Journal of Choice Modelling 48 (2023) 100412

Contents lists available at ScienceDirect

Journal of Choice Modelling

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

Formative versus reflective attitude measures: Extending the hybrid choice model

J.M. Rose [a,*], A. Borriello [b], A. Pellegrini [c]

[a] The University of Sydney Business School, Institute of Transport and Logistics Studies, The University of Sydney, NSW, 2006, Australia
[b] European Commission, Joint Research Centre (JRC), Ispra, Italy
[c] The University of Sydney Business School, Institute of Transport and Logistics Studies, The University of Sydney, NSW, 2006, Australia

ARTICLE INFO

Keywords:
Formative measurement
Reflective measurement
Attitudes
Perceptions
Discrete choice modelling
Hybrid choice models

ABSTRACT

The inclusion of attitudinal indicator variables within discrete choice models is now largely common practice. Typically, this involves the estimation of multiple indicator multiple cause (MIMIC) type models which are used to construct latent attitudinal variables that are then employed as independent variables within standard discrete choice models. Such models, collectively termed hybrid choice models (HCM) assume a particular causal relationship between the indicator variables, latent construct, and choice. In effect, the underlying assumption of such a model system is that latent variables of interest exist independent of the indicator variables used to measure them, and that the survey items used are reflective in nature insofar as responses to such questions reflect the underlying constructs. In this paper, we describe an alternative form of attitude measure, known as formative measures, where the items themselves are used to create the latent variable rather than the other way around. In addition to making a distinction between formative and reflective attitudinal measures, the paper seeks to describe how the HCM can be adapted to model different types of attitude question formats. Further the paper seeks to act as a catalyst for choice modellers to think more about the quality and validity of attitudinal items capture in survey questionnaires, by placing more emphasis on proper scale development techniques.

1. Introduction

Discrete choice models (DCM) have become the dominant method for understanding preferences and behaviour of economic agents observed to make decisions either in real markets or in response to hypothetical scenarios presented to them as part of a broader questionnaire. Estimation of DCMs typically involves defining alternative specific utility functions that are used to describe the role that the attributes of the alternatives being examined and/or the characteristics of the decision makers' play in the decision-making process. By observing how decisions vary as the attributes of the alternatives or characteristics of the decision makers differ within data, the analyst is able to derive utility weights for each that reflect how the modelled variables influence choice.

As noted over 40 years ago by McFadden (1980) however, "The theory of the economically rational utility-maximizing consumer, interpreted broadly to admit the effects of perception, state of mind, and imperfect discrimination, provides a plausible, logically unified foundation for the development of models of various aspects of market behaviour." Inspired by the development of the hybrid

* Corresponding author.
E-mail address: john.rose@sydney.edu.au (J.M. Rose).

https://doi.org/10.1016/j.jocm.2023.100412
Received 1 July 2022; Received in revised form 19 April 2023; Accepted 21 April 2023
Available online 14 May 2023

conjoint model (e.g., Green et al., 1981; Green 1984), McFadden (1986) went on to derive the theoretical foundations for the hybrid choice model (HCM) designed specifically for the purposes of exploring the role that psychological constructs (e.g., attitudes, perceptions) play in discrete choice situations. First operationalized by Train et al. (1987), the model as it currently is utilized assumes that attitudes and perceptions are latent constructs that can be uncovered using indicator variables collected via surveys of the decision makers concurrently with the relevant choice data. Despite early acknowledgement of the importance that psychological factors play when individuals make decisions, it has only been in the last ten to fifteen years that research has really sought to examine how psychological, cognitive, emotional, and social factors, including attitudes and perceptions, influence choice behaviour.

To date, analysts employing the HCM have tended to assume the relationship between indicator variables and latent constructs be reflective, meaning that changes in the unobserved constructs manifest as variations in the corresponding observed indicator variables. As such, these models assume a particular directionality in the relationship between underlying unobserved latent constructs and the responses decision makers provide to survey questions measuring attitudes and perceptions. For example, an individual not concerned about climate change may respond accordingly by selecting values corresponding with such an attitude to a battery of questions dealing with climate change embedded within some survey. If later in life, due to personal circumstances, the same individual changes their underlying attitude towards climate change, then the responses to the same battery of questions would reflect this change. In this way, the latent construct informs the responses to the indicator variables, not the other way around. Such indicators have come to be termed reflective (or effects) indicators, and models to which they are applied to, effects models.

MacCallum and Browne (1993), however, note that in some cases indicators may be viewed as causing, rather than being caused by, the latent variable. Termed formative (or causal) indicators, changes in the values of the indicators result in changes in the underlying latent construct, the exact opposite to reflective indicator variables. Derived from research in psychology and sociology (e.g., Blalock 1964; Bollen 1984; Bollen and Lennox 1991), formative measures of attitudes have been used in marketing (e.g., Diamantopoulos 1999; Diamantopoulos and Winklhofer 2001; Rossiter 2002) and strategy (see e.g., Venaik et al., 2005).

Formative measurement scales differ to reflective scales in that the latent construct being measured is a function of the indicators, and as such, the causal direction between the latent construct and the indicators is reversed. Further, the characteristics of the indicators differ between the two approaches. In reflective models, the direction of concomitant variation is such that changes in the latent construct precede changes in the indicator variables employed to measure the latent construct. This suggests that whilst the indicator variables must be correlated with the latent construct, they may also be somewhat interchangeable in that the specific questions asked, subject to internal validity considerations, have no influence on the underlying latent construct being measured. On the other hand, the latent construct is assumed to be a function of the indicator variables used in formative measurement scales, meaning that the number and specific types of indicator variables adopted will impact on the latent construct. Further, in reflective measures, the expectation is that the indicator variables will be correlated with one another, otherwise they are not measuring the same underlying latent construct. With formative measurement scales, no such correlation is required (see Venaik et al., 2005).

The majority of choice modelling studies have to date employed reflective measures to capture the attitudes of those sampled, then employ the HCM framework to link underlying latent attitudinal constructs to individuals' choice behaviour. The empirical analysis of such data in this setting is now widely accepted. However, it is worth noting that in some cases, researchers in different fields have inadvertently modelled formative scales as if they were reflective. Within the transportation literature, attitude measures have been defined as being either individual specific or alternative specific in nature (see e.g., Bahamonde-Birke et al., 2017). Under this taxonomy, individual specific measures are assumed to reflect more generally held attitudes of a global nature (e.g., attitudes towards the environment) whereas alternative specific measures relate to perceptions linked to specific goods or services within a market place (e.g., attitudes towards buses). Such a distinction, whilst useful, ignores the fact that the former type of measures is more likely to utilise formative measurement scales whilst the latter are more likely to make use of reflective measures. For example, Daziano (2012) defines appreciation of car features using items measuring eight aspects, these being purchase price, vehicle type, fuel economy, horsepower, safety, seating capacity, reliability, and styling. Such constructs are formative in nature insofar as they represent aspects of the vehicle rather than attempt to measure underlying attitudes towards vehicles. In a more recent study, Guzman et al. (2021) measure the satisfaction of public transport system considering the specific satisfaction for fare, comfort, security, as well as general overall satisfaction towards the public transport system. As with the measurement items formalized in Daziano (2012), these indicator variables appear to measure constructs that contribute to the satisfaction of a public transport system, rather than measure latent satisfaction. In other words, levels of comfort, security, and fare influence satisfaction, rather than being measures of underlying satisfaction. Similarly, Jin et al. (2020) attempt to measure satisfaction with a transport system in a reflective fashion by means of indicators measuring convenience, comfort, timeliness, cost, privacy and safety of the system, with the indicators forming (as opposed to depending on) the latent measure. Again, subjects build the satisfaction with the transport system on different domains, such as convenience, comfort and cost, and it is a change in these indicators that leads to a change in the satisfaction.

