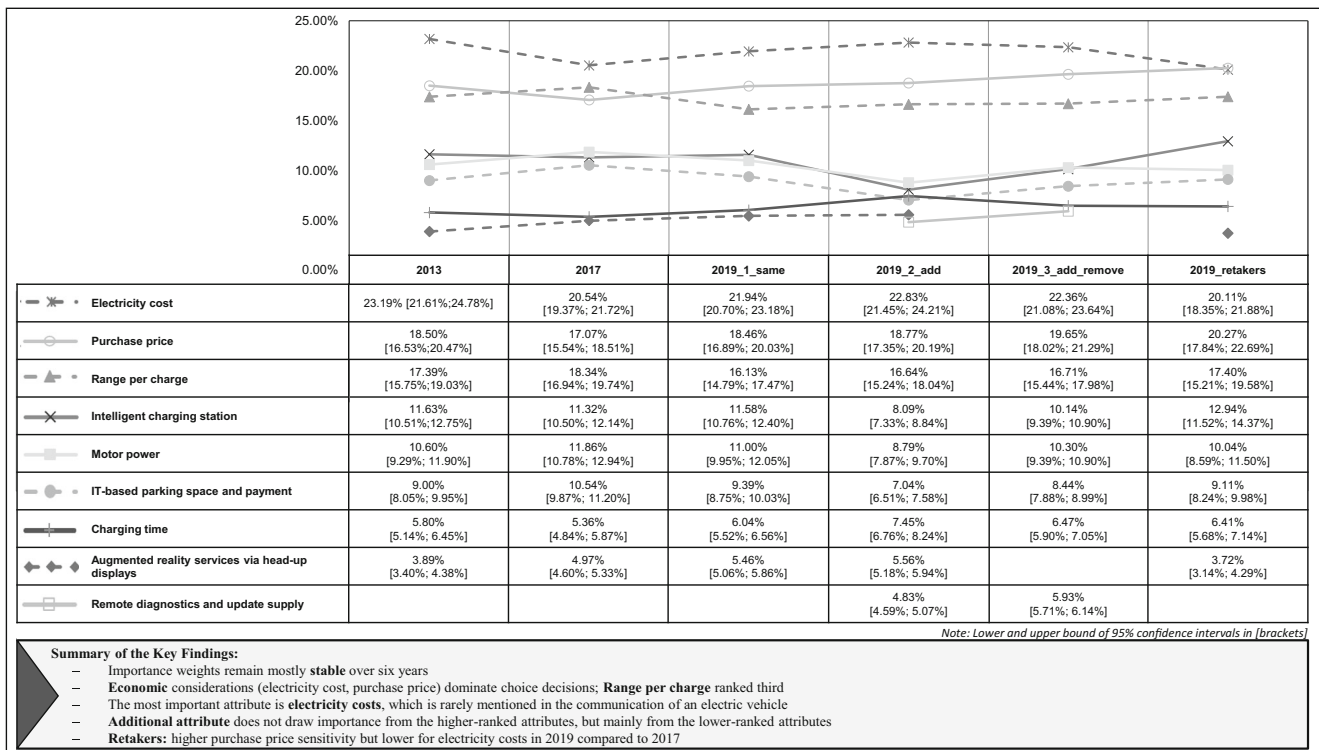


Fig. 1 Average importance weights and key findings

for the complementary mobility services “intelligent charging station” and “augmented reality services via head-up displays”; however, we consider these changes as minor because the former one slightly gained in importance and the latter lost some. When comparing the sample of retakers against the same study with completely new respondents (2019\_1\_same), we detected nearly no differences in the attribute parameters.

Overall, we conclude that it is valuable—if possible—to invite the same respondents again. There is great value in the potential to modify certain aspects in the selected list of attributes and levels, as proposed by the sample-based longitudinal discrete choice experiment. However, if it is not possible, then entirely new respondents provided similar insights in our study.

### Changes in purchase probabilities

In the last column of Table 6, we report the share of purchase decisions. Whereas the general purchase intention decreased from 2017 to 2019, the share of purchase decisions increased monotonously among those respondents with a general purchase intention.