Within the Health Economics literature for example, Kløjgaard and Hess (2014) measure pro-surgery attitude via reflective indicators proxying back pain at present and past, leg pain at present and past, influence of pain onto relationships, and physical support needed. Once again, it would be more appropriate to consider the construct as formative, in that changes to the indicators (e.g., higher back pain) will change the subjective attitude towards surgery and not the other way around. In the same field, Santos et al. (2011) incorrectly framed the attitude towards toilets and sewerage connections, which is in fact formed by the indicators measuring prestige, monetary valorisation, comfort, better image, among others.

In a different field of research, Fantechi et al. (2022) explore the effect of three attitudes, namely attitude towards animal welfare, towards hunting and towards wild game meat in a preference study about meat consumption. The authors model all the latent constructs through a reflective measurement setting. This choice seems appropriate for the first two types of attitudes, which "cause" an impact on the corresponding measurement items. For example, the attitude towards animal welfare will influence the rating that the subject assign to the items “It is important that the food I normally eat has been produced in a way that animals have not experienced pain” and “It is important that the food I normally eat has been produced in a way that animals’ rights have been respected”. Opposite is the case of the third latent construct, attitude towards wild game meat, which is composed by, rather than reflects onto, the indicators safety (“It is safe to eat”), taste (“It tastes good”), and convenience (“Its price is fair compared to product quality”).

The purpose of the current paper is two-fold. Firstly, we seek to introduce the concept of reflective versus formative attitudes to the choice modelling community by extending the current HCM framework to include formative attitudinal measures. In doing so, we expose a known issue with the use of reflective attitude measures, that being the difficulty of utilising such measures for forecasting purposes, particularly when applied to cross-sectional survey data (see Chorus and Kroesen 2014). The second purpose of this paper seeks to demonstrate that the limitation of reflective measures in forecasting can be somewhat overcome when using cross sectional data if different respondents are exposed to alternate information states, that can serve as covariates alongside socio-demographic variables. For completeness, we also test the impact of different information states on models incorporating formative measures.

The remainder of the paper is organised as follows. In the following section, we discuss the theory and different modelling approaches that can be applied to formative and reflective attitude measures. Section 3 of the paper then describes an empirical study that is used to demonstrate the various models presented in Section 2. Next, the model results are presented, before section 5 demonstrates how each model can be applied in practice. The paper concludes with sections involving a general discussion and conclusion drawn.

2. Reflective versus formative measures

In this section, we outline the theories and modelling approaches adopted to estimate both formative and reflective measures as applied to discrete choice models. We begin with a discussion of reflective measurement theory given that such measures represent the dominant form of indicators used to date by choice modellers across all applied areas.

2.1. Reflective measurement and choice

The desire to include attitudinal influences into discrete choice models requires that the analyst identifies the underlying structural relationships between different latent unobserved constructs, one set associated with the attitudes and the other with the utilities derived from each of the modelled alternatives. Typically, the modelling process employed by discrete choice modellers consists of a series of two or more linked models estimated simultaneously. The first set of models are designed to measure attitudes, with the second capturing preferences across the observed set of discrete choices, with the latter models assumed to be influenced by the output of the former. With respect to the modelling of attitudes, the conventional approach involves identifying covariation between assumed underlying latent attitudinal constructs, and observed indicators, which are designed to measure the latent constructs. The most common model framework adopted assumes that any observed variation in indicator I is associated with variation in some underlying latent construct, Z, such that any exogenous intervention that impacts on Z will be reflected by changes to I. Thus, the relationship between indicator and construct is assumed to be reflective, resulting in a causality flows in the direction from the latent construct to the indicator variables.

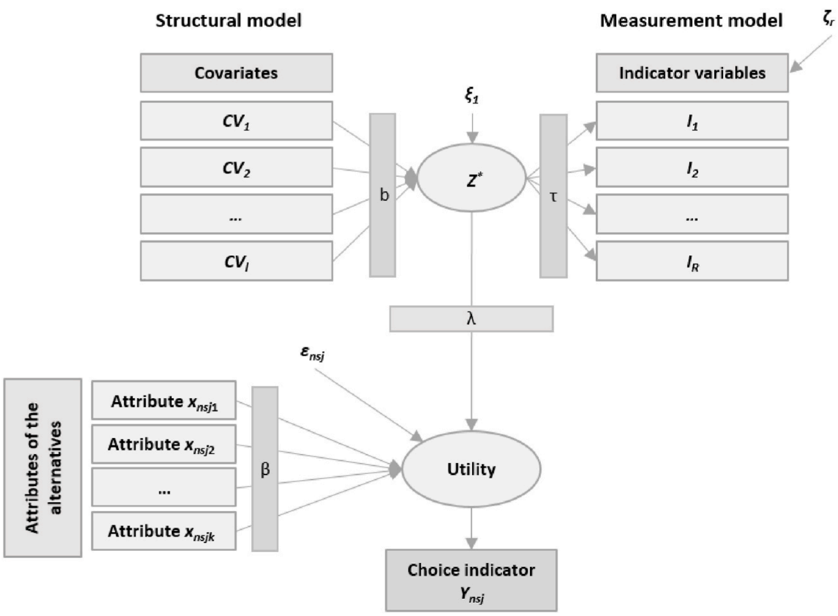


Fig. 1. Traditional HCM model structure.

Within the HCM framework, a structural equation model (SEM) is typically utilized to model psychological constructs and consists of latent variables that then enter the utility functions of a classical DCM. Traditionally, HCMs have often been specified as multiple indicator multiple cause (MIMIC) models (Joreskog and Goldberger 1975). In a MIMIC model, structural equations are defined such that the latent variables are explained by either observable variables or causal indicators. Because the dependent variable in the structural equation is latent, a (set of) measurement equation(s) is needed to define the links with its observable variables (i.e., effect indicators). Finally, the latent variables are then related to the discrete choice model either as main effects, interaction effects or as an influence on the overall scale of utility. The traditional HCM structure in the case of a single latent variable is diagrammatically represented in Fig. 1.

2.1.1. Structural model

Let n designate decision maker $n = 1, 2, \ldots, N$, and $m = 1, 2, \ldots, M$ represent the latent constructs assumed by the analyst. Then let each latent variable, z^*_{nm}, be related to L covariates, CV_{nl}, each with a decision weight, b_{ml}, then

$$z^*_{nm} = \sum_{l=1}^{L} b_{ml} CV_{\mathrm{nl}} + \xi_{nm}, \tag{1}$$

where $\xi_{nm} \sim N(0,1)$. The error terms of the structural model can be correlated by the means of a Cholesky decomposition procedure, albeit most models estimated in practice rely on other approaches. Next, the model is normalized by setting a unit variance for each latent variable. By setting one or more of the parameters, b_{ml} equal to zero, the associated covariates will not be related to the corresponding latent construct. As demonstrated by Equation (1), the model structure is flexible enough to allow for multiple latent variables.

2.1.2. Measurement model

The measurement model links the latent variables estimated via Equation (1) to the indicator variables. Let I_r represent the rth indicator variable, then the measurement model can be represented as follows:

$$I_{nr} = \sum_{m=1}^{M} \tau_m z^*_{nm} + \zeta_{nr}, \tag{2}$$

where $\zeta_{nr} \sim N(0, \omega^2)$, such that the variances of the measurements are not required to be the same. Even though a heteroskedastic general covariance matrix is possible for the measurement equations, because of identification restrictions, the measurement equations cannot be correlated (see, Daziano and Bolduc 2012). In the case that the analyst wants to estimate the variance of the latent variable in the structural model, then for identification purposes one τ (per each latent variable) should be normalized to one.

2.1.3. Discrete choice model

The last component of the HCM is the discrete choice model. The discrete choice model is specified in a similar fashion to traditional choice models, with the difference being that the latent variable obtained from the structural model will enter into one or more of the utility functions of the discrete choice models. Let $j = 1, 2, \ldots, J$ represent alternative and assuming a linear in the parameters, the utility specification for model can be represented as per Equation (3). Alternative specific constants (ASC, index $k = 0$) can be estimated for up to J-1 alternatives by setting one constant to zero. x_{nsjk} in Equation (3) represents the kth attribute associated with alternative j in choice observation s as observed by respondent n. z^*_{nm} are the latent variables estimated via Equation (1) and λ_{mj} are parameters estimated that reflect the influence of the various latent variables on choice.

$$U^*_{nsj} = \sum_{k=0}^{K} \beta_{jk} x_{nsjk} + \sum_{m=1}^{M} \lambda_{jm} z^*_{nm} + \varepsilon_{nsj}, \tag{3}$$

where ε_{nsj} is IID EV1 distributed, and hence follows a Logit model specification. It is possible for different latent variables to be associated with different utility functions, by setting different λ_{mj} equal to zero. Notice that latent variables measured through the same parameter can enter maximum *J-1* utility functions. The specification shown in Equation (3) assumes that the latent variables enter utility as main effects. Other specifications allow the latent variables to be interacted with other the attributes or socio-demographic variables contained within the model utility functions or enter the model via scale.

2.1.4. Issues and concerns regarding reflective measures and discrete choice models

Whilst the HCM framework is now widely accepted as to how best to link attitudinal data with choice data, the approach is not without criticism. Chorus and Kroesen (2014) note three issues that may arise within the HCM framework. First, attitudes and choice behaviour are likely to be influenced by similar experiences, and hence the two may be endogenously related. This is problematic given the assumed directional nature of the relationship within the model framework, with latent attitudes influencing choices, but not the other way around (see Fig. 1). Second, the widespread use of cross-sectional data precludes causal inferences from being drawn from such models, insofar as there does not exist any within-subject-variation with respect to the attitudinal elements of most surveys necessary to understand how attitudes may change over time. Third, individual respondents are likely to engage with surveys in unexpected ways, such as answering questions so as to be consistent with the answers to previous questions posed, rather than providing truthful replies. The current paper seeks to address the first two of these issues, whilst also acknowledging the importance of the third.

2.2. Formative measurement and choice

Formative measures differ substantially to reflective measures, and hence the modelling of such data requires the HCM approach be adapted somewhat. Bollen and Bauldry (2011) identify three approaches to the modelling of formative measures, with the appropriate technique to use determined by the characteristics of the measures under consideration. The first approach is employed when the formative indicators share no common underlying concept and are only grouped together under some broad overarching theme. The second approach is used when the indicator variables of the formative measurement scale have conceptual unity. The third approach involves the use of covariates, albeit we do not explore this approach further within this paper.

2.2.1. Composite model approach

Formative indicators can be considered as a form of composite indicator, insofar as the underlying constructs formed are derived as explanatory combinations of indicators which are determined by a mix of variables (Fornell and Bookstein 1982). As such, composite indicators are weighted elements that form a composite variable for which there is no disturbance term. That is, the composite variable is assumed to be an exact linear combination of the composite indicator variables, and the estimated coefficients represent neither structural nor causal coefficients, but rather being simply weights used to derive the composite score (see Bollen and Bauldry 2011). Typically, a linear function is imposed for this purpose, as given by Equation (4),

$$C_{nm} = \sum_{r=1}^{R} \mu_{mr} I_{nr}, \tag{4}$$

where n denotes the nth decision maker, m the mth construct, and I_{nr} the rth indicator variable.

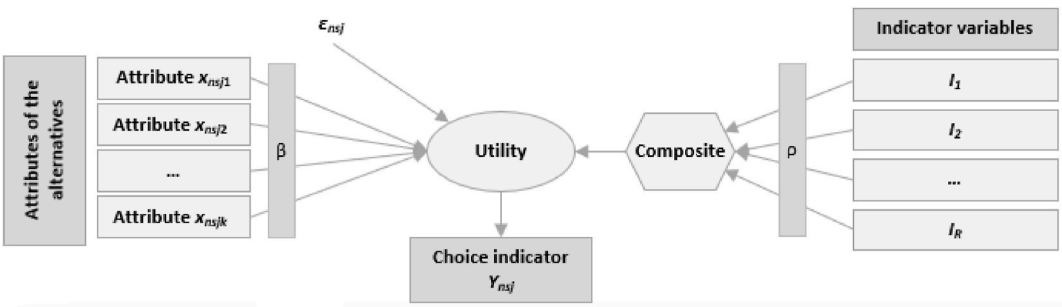
Unlike the traditional HCM based on the MIMIC approach, the composite measure HCM does not require the estimation of structural and measurement models, and hence is far simpler to estimate. The model simply estimates Equation (4) and embeds the results within the utility functions of the discrete choice model, as shown in Equation (5), and diagrammatically in Fig. 2a. For identification purposes, the constant in Equation (4) may need to be normalized to zero. The parameter ρ_{jm} takes values 1 or 0, according to which latent constructs the analyst wants to include in the specific utility functions. It is worth noting that in the case of a single composite variable, the model functionally collapses to one in which the indicator variables enter utility as separate items, as shown in Fig. 2b. The model can be estimated simultaneously or sequentially, and as with the traditional HCM, the composite measure may enter utility either as a main effect, as an interaction effect with another variable, or via the model's scale.

$$U_{nsj}^{*} = \sum_{k=0}^{K} \beta_{jk} x_{nsjk} + \sum_{m=1}^{M} \rho_{jm} C_{nm} + \varepsilon_{nsj}. \tag{5}$$

It is generally advised in practice to avoid including covariates when computing composite scores based on formative scale measures. This is because the theoretical framework under which formative measures are derived is such that indicator variables are assumed to form the composite measure, whereas covariates are often viewed within the psychometric literature as control variables designed to account for omitted variable bias. Alternatively, covariates may represent determinates of a latent construct, as opposed to representing the very concept of the latent construct, as is the case for formative measures (see Bollen and Bauldry 2011).

2.2.2. Factor analytical model approach

Formative scale indicators need not be related to each other, and are often represented by a set of quite dissimilar variables with little or even no apparent conceptual overlap, other than they combined to form some aggregate measure of interest. Therefore, formative scale indicators can in practice be positively, negatively, or completely unrelated to one another. Nevertheless, scale



a: Composite measure HCM model structure

Fig. 2a. Composite measure HCM model structure.