The subsequent counterfactual simulation combines the joint effect of the two trends. We created a stylized scenario, in which we report the purchase probabilities for a “status quo” scenario. We derived this scenario by analyzing a basic electric vehicle offer (here: range per charge of 250 km; charging time of four hours; motor power of 40 kW; purchase price

of 40,000€; electricity cost per 100 km of 5€, and no complementary mobility services) and then alternately varied several attribute levels one by one. We summarize the results in Table 6.<sup>2</sup> For each scenario, we averaged the individual purchase probability by employing the second term in Equation (A1) (c.f., Technical Appendix in the Web Appendix) and then multiplied this probability by the average purchase intention of the respective year. For better readability, we combined the three versions in 2019.

In all scenarios, purchase probabilities increased from 2013 to 2017, but decreased in 2019. In 2013, only 4.60% of respondents would buy the status quo product, compared to 7.21% in 2017 and 4.38% in 2019. Considering the 4,000€ governmental incentive in 2016, we observed a purchase probability of 4.68% in 2019, but 5.23% when the subsidy increased further to 9,000€ in June 2020. The additional 5,000 € subsidy increases the probability of buying an electric vehicle by +0.55% or, in absolute figures, by 19,855 electric vehicles (i.e., 0.55% · 3.61 million newly registered vehicles in Germany in 2019, according to Statista).

<sup>2</sup> We replicate the counterfactual simulation for the separate and generic estimation. The market shares of the generic estimation are close to the ones of Table 6 with a mean absolute difference in the percentages of .39% (std.dev. .18%). For the separate estimation, the percentage differences between each simulated condition are similar in size compared to the ones of Table 6. However, the overall market shares discriminate more between samples: the separate 2017 shares are on average .78% (std.dev. .79%) higher and the other market shares are 1.57% (.47%) lower.

Comparing the predictions of the counterfactual between the new 2019 samples and the retakers, we also observed only minor differences. These results further suggest that surveying a new sample vs. asking respondents from previous studies provides comparable insights.

### Model-based managerial decision

Finally, we used the 2019 estimates to present an illustrative optimization problem for automobile manufacturers. Since June 2020, the government supports the purchase of an electric vehicle with a discount of 9,000€. This discount comprises a 6,000€ federal subsidy and an obliged manufacturer’s share of 3,000€. The market simulation in Table 6 assumed that the car manufacturer passes on the whole discount directly to the consumer. A more realistic view is that the manufacturer would have offered a voluntary discount, even without the governmental program, and thus must decide on how much of the 9,000€ it should pass on to its customers to maximize profits. This maximization problem is also relevant beyond the context of electric vehicles. Essentially, the 6,000€ federal subsidy represents a reduction of the manufacturer’s variable costs. The underlying general decision problem is: What percentage should a manufacturer pass on, in case the variable costs decrease?

We assumed a list price  $l$  and an average discount  $d$ , which would have been offered without an incentive program, such that the actual purchase price  $p$  is  $p = l - d$ . Since manufacturers vary in their discounts, we examined the optimal decision-making of a car manufacturer under different pricing strategies. While Tesla, for example, communicates a zero-discount strategy, other manufacturers usually grant a certain

discount, although the exact amount is not known and varies between manufacturers. Thus, we distinguished three pricing strategies in order to determine the effects of a subsidy. We considered a manufacturer who, with a previous discount of 11,000€, has already granted a discount of more than 9,000€ before the government subsidy. As a second pricing strategy, we assumed a previous discount of less than 9,000€ (i.e., 7,000€). Finally, we also considered a manufacturer that does not offer a discount at all.

Assuming a manufacturer’s margin  $m$ , the variable costs  $c_{var}$  without governmental support are  $c_{var} = p \cdot (1 - m)$ . The new regulations of the federal government reduce the variable costs by 6,000€, such that  $c_{var}^* = p^* \cdot (1 - m) - 6000€$ , where  $p^*$  is the new purchase price corresponding to the new discount  $d^*$ . Let customer  $h$ ’s purchase probability equals (c.f., Equation (A.1)):  $Pr_h(p) = \frac{\exp(v_h(p))}{(1 + \exp(v_h(p)))}$ . For better readability, we omitted the index  $i$  for an electric vehicle. The maximization problem is:

$$\max \pi(d^*) = (l - d^*) \cdot (1 - c_{var}^*) \cdot \sum_{h \in H} [Pr_h(p^* = l - d^*)] \tag{7}$$