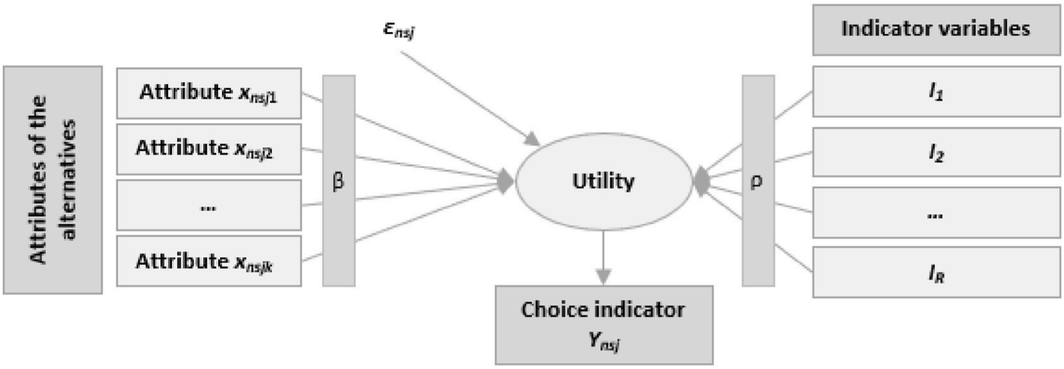


Fig. 2b. Collapsed composite measure HCM model structure.

development in the social sciences typically involves the creation of multiple indicators that are conceptually related to a common unifying construct. As such, measurement scale items used in survey work often exhibit some form of inter item correlation. Where formative scale indicators are conceptually related, and provided that they all share the same directional relationship with the latent construct of interest, then their effects on the latent construct can be considered to be structural. This means that formative scales with causal links to one or more latent constructs can be analysed using such methods as factor analysis, which can then be utilized to obtain measures of the latent constructs. In the current study, we propose using the factor scores from the factor analysis to measure the latent variables, albeit we acknowledge the existence of other different approaches. Specifically, we adopt the non-refined weighted sum approach where the indicator variables are simply weighted by their respective factor loadings and summed to form a single composite measure.

With this technique, it is first necessary to derive the factor loadings based by applying factor analysis to the indicator variables. To do this, we standardize the indicator variables and regressing these against a random variable, F, such that

$$I^z_{nr} = \sum_{g=1}^{G} \theta_{rg} F_{nr} + \delta_{nr}, \tag{6}$$

where I^z_{rn} refer to the standardized indicator for item r, F is a random term drawn from a standard normal distribution, and θ_{rg} is the factor loading associated with the rth indicator variable, and the gth latent construct. δ_{nr} are independently distributed stochastic terms. Next, the composite measure stems from weighting the non-standardized indicator variables by the factor loadings obtained in Equation (6) and summing the result. Thus, for the gth construct, we obtain

$$C_{ng} = \sum_{G=1}^{G} \theta_{rg} I_{nr}, \tag{7}$$

The resulting composite scores then enter into the discrete choice model in the exact same manner as described by the composite model approach. The overall model structure is shown in Fig. 3. Whilst somewhat more complicated than the composite model approach described in Section 2.2.1, the model can easily be estimated using Maximum Likelihood estimation methods.

2.2.3. Issues and concerns regarding formative measures and discrete choice models

One objective of this paper is to introduce formative measures into discrete choice models. To the best of the authors' knowledge, no prior research has investigated potential issues regarding this type of constructs in the DCM setting. Nevertheless, the first issue noted by Chorus and Kroesen (2014) may be valid also for formative constructs insofar as choice behaviour and latent (formative) attitudes may be influenced by similar experiences. Also, as with any other type of survey dealing with psychological constructs, respondents may attempt to align their perceptions (and attitudes) with the choices, forcing endogeneity on the latent constructs. More in general and not specifically in DCM setting, several critics have been brought forward to formative indicators. We refer the interested reader to Bollen and Diamantopoulos (2017) and Hardin (2017), who examine in details the most common criticisms.

One relevant (to discrete choice modellers) criticism relates to the absence of measurement error. Indeed, some scholars argue that measurement error is likely present in most "formative" indicators, but this error is implicitly assumed absent. First, it is important to distinguish between different formative measurement approaches. As outlined above, in the factor analytical model the (formative) indicators include an unobserved stochastic error term and therefore the criticism is not justified under this approach. Different is the case of the composite model approach, under which the latent variable is a linear combination of the (formative) indicators. Here, the indicators are error-free in their measurement stage. Although this may seem a strong assumption, it should be noted that under this approach the relation between indicators and (latent) variable "is generally not seen as causal but simply as a convenient way to form a summary composite variable" (Bollen and Diamantopoulos 2017). Indeed, "the composite variable is referred to as a "latent" variable, but the researcher assumes the composite indicators perfectly determine the composite" (Bollen and Bouldry 2011). For example,

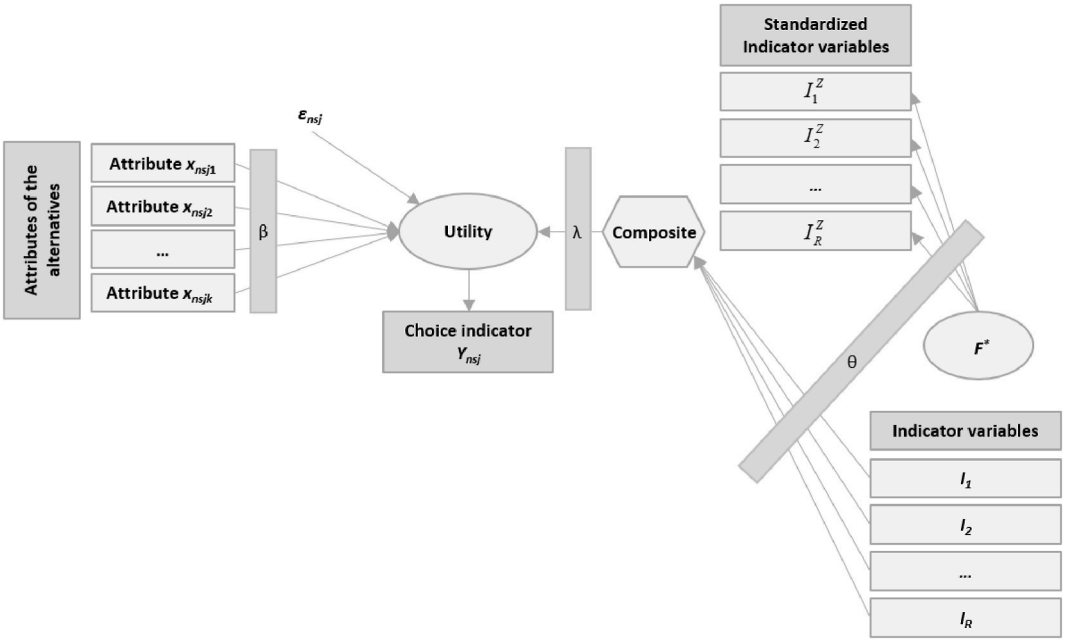


Fig. 3. Factor Analytical composite measure HCM model structure.

Heise (1972) creates standardized coefficients by combining race, gender and age. In this setting, the measurement error is very marginal if any, so will be its effect onto the (latent) variable (Bollen and Diamantopoulos 2017). In other circumstances when the measurement error is not marginal, then the model can be adapted by creating a separate latent variable, which will include a measurement error, for each formative indicator.

In the context of DCM, the formative indicators measured via a factor analytical approach will bear no complications nor concerns. Under the composite approach, where indicators enter directly the utility functions of the alternative(s), these can be thought of as proxies of the (latent) variable in the sense described by Guevara (2015) .[1]

2.3. *Formative versus reflective measures: empirical considerations*

Table 1 presents a summary of the theoretical and empirical considerations related to formative versus reflective measures adapted from Coltman et al. (2008). To summarise the differences between the two measures, reflective measures assume that the underlying construct exists independent of any measurement and that the indicator variables measure the degree to which an individual exhibits the underlying latent construct. In contrast, some formative measures (composite measures) exist only as a result of measurement taking place with the items forming the latent construct. This difference implies opposite directional causality with reflective measures assuming that the latent construct causes the responses to the indicator variables, whilst formative measures assume that responses to the indicator variables cause the latent construct. The concept of causality in measurement is important with many authors and reviewers rejecting the idea of the existence causal inference based on survey data. Unfortunately, scale development and the use of indicator variables assume casual inference as do the models used to estimate such data. Indeed, the primary fundamental difference lies in the direction of causality between the indicators and the latent construct of interest.

Although not shown in the table, reflective measures imply the need for multiple items or indicators related to each latent construct, whereas formative measures do not. This does not mean that formative measures cannot benefit from utilising multiple items, with best practice suggesting that this be so. For reflective measures, each latent construct should have several related indicators (preferably no less than three). Given that the indicators are assumed to be related, and the fact that the underlying latent variable induces the responses to the indicators and not the other way around, removing one or more indicators should firstly not influence the latent variable, and secondly, not dramatically change the modelled outcomes. Further, the fact that the items or indicators are conceptually related implies that answers to reflective indicator measures should be correlated with one another. None of these propositions are true for formative measures, as the indicators are used to form the latent construct such that removing one or more indicators will change the value of the construct being measured. With respect to the correlation structure of formative measures, the measures themselves may or may not be correlated in practice, with the correlation structure assumed dictating how they should be modelled.

[1] We thank the editor for pointing this out.

Table 1
Theoretical and empirical considerations of formative and reflective measures.

Consideration	Reflective measure	Formative measure
Theoretical considerations		
1. Nature of construct	Latent construct exists • Latent construct exists independent of the measures used	Latent construct is either formed or exists • Composite latent construct is determined as a combination of its indicators • Latent construct with error exists independent of the measures used
2. Direction of causality between items and latent construct	Causality from construct to items • Variation in the construct causes variation in the item measures • Variation in the item measures does not cause variation in the construct	Causality from items to construct • Variation in the construct does not cause variation in the item measures • Variation in the item measures causes variation in the construct
3. Characteristics of items used to measure the construct	Items are manifested by the construct • Items share a common theme • Items are interchangeable • Adding or dropping an item does not change the conceptual domain of the construct	Items define the construct • Items need not share a common theme • Items are not interchangeable • Adding or dropping an item may change the conceptual domain of the construct
Empirical considerations		
4. Item intercorrelation	Items should have high intercorrelations • Empirical test: Internal consistency and reliability via Cronbach alpha, average variance extracted and factor loadings	Items can have any pattern of intercorrelation but should possess the same directional relationship • Empirical test: Indicator reliability cannot be assessed empirically
5. Item relationships with construct antecedents and consequences	Items have similar sign and significance of relationships with antecedents/consequences as the construct • Empirical test: Content validity is established based on theoretical considerations, and assessed via convergent and discriminant validity	Items may not have similar significance of relationships with the antecedents/consequences as the construct • Empirical test: Nomological validity can be used to empirically structural linkage with another criterion variable
6. Measurement error and collinearity	Error term in items can be identified • Empirical test: common factor analysis can be used to identify and extract out measurement error	Error term cannot be identified if the formative model is estimated in isolation • Empirical test: vanishing tetrad test can be used to determine if the formative items behave as predicted • Collinearity should be ruled out by standard diagnostics such as the conditions index

2.4. Formative versus reflective measures: how to choose

The correct specification of the measurement model is relevant for the design of the latent construct, for the estimated parameters and for the overall theoretical framework. First, the misspecification can lead to the discard of valid measures only because they are assessed using the incorrect tools. For example, if researchers fail to appropriately design a "formative in nature" construct and test the items for reliability (which is typical of reflective constructs) via the Chronbach's alpha index, they can discard some items only because they do not display internal consistency. Furthermore, MacKenzie et al. (2005) demonstrated that measurement model misspecification can largely bias unstandardized structural parameter estimates. Overall, the usefulness of a theory is lowered when the incorrect measurement type is adopted (Coltman et al., 2008).

Adapted from MacKenzie et al. (2005) and Bollen and Diamantopoulos (2017) and based on Table 1 above, we list four questions that can be useful to address the choice.

1. Is the latent construct exogenous or endogenous to the indicators? If a change in the latent construct will more likely lead to a change in the indicators (exogenous latent construct), then the construct should be reflective. If a change in the indicators will more likely lead to a change in the latent construct, then the construct should be formative.
2. Are the indicators conceptually interchangeable? In other words, do they share a common domain? If so, the latent construct is reflective.
3. Are the items intercorrelated? If not, the construct is surely not reflective. Reflective constructs are measured through extremely intercorrelated items. If the items are interrelated, then either model can be appropriate and other criteria should be used to determine the nature of the construct.
4. Do all the indicators have the same antecedents/consequences? Due to their interchangeability, indicators of a reflective construct should have the same antecedents and consequences.

When this conceptual guide still leaves the researcher unsure about the nature of the latent construct, it is possible to use the empirical Vanishing Tetrad Test proposed by Bollen and Ting (2000).

3. Empirical data

The empirical analysis of this paper is based on a data set collected from 2490 respondents drawn from across all of Australia between the 4th March till the 10th March 2022. A total of 3176 respondents were contacted from the online panel QOR Surveys (https://qorsurveys.com.au) of which 2556 completed an internet-based questionnaire suitable only for use with a PC, laptop or tablet. As such, the majority of incompletes arose due to respondents using incompatible devices such as a mobile phone. Data cleaning, included removing respondents who provided straight-line responses to attitudinal questions (i.e., providing the same rating for all indicator items), as well as respondents who completed the survey in an implausible time period. This resulted in the loss of an additional 66 respondents.

Table 2 presents the socio-demographic characteristics of the final sample compared to the known population based on the 2016 Census (data from the more recent 2021 Census is as yet unavailable). With respect to geographical representation, the sample does a decent job matching the population distribution. The final sample does however slightly over represent respondents from Queensland, and underrepresents respondents from Western Australia. The sample also consists of higher proportion of individuals who reside in Capital cities than should otherwise be, presents a higher mean age and slightly overrepresents the female group. Also, the average gross weekly income is slightly higher than the population average in 2016 (six years prior to the survey taking place).

With respect to education, the sample reflects better those with only a school diploma, under sampling those with an educational attainment achieved post schooling. The other education category includes those who left school in year 10 (the earliest year possible), those whose education was obtained overseas and does not meet any of the other nominated criteria, and those who did not specify the level of their highest education attained. Although we provide sample statistics on the labour market status of respondents, no such data is given from the Census. This is because the reporting of such data does not fully align with the categories used in the survey.

Despite differences between the sample and the population, we choose not to weight the data. This is because firstly, we are not interested in forecasting, but rather demonstrating theoretical advancements and nothing will be gained from exogenously weighting the data. Second, when data are obtained at the level of individual respondents, it is common for sampling weights designed to make the sample representative of the target population be applied to each observation. Yet, evidence suggests that such weights are inappropriate for estimation or hypothesis testings, but instead should be applied to market simulation contexts (see e.g., McFadden et al., 2006).

Table 2
Demographic characteristics of the sample.