$$d^* \geq 9000€ \tag{8}$$

$$d^* \geq d \tag{9}$$

**Table 6** Counterfactual simulations on purchase probabilities

	Status Quo	40,000€ →36,000€ (government’s financial incentive 2016)	40,000€ →31,000€ (government’s financial incentive 2020)	5€ per 100 km →3€ per 100 km	250 km range →325 km range	IT-based parking space and payment
2013	4.60%	5.04%	5.89%	5.76%	5.29%	5.96%
2017	7.21%	7.51%	8.12%	8.13%	8.30%	9.25%
2019	4.38%	4.68%	5.23%	5.39%	5.02%	5.28%
(average across the 3 versions)						
2019_retakers	3.87%	4.35%	5.09%	4.57%	4.61%	5.27%
	Intelligent charging station	Augmented reality services via head-up displays	Remote diagnostics and update supply	With top two mobility services	Share of Purchase Decisions	
2013	6.37%	5.25%	4.60%	8.07%	42.64%	
2017	9.29%	8.30%	7.21%	11.47%	45.33%	
2019	5.57%	4.80%	4.88%	6.67%	46.03%	
(average across the 3 versions)						
2019_retakers	5.77%	4.67%	3.87%	7.51%	45.88%	

Status Quo: range per charge: 250 km; charging time: four hours; motor power: 40 kW; purchase price: 40,000€; electricity cost per 100 km: 5€, no complementary mobility services

The decision variable in the model (7)-(9) is the new discount  $d^*$ . The objective function (7) maximizes the profit  $\pi$  of an electric vehicle manufacturer. The constraint in (8) ensures that at least 9,000€ will be passed on to customers, and (9) ensures that the manufacturer offers at least the discount it would have offered, even without governmental support.

Besides the pricing strategy (i.e., the discount), we also vary the margins in two levels (5% and 20%). Thus, we examine the maximization problem for a total of six settings. For this analysis, we map the characteristics of one of the best-selling electric vehicles in Germany in 2019 (according to Statista) to the corresponding attribute levels of our study design (List price: 35,900€; range per charge: 300 km, charging time: 1 h; motor power: 80 kW; electricity cost per 100 km: 7€, none of the complementary mobility services). In Table 7, we report the results for the (3 · 2 =) six settings.

The model recommends the minimum discount in all settings, determined by the constraints (2) and (3). Thus, manufacturers with previous discounts lower than 9,000€ (e.g., 7,000€ or no discount at all) should now offer a discount of 9,000€, and manufacturers with a previous discount greater than 9,000€ should stick to their previous discount (i.e., 11,000€). The pricing policy of the manufacturer determines how much it benefits from the governmental incentive program. Manufacturers, who charged list price without a discount before the program, face a decrease in their margin or even losses. Their rationale for taking part in such a program is probably rather due to reasons of competitiveness. Those with originally large discounts benefit from the federal subsidy. The larger the original margin, the more they benefit.

The results indicate that the purchase probability does not increase sufficiently if manufacturers partly pass on the federal subsidy. This observation implies that the price sensitivity is too low (respectively, the parameter for the purchase price is too close to zero). The electric vehicle’s features and the consumers’ general attitudes toward buying one mainly drive the decision to buy an electric vehicle—the price is less of an

issue, thereby indicating a price-inelastic demand. To compensate for a passed-on discount of 1,000€, the average purchase probabilities would have to change by at least 1.81% to increase the manufacturer’s profits (see the last column in Table 7).

Although the results of our longitudinal study indicate an increase in the price sensitivity between 2017 and 2019 (see Table 5), a sufficient increase in purchase probabilities is currently out of reach. From today’s perspective, the governmental incentive program encourages consumers to buy the electric vehicle before the program runs out, but it is barely able to stimulate additional purchases. Nevertheless, it is important to observe whether changes will occur to the price sensitivity in the near future.

## Conclusion

### Methodological contribution and contribution to theory

Discrete choice experiments are well known, but they are rarely used to study preferences over time. In this research, we proposed a sample-based longitudinal discrete choice experiment, together with the covariate-extended hierarchical Bayes logit estimator, to track changes in preference over time using different samples of respondents. We also structured all elements that can change from a respondent’s and the researcher’s perspective in order to guide researchers in their modeling. Hence, the sample-based longitudinal discrete choice experimental approach can be easily transferred to other research topics.