	Sample	Population
NSW	0.29	0.29
Victoria	0.26	0.23
South Australia	0.08	0.07
Queensland	0.23	0.18
Western Australia	0.10	0.18
ACT	0.02	0.02
Tasmania	0.02	0.02
Northern Territory	0.00	0.01
Capital City	0.64	0.54
Age	46.56	38.00
Female	0.54	0.51
Gross Income	1 510.27	1 438.00
Paid employment	0.59	–
Unpaid employment	0.01	–
Not seeking work	0.06	–
Leave	0.01	–
Retired	0.22	–
Student	0.03	–
Unemployed	0.05	–
Employed other	0.03	–
Yr11	0.12	0.16
Yr12	0.15	0.16
Certificate	0.18	0.16
Diploma	0.12	0.09
Graduate certificate	0.07	
Bachelor's degree	0.26	0.22
Post Graduate degree	0.09	
Education other	0.13	0.37
Number of cars	1.76	1.80

Respondents were randomly assigned to different treatment groups as part of the study. Firstly, of the 2490 respondents used in the analysis in this paper, 1242 were assigned to a treatment group in which they were asked to complete a reflective type of attitude scale, with the remaining 1248 respondents asked to complete a formative scale. Respondents were further randomly assigned to one of three different information states. A third of respondents were presented with a screen prior at the outset of the survey presenting negatively framed information about electric vehicles (see Fig. 4) with another third shown a screen with positively framed information (Fig. 5). The last third of respondents received no such information framing during the survey, and hence constitute a control group. The final breakdown of the number of respondents assigned to each experimental condition is given in Table 3.

After having to either read negative, positive or no information about electric vehicles, respondents were exposed to a discrete choice experiment, consisting of five tasks describing three hypothetical vehicles and a no choice alternative. Each vehicle was described by five attributes, including the type of vehicle (hybrid or battery electric), the charging time required to completely charge the vehicle, and the vehicles driving range based on a fully charged battery. Each vehicle was further described by its purchase price and average weekly running costs. Fig. 6 provides an example screen from the discrete choice experiment. The attributes and attribute levels used in the design are reported in Table 4. Respondents were provided with a detailed description of the task and attributes before being asked to complete the experiment.

A Bayesian D-optimal design was constructed using Ngene Software (www.choice-metrics.com). One thousand Sobol draws (see Hensher et al., 2015) were used to simulate uninformative priors that were employed to generate the design based on (see Rose and Bliemer 2009). The final design consisted of 20 choice tasks, which was blocked into four blocks of five tasks each, with respondents randomly assigned to one block during the survey.

Immediately after completing the discrete choice experiment, respondents had to complete a battery of attitudinal type questions that were either reflective or formative in nature. For respondents assigned to the reflective questionnaire, they were asked how strongly they agreed or disagreed with 15 statements based on the New Ecological Paradigm (NEP) scale developed by Dunlap and Van Liere (1978) and later refined by Dunlap et al. (2000). Using a seven-point Likert scale (from totally disagree to totally agree), the NEP scale consists of eight pro-environmental attitude (NEP+) items and seven non-pro or anti-environmental attitude (NEP-) items. The fifteen items are reported in Table 5, alongside the mean, median and standard deviation of the responses provided to each statement.

Those who were not assigned to the NEP scale had to complete the Modified Innovative Resistance (MIR) scale developed by Ram and Sheth (1989). The MIR seven-point scale represents a formative measurement designed to measure an individual's level of resistance to innovation and technology adoption. It consists of five statements that relate to five different dimensions such as usage, value, risk, tradition and image. In this way, the scale is designed such that the indicators inform a respondent's resistance to innovation latent construct, rather than the latent construct of resistance to innovation driving the respondent's responses to the indicator variables. Consequently, the scale is formative as opposed to reflective in nature. Table 6 shows the five statements of the scale, alongside descriptive statistics of the responses generated from the scale in the sample.

At this point, it is worth noting that two scales used to represent formative and reflective measures relate to different concepts, as well as differing in terms of the number of items used. The NEP scale has 15 items, and aims to capture information about a person's attitudes towards the environment. The MIR scale, on the other hand, uses only five items and measures a person's resistance to innovation. We see these differences as being immaterial to the current study given the objectives of the paper, which is to first

IS THE HYPE FOR ELECTRIC VEHICLES JUSTIFIED?

Scientists worldwide agree that **transport is one of the main causes of greenhouse gas emissions** into the atmosphere.

Technological advancements can mitigate the negative impact on the environment, however **there is no solid** evidence **that battery electric vehicles** (or commonly, electric vehicles or EVs) **are the solution to this issue**. Battery electric vehicles are vehicles with motors that are powered by electricity rather than liquid fuel.

The scientific community lists few arguments against the (profitable) link EVs = sustainability.

First, EVs are charged via the electricity grid, which in most cases distributes **energy obtained from non-renewable sources**. On average, in Australia less than 20% of electricity on the grid comes from renewable sources. This means that about **80% of the electricity used to power a "supposedly-clean" EV is generated by burning coal, gas or oil**.

Another related issue is **the reliance on the electric grid in specific times of the day**. Most EVs owners charge their vehicle at evening, when the amount of green electricity (derived from renewable sources) is at its lowest peak. This means that **the electricty used to charge the vehicles will most likely be coming from burning coal, gas or oil**. Also, the electricity demand will spike during these hours if all the owners of EVs plug in their cars, with **the risk of affecting the provision to everyone connected to the grid (e.g., more frequent blackouts, higher costs)**.

Finally, the batteries used in these vehicles will most likely need to be replaced during the vehicle lifespan and the repurposing of these batteries is very complicated. **The disposal of longer-lasting lithium-ion batteries is potentially very toxic for the environment**.

Fig. 4. Negative information frame.

THE SOCIETAL BENEFITS OF ELECTRIC VEHICLES

Scientists worldwide agree that **transport is one of the main causes of greenhouse gas emissions** into the atmosphere.

Technological advancements in this field can mitigate the negative impact on the environment. Battery electric vehicles (or commonly, electric vehicles or EVs) are vehicles with motors that are powered by electricity rather than liquid fuel.

Compared with traditional gasoline vehicles, **EVs have several societal benefits**. The most important one is **environmental**, which is achieved thanks to **carbon emissions reduction**.

Battery electric vehicles, unlike combustion engines (e.g., engines powered with gasoline or diesel) which convert on average 20-25% of energy into propulsion, are able to convert between 80 to 90% of the energy into propulsion.

Other societal benefits include, **costs saving over life of vehicle** (e.g., fuel, operation and maintanance), **health impact** (EVs do not emit fine particle pollution typical of combustion engine vehicles), the opportunity of **economic develoment** (e.g., direct job creation in the auto industry in manufacturing, research and development, battery manufactuuring, installation and maintenance of EVs supply equipment), **enhancement of energy security** (reducing reliance on exported energy from other countries).

Fig. 5. Positive information frame.

Table 3
Number of respondents assigned to each experimental condition.

	Negative information	Positive information	Control group	Total sample
Reflective	421	410	411	1 242
Formative	414	422	412	1 248

introduce the wider choice modelling community that not all scales are reflective in nature, and second that formative measures require different modelling than do reflective ones. Thirdly, we seek to demonstrate how the different scales can be used in practice when paired with discrete choice models. As such, the paper does not set out to directly compare and contrast the two scale approaches within a single empirical study. Such a study, we leave to future research.