When examining the performance of the covariate-extended model, we found that its utility is motivated more by theoretical arguments and its statistical testing ability than by its potential gains in internal and predictive validity (in line with Orme and Howell 2009; Sentis and Geller 2010). We

**Table 7** Optimal discount strategy for manufacturers with new governmental program

Purchase price p without governmental support	Discount d without governmental support	Margin per electric vehicle	Original average purchase probability Pr(p)	Optimal new purchase price p*	Optimal discount d*	Change in average purchase probability	New margin per electric vehicle	Required change in purchase probability to compensate for passing on an additional 1,000€
24,900€	11,000€	5%	18.05%	24,900€	11,000€	+0%	29.10%	+2.89%
24,900€	11,000€	20%	18.05%	24,900€	11,000€	+0%	44.10%	+1.81%
28,900€	7,000€	5%	16.10%	26,900€	9,000€	+0.93%	20.24%	+3.83%
28,900€	7,000€	20%	16.10%	26,900€	9,000€	+0.93%	36.36%	+1.94%
35,900€	0€	5%	13.62%	26,900€	9,000€	+3.41%	-4.48%	not applicable
35,900€	0€	20%	13.62%	26,900€	9,000€	+3.41%	15.54%	+5.35%



consider its testing ability to be the core benefit because it answers a range of questions that researchers typically have when conducting longitudinal studies. One of these questions is whether it is appropriate to survey different samples of respondents. Here, we statistically observed that upper layer preferences do not differ for the new sample relative to the respondents from previous studies. While we cannot generalize from this single study, this finding suggests that interviewing new participants is viable. At the same time, we demonstrated the need to ask different respondents in later studies: After 2.5 years, only 55.15% of the respondents from 2017 responded to our survey invitation in 2019. Without our approach, it would have been necessary to gather large, costly sample sizes at the beginning of the studies to guarantee an appropriate sample size for later studies. Our approach relaxes this requirement. Thus, it reduces the expenses for conducting longitudinal studies and even allows researchers to expand sample sizes in later studies—for example, to experimentally test for different attribute specifications, as we did in the 2019 study.

Concerning the testing ability of the covariate-extended model, we showed that statistically testing for differences is necessary; looking only at absolute or relative changes in parameter values can be misleading in detecting significant changes. Thereby, our approach can be more general than comparing groups of respondents over time: It enables a new level of flexibility in analyzing multiple studies that only partially overlap. For example, a globally operating company has data sets available at not only several points in time, but also for different groups (e.g., different market regions like U.S. vs. European car buyers) that the management would like to combine and evaluate. For these groups, different attributes or attribute characteristics are conceivable depending on the country, and researchers can now account for them separately in our proposed model.

Moreover, adding and changing attributes enables researchers to react to future changes in the market. There are manifold reasons why the ability to change attributes may be beneficial to a longitudinal study (c.f., Table 2). For example, the importance of some attributes may not be recognized until later, or new technology features arise after the longitudinal study begins. We explicitly considered this prospect in our modeling and provided a theoretical base, with explicit assumptions, about how respondents of earlier studies take unseen information into account. We also help explain changes in the continuous levels of attributes by outlining that they mostly affect the constant, but not the parameter related to the changed attribute. Concerning adding attributes, we experimentally demonstrated that their percentage in importance weights were not drawn proportionally from all other attributes; rather, some of the attributes were unaffected. Having knowledge about which attributes are affected can support product managers in their decision-making and communication.

A final finding relates to the evolution of preferences for innovations. As one of the very few who have conducted a longitudinal study over a longer period of time, we show that preferences do not necessarily evolve monotonously. Accordingly, forecasting models that extrapolate the development of preferences into the future are rather difficult. Longitudinal studies are all the more suitable for this purpose.

### Managerial contribution

Let us return to the previously mentioned manufacturer ACME, which might use these analyses and the corresponding information for future management decisions. Based on the first two studies in 2013 and 2017, ACME might have made larger investments since 2017, as the curve showed a strong growth trend promising a constant and rapid breakthrough of the technology. However, since we observed an inverted U-shaped curve with a peak in 2017 as the overall effect over the entire period, these investments would not have paid off to the extent hoped for so far. The sales market for electric cars has continued to grow, but not to the extent that firms expected. With the third major study in 2019, the manufacturer would gain more up-to-date information about the market and adjust its marketing measures accordingly. According to the Gartner hype cycle, innovation diffusion can fluctuate in the pre-mass market phase, which is why not all investments should be discontinued. It may be that the breakthrough to the plateau of productivity is imminent.

Notably, the preferences for the different attributes remained largely constant. Price sensitivity was relatively low, although electricity costs and purchase price were the most important attributes. From this, we can deduce that manufacturers should more directly communicate information about electricity costs, as these are consistently among the most important attributes. The low price sensitivity suggests that the purchase price is not that important: People either have strong preferences toward owning an electric vehicle or continuing to drive a conventional fuel-based vehicle. A potential reason for this is that the media excessively covers the problems of electric vehicles. For example, people might hear about the low number of charging stations in Germany, or that small accidents can damage the battery—the car's main value—and lead to a total loss relatively quickly. Moreover, prospective buyers were concerned about batteries setting fire to the entire electric vehicle or the garage in the house and that the monetary advantage of an electric vehicle may be lost when charging at commercially operated charging stations. Finally, critical reports question the environmental friendliness of electric vehicles, e.g., if the electricity is generated in coal-fired power plants. In sum, researchers can use sample-based longitudinal studies to analyze the overall perception of an innovation within the population and, above all, test

possible solutions that will achieve acceptance among end customers.

## Limitations

Like any research, our study features some limitations. First, we only studied one product in our longitudinal study. We encourage future studies to examine other products over a longer time to observe possible similarities or differences in terms of product development or market penetration. By including different product categories or product lifecycles, future studies could address cases #4 and #6 (Table 2), which we left out empirically.

Second, the available options were characterized solely by their attribute levels. The preference and purchase decisions were, therefore, hypothetical. This is exacerbated by the fact that the market penetration of electric vehicles in the focal country is still very low, so our respondents may have felt a lack of experience with the product under consideration.

Third, in using a D-efficient design, we only measured the main effects and could not observe interaction effects. This limitation probably applies to almost all discrete choice experiments. Some authors have proposed designs that can also handle a selected set of attribute-interaction effects (e.g., Yu et al. 2006). However, we are not aware of any research that applies them in an empirical setting, as they require a substantially higher number of choice sets—a situation we wanted to avoid for our study. Follow-up studies could conceivably address this limitation.

Finally, our study results suggest that surveying the same respondents over time provides similar insights compared to questioning a new sample of respondents. Future research could challenge the boundary conditions and the generalizability of this finding. Research on biases, such as the mere exposure effect (e.g., Zajonc 1968), could serve as a suitable starting point together with dynamic models that allow for individual-level heterogeneity around an aggregated trend (e.g., Liechty et al. 2005). Another boundary condition is the number of changed attributes and levels, i.e., how much change is possible to guarantee a certain level of temporal stability and structural reliability of the results.

**Supplementary information** The online version of this article (<https://doi.org/10.1007/s11747-020-00758-8>) contains supplementary material, which is available to authorized users.