4. Results

In this section, we present the results of the different models estimated on the empirical data collected. All models were estimated using PandasBiogeme (Bierlaire 2020) using maximum likelihood estimation techniques to estimate the parameters. The discrete choice models all assume an error component functional form, with the error component attached to the utility functions of the three vehicle alternatives. The remaining parameters of the model are treated as fixed. All models were estimated simultaneously with 1000 MLHS draws being employed to simulate the error components attached to the discrete choice models and error terms associated with the latent variable component of the attitude models. In what follows we begin with a discussion of the model estimated on the sample of respondents assigned to complete the reflective measurement scale, after which we present formative composite measure results.

4.1. Reflective measure results

Table 7 reports the findings of the MIMIC HCM discussed in Section 2.1 and applied to the data for which respondents completed the reflective NEP scale. The NEP scale is theoretically divided into two parts, one in which the questions are framed to capture pro environmental attitudes, and the other to capture anti-environmental attitudes. Whilst within the survey itself, the questions are designed to alternate between positive and negative wording frames, the two sets of questions are intended to relate to separate latent constructs. With this in mind, the HCM reported here assumes two latent constructs, one for each set of questions. This gives rise to the estimation of two separate structural models, with both models comprising explanatory variables related to the respondent's gender, income, and age. Also, two dummy variables representing the various information treatment groups to which different respondents were randomly assigned as part of the study are included in both specifications, with these being one for respondents exposed to the negative information message, and the other for positive. Hence, the dummy variables are meant to be interpreted relative to the base control group.

All parameters in the two structural models are statistically significant. For the first structural model, defining a pro-environmental attitude, the parameters are all positive, with the opposite being found for the second construct, which defines an anti-environmental

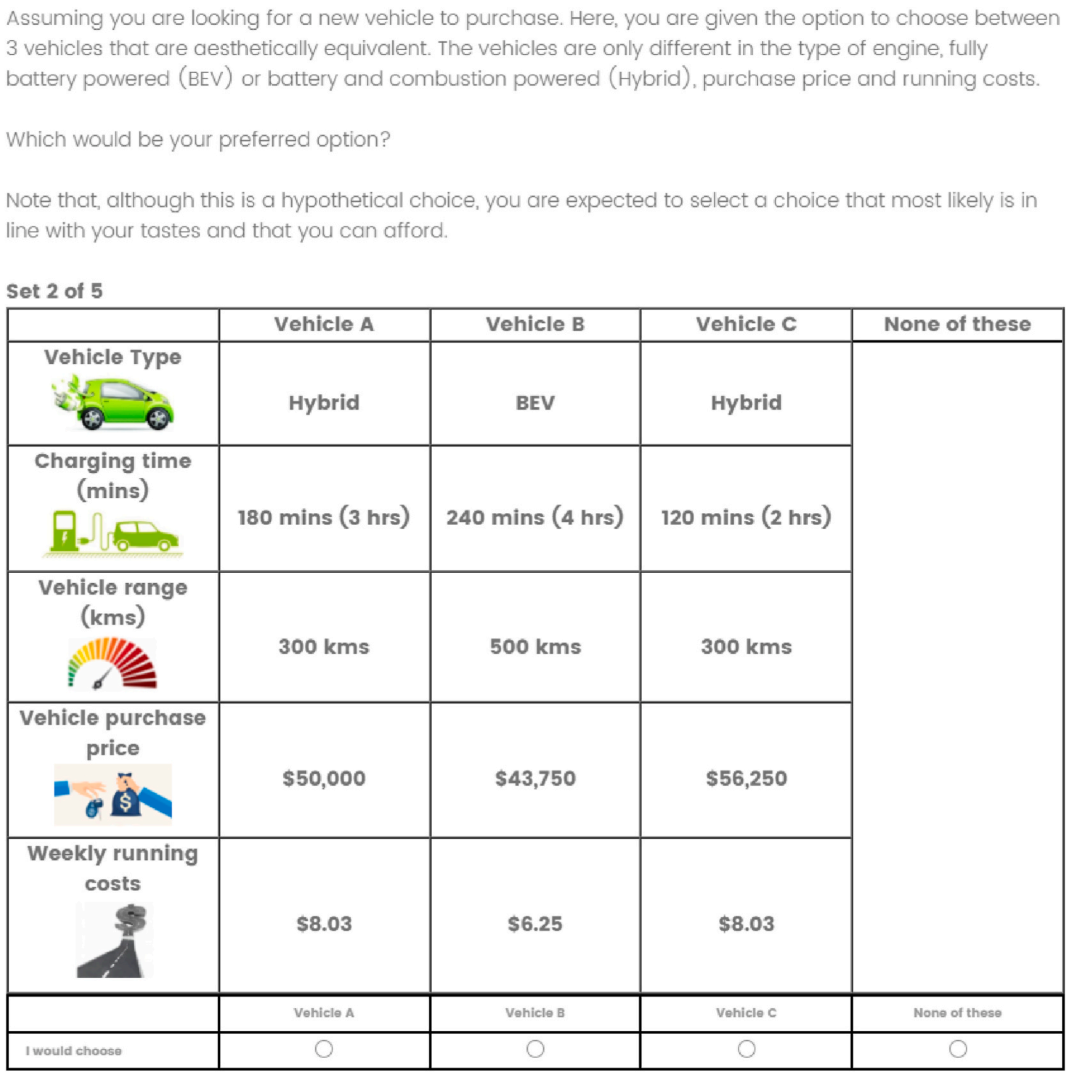


Fig. 6. Example discrete choice experiment screen.

Table 4
Attributes and attribute levels.

Attributes	Attribute levels
Vehicle type	Hybrid, Battery electric
Recharging time (mins)	15, 30, 60, 120, 180, 240, 300, 360, 420, 480
Driving range (kms)	200, 300, 400, 500, 600
Purchase price (Au$)	37500, 43750, 50000, 56250, 62500
Weekly running costs (Au$)	5.36, 6.25, 7.14, 8.03, 8.93

attitude. Ignoring the error terms of the structural models, the deterministic components of the models therefore allow for only positive values for the first latent construct, and negative values for second latent variable. With respect to the first latent construct, females, higher income earners and older respondents are expected to have more positive values for the latent construct, all else being equal. For the second latent construct, the same demographic groups are predicted to have more negative values than other types of respondents. Further, those exposed to the different information messages have statistically significantly different latent variables than those assigned to the no information control group (we discuss this in Section 5 when we apply the models). It is important to highlight, however, that the latent constructs derived from the structural models should not be interpreted without reference to the measurement models that they are linked to or the estimates obtained from those same measurement models. This is because the indicator variables allow one to translate the meaning of the latent variable (e.g., if the indicator variables are all about the environment, then the latent

Table 5
NEP reflective indicators.