**Acknowledgements** The authors gratefully thank Luigi Bianco for his support in data collection for the 2013 study and Prof. Dr. Bernd Skiera for his help and feedback during the colloquium. This work has been funded by the German Research Foundation (DFG) within the Collaborative Research Center (CRC) 1053—MAKI.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agarwal, J., DeSarbo, W. S., Malhotra, N. K., & Rao, V. R. (2015). An interdisciplinary review of research in conjoint analysis: recent developments and directions for future research. *Customer Needs and Solutions*, 2, 19–40.
- Ambos, T. C., Cesinger, B., Eggers, F., & Kraus, S. (2019). How does de-globalization affect location decisions? A study of managerial perceptions of risk and return. *Global Strategy Journal*, 16, 210–236.
- Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, 25, 187–217.
- Bhattacharjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly*, 25, 351–370.
- Bradlow, E. T., Hu, Y., & Ho, T.-H. (2004). A learning-based model for imputing missing levels in partial conjoint profiles. *Journal of Marketing Research*, 41, 369–381.
- Branco, F., Sun, M., & Villas-Boas, J. M. (2012). Optimal search for product information. *Management Science*, 58, 2037–2056.
- Brazell, J. D., Diener, C. G., Karniouchina, E., Moore, W. L., Séverin, V., & Uldry, P.-F. (2006). The no-choice option and dual response choice designs. *Marketing Letters*, 17, 255–268.
- Coupey, E., Irwin, J. R., & Payne, J. W. (1998). Product category familiarity and preference construction. *Journal of Consumer Research*, 24, 459–468.
- Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 28, 307–319.
- Ellickson, P. B., Lovett, M. J., & Ranjan, B. (2019). Product launches with new attributes: a hybrid conjoint-consumer panel technique for estimating demand. *Journal of Marketing Research*, 56, 709–731.
- Fourt, L. A., & Woodlock, J. W. (1960). Early prediction of market success for new grocery products. *Journal of Marketing*, 25, 31–38.
- Gensler, S., Hinz, O., Skiera, B., & Theysohn, S. (2012). Willingness-to-pay estimation with choice-based conjoint analysis: Addressing extreme response behavior with individually adapted designs. *European Journal of Operational Research*, 219, 368–378.
- Gilbride, T. J., Currim, I. S., Mintz, O., & Siddarth, S. (2016). a model for inferring market preferences from online retail product information matrices. *Journal of Retailing*, 92, 470–485.
- Graham-Rowe, E., Gardner, B., Abraham, C., Skippon, S., Dittmar, H., Hutchins, R., et al. (2012). Mainstream consumers driving plug-in battery-electric and plug-in hybrid electric cars: a qualitative analysis of responses and evaluations. *Transportation Research Part A: Policy and Practice*, 46, 140–153.
- Hackbarth, A., & Madlener, R. (2016). Willingness-to-Pay for alternative fuel vehicle characteristics: a stated choice study for Germany. *Transportation Research Part A: Policy and Practice*, 85, 89–111.
- Hinz, O., Schlereth, C., & Zhou, W. (2015). Fostering the adoption of electric vehicles by providing complementary mobility services: a

- two-step approach using best-worst scaling and dual response. *Journal of Business Economics*, 85, 921–951.
- Jensen, A. F., Cherchi, E., & de Dios Ortúzar, J. (2014). A long panel survey to elicit variation in preferences and attitudes in the choice of electric vehicles. *Transportation*, 41, 973–993.
- Kurz, P., & Binner, S. (2010). Added Value through Covariates in HB Modeling? *Proceedings of the Sawtooth Software Conference*.
- Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical bayes conjoint analysis: recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, 15, 173–191.
- Liakhovitski, D., & Shmulyian, F. (2011). Covariates in Discrete Choice Models: Are They Worth the Trouble? *ART Forum*.
- Liechty, J. C., Fong, D. K. H., & DeSarbo, W. S. (2005). Dynamic models incorporating individual heterogeneity: utility evolution in conjoint analysis. *Marketing Science*, 24, 285–293.
- Linden, A., & Fenn, J. (2003). *Understanding gartner's hype cycles*. Accessed on January 28, 2020 from <https://www.bus.umich.edu/KresgePublic/Journals/Gartner/research/115200/115274/115274.pdf>.
- Liu, Q., Dean, A., Bakken, D., & Allenby, G. M. (2009). Studying the level-effect in conjoint analysis: an application of efficient experimental designs for hyper-parameter estimation. *Quantitative Marketing and Economics*, 7, 69–93.
- Louviere, J., Lings, I., Islam, T., Gudergan, S., & Flynn, T. (2013). An introduction to the application of (case 1) best–worst scaling in marketing research. *International Journal of Research in Marketing*, 30, 292–303.
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge: Cambridge University Press.
- Louviere, J. J., & Islam, T. (2008). A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. *Journal of Business Research*, 61, 903–911.
- Lucker, J., Hogan, S. K., & Sniderman, B. (2018). Fooled by the hype: Is it the next big thing or merely a shiny new object? *Deloitte Review*, 23, 84–95.
- McCullough, J., & Best, R. (1979). Conjoint measurement: temporal stability and structural reliability. *Journal of Marketing Research*, 16, 26–31.
- Meeran, S., Jahanbin, S., Goodwin, P., Neto, Q. F., & J. (2017). When do changes in consumer preferences make forecasts from choice-based conjoint models unreliable? *European Journal of Operational Research*, 258, 512–524.
- Meyer, J., Shankar, V., & Berry, L. L. (2018). Pricing hybrid bundles by understanding the drivers of willingness to pay. *Journal of the Academy of Marketing Science*, 46, 497–515.
- Myers, J. H., & Alpert, M. I. (1968). Determinant buying attitudes: meaning and measurement. *Journal of Marketing*, 32, 13–20.
- Ohler, T., Le, A., Louviere, J., & Swait, J. (2000). Attribute range effects in binary response tasks. *Marketing Letters*, 11, 249–260.
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17, 460–469.
- Orme, B., & Howell, J. (2009). Application of covariates within sawtooth software's CBC/HB program: theory and practical example. *Proceedings of the sawtooth software conference*.
- Papies, D., Eggers, F., & Wlömert, N. (2011). Music for free? How free ad-funded downloads affect consumer choice. *Journal of the Academy of Marketing Science*, 39, 777–794.
- Ramsey, M. (2019). *Hype Cycle for Connected Vehicles and Smart Mobility*. Accessed on February, 13, 2020 from <https://www.gartner.com/en/documents/3955767/hype-cycle-for-connected-vehicles-and-smart-mobility-201>.
- Reinders, M. J., Frambach, R. T., & Schoormans, J. P. L. (2010). Using product bundling to facilitate the adoption process of radical innovations. *Journal of Product Innovation Management*, 27, 1127–1140.
- Rogers, E. M. (1962). *Diffusion of innovations (Social science)*. New York: Free Press.
- Schlereth, C., & Skiera, B. (2012). DISE: Dynamic intelligent survey engine. In A. Diamantopoulos, W. Fritz, & L. Hildebrandt (Eds.), *Quantitative marketing and marketing management* (pp. 225–243). Wiesbaden: Gabler Verlag.
- Schlereth, C., & Skiera, B. (2017). Two new features in discrete choice experiments to improve willingness-to-pay estimation that result in SDR and SADR: separated (Adaptive) dual response. *Management Science*, 63, 829–842.
- Schlereth, C., Skiera, B., & Schulz, F. (2018). Why do consumers prefer static instead of dynamic pricing plans? An empirical study for a better understanding of the low preferences for time-variant pricing plans. *European Journal of Operational Research*, 269, 1165–1179.
- Sentis, K., & Geller, V. (2010). The impact of covariates on HB estimates. *Proceedings of the sawtooth software conference*.
- Simon, H. A. (1955). A Behavioral model of rational choice. *The Quarterly Journal of Economics*, 69, 99–118.
- Street, D. J., & Burgess, L. (2007). *The construction of optimal stated choice experiments: theory and methods*. Hoboken.: Wiley-Interscience.
- Swait, J., & Andrews, R. L. (2003). Enriching scanner panel models with choice experiments. *Marketing Science*, 22, 442–460.
- Teas, R. K. (1985). An analysis of the temporal stability and structural reliability of metric conjoint analysis procedures. *Journal of the Academy of Marketing Science*, 13, 122–142.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34, 273–286.
- Train, K. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.
- Verlegh, P. W. J., Schifferstein, H. N. J., & Wittink, D. R. (2002). Range and number-of-levels effects in derived and stated measures of attribute importance. *Marketing Letters*, 13, 41–52.
- Verlegh, P. W. J., Schifferstein, H. N. J., & Wittink, D. R. (2002). Range and number-of-levels effects in derived and stated measures of attribute importance. *Marketing Letters*, 13, 41–52.
- Völckner, F. (2008). The dual role of price: decomposing consumers' reactions to price. *Journal of the Academy of Marketing Science*, 36, 359–377.
- Yang, L., Toubia, O., & DeJong, M. G. (2015). A bounded rationality model of information search and choice in preference measurement. *Journal of Marketing Research*, 52, 166–183.
- Yu, J., Goos, P., & Vandebroek, M. L. (2006). The importance of attribute interactions in conjoint choice design and modeling. *Working paper 0601, Department of Decision Sciences and Information Management, Katholieke Universiteit Louven*.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1–27.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.