Pro-Environmental Attitude Measurement Items (NEP+)	Average	1	2	3	4	5	6	7
We are approaching the limit of the number of people the Earth can support	4.84	0.067	0.057	0.070	0.196	0.226	0.153	0.230
When humans interfere with nature it often produces disastrous consequences	5.43	0.020	0.030	0.052	0.161	0.199	0.209	0.329
Humans are seriously abusing the environment	5.33	0.029	0.026	0.069	0.143	0.239	0.180	0.315
Plants and animals have as much right as humans to exist	5.74	0.014	0.025	0.039	0.118	0.168	0.202	0.434
Despite our special abilities, humans are still subject to the laws of nature	5.69	0.010	0.013	0.033	0.126	0.219	0.230	0.368
The Earth is like a spaceship with very limited room and resources	4.89	0.041	0.046	0.093	0.215	0.216	0.183	0.207
The balance of nature is very delicate and easily upset	5.36	0.019	0.016	0.056	0.176	0.240	0.208	0.285
If things continue on their present course, we will soon experience a major ecological catastrophe	5.27	0.034	0.033	0.064	0.177	0.194	0.192	0.306

Anti-Environmental Attitude Measurement Items (NEP-)	Average	1	2	3	4	5	6	7
Humans have the right to modify the natural environment to suit their needs	3.47	0.174	0.157	0.163	0.234	0.139	0.066	0.067
Human ingenuity will ensure that we do not make the Earth unliveable	4.18	0.079	0.086	0.134	0.288	0.196	0.126	0.091
The Earth has plenty of natural resources if we just learn how to develop them	4.75	0.059	0.060	0.093	0.205	0.225	0.157	0.201
The balance of nature is strong enough to cope with the impacts of modern industrial nations	3.54	0.176	0.143	0.186	0.192	0.146	0.082	0.075
The so-called “ecological crisis” facing humankind has been greatly exaggerated	3.44	0.232	0.135	0.131	0.196	0.144	0.086	0.076
Humans were meant to rule over the rest of nature	3.27	0.260	0.136	0.143	0.198	0.126	0.077	0.060
Humans will eventually learn enough about how nature works to be able to control it	3.93	0.114	0.106	0.143	0.271	0.185	0.105	0.076

Table 6
Modified Innovation Resistance Theory formative indicators.

Statement	Average	1	2	3	4	5	6	7
In my opinion, new technology is less reliable than well-established one	4.05	0.076	0.107	0.167	0.266	0.195	0.107	0.083
Newly developed products are always too expensive for the benefit they produce	4.92	0.028	0.051	0.089	0.197	0.263	0.184	0.188
I fear that newly developed technology may be risky	4.46	0.034	0.083	0.127	0.262	0.242	0.144	0.108
It is extremely hard to get used to the idea of new technology replacing the current one	4.01	0.094	0.137	0.151	0.215	0.188	0.123	0.093
In my opinion, new technology is often too complicated to be useful	3.97	0.093	0.139	0.159	0.220	0.183	0.111	0.093
In my opinion, new technology is less reliable than well-established one	4.05	0.076	0.107	0.167	0.266	0.195	0.107	0.083

construct is likely to be environmental in nature) and the signs of the parameters from the measurement models are unconstrained such that a more positive latent variable may actually result in a lower probability of responding more positively on an indicator variable, if the parameter associated with that indicator variable in the measurement model is negative.

As the indicators in the survey were all measured using a seven-point Likert rating scale, we estimate the measurement models using ordered logit models. The first latent construct is employed as an explanatory variable to explain the eight positive indicators of the NEP scale. Examination of the parameters of the first measurement model associated with each indicator reveal that they are all positive and statistically significant. This suggests that as the underlying latent variable becomes more positive, a person is predicted to provide a higher score for each indicator variable. Hence, given the structural model, all else being equal, females, higher income earners and older respondents are predicted to have a greater probability of providing higher scores on the indicator variables measuring pro-environmental attitudes.

The second latent variable is included as an explanatory variable in models utilized to explain negative environmental indicator variables of the scale. The results of the second measurement model suggest that the underlying latent variable is positively related to each of the eight items associated with positive attitudes towards the environment. Because latent variable derived from the deterministic component of the second structural model are constrained to always be negative, interpretation of the second latent construct

Table 7
Reflective HCM (MIMIC) results.

Variable	Par.	(rob *t*-rat.)
Structural model		
Pro environmental attitudes		
Female	0.809	(6.92)
Income	9.10×10^{-05}	(3.71)
Age	0.026	(9.99)
Treatment group 1 (Positive information message)	0.397	(2.89)
Treatment group 2 (Negative information message)	0.365	(2.51)
Sigma	1.687	(20.09)
Anti-environmental attitudes		
Female	−1.027	(-7.08)
Income	-1.3×10^{-04}	(-3.87)
Age	−0.034	(-10.93)
Treatment group 1 (Positive information message)	−0.193	(-2.11)
Treatment group 2 (Negative information message)	−0.204	(-2.09)
Sigma	2.266	(24.23)
Measurement model		
Pro environmental attitudes		
We are approaching the limit of the number of people the earth can support	1.000	–
When humans interfere with nature it often produces disastrous consequences	1.391	(30.21)
Humans are severely abusing the environment	1.355	(30.15)
Plants and animals have as much right as humans to exist	1.643	(28.18)
Despite our special abilities humans are still subject to the laws of nature	1.477	(31.59)
The earth is like a spaceship with very limited room and resources	1.014	(31.93)
The balance of nature is very delicate and easily upset	1.308	(31.27)
If things continue on their present course, we will soon experience a major ecological catastrophe	1.329	(29.41)
Threshold1 (+ve)	−2.081	(-11.50)
Threshold2 (+ve)	−1.162	(13.97)
Threshold3 (+ve)	−0.187	(18.20)
Threshold4 (+ve)	1.344	(28.27)
Threshold5 (+ve)	2.805	(35.49)
Threshold6 (+ve)	4.229	(33.85)
Anti-environmental attitudes		
Humans have the right to modify the natural environment to suit their needs	1.000	–
Human ingenuity will insure that we do NOT make the earth unliveable	0.683	(30.45)
The earth has plenty of natural resources if we just learn how to develop them	0.468	(14.52)
The balance of nature is strong enough to cope with the impacts of modern industrial nations	1.033	(44.89)
The so-called "ecological crisis" facing humankind has been greatly exaggerated	1.089	(39.43)
Humans were meant to rule over the rest of nature	1.154	(43.38)
Humans will eventually learn enough about how nature works to be able to control it	0.808	(37.07)
Threshold1 (-ve)	−5.124	(-31.39)
Threshold2 (-ve)	−3.963	(25.99)
Threshold3 (-ve)	−2.936	(28.18)
Threshold4 (-ve)	−1.482	(35.65)
Threshold5 (-ve)	−0.222	(30.25)
Threshold6 (-ve)	0.954	(21.69)
Discrete choice model		
No choice ASC	−6.127	(-8.70)
Attitude positive (no choice)	−0.273	(-2.78)
Attitude negative (no choice)	−0.157	(-2.21)
ASC1	−0.203	(-5.22)
ASC2	−0.024	(-0.65)
Battery electric vehicle	0.057	(2.49)
Vehicle recharging time	−0.002	(-11.57)
Vehicle range	0.003	(20.61)
Weekly running cost	−0.174	(-10.74)
Vehicle purchase price	-4.36×10^{-05}	(-20.21)
Error component (SP alternatives)	4.804	(14.34)
Model fit statistics		
LL(0)	−62481.967	
LL(B)	−36040.177	

(*continued on next page*)

Table 7 (*continued*)

Variable	Par.	(rob *t*-rat.)
ρ^2	0.423	
Adj. ρ^2	0.419	
AIC	72176.354	

needs to be considered in this light. From the second structural model, females, higher income earners and older respondents are predicted to larger negative values for the second latent construct relative to their counterparts. Given that respondents with a greater negative value for the latent construct are more likely to provide answers consistent with a response towards a lower end of the indicator variable scale, the model predicts that females, higher income earners and older respondents are more likely to report lower values for the indicator variables related to anti-environmental sentiments than other types of respondents, all else being equal.

Turning to the choice model, three alternative specific constants are estimated, two for the first two hypothetical alternatives (ASC1 and ASC2) and one for the no choice alternative. The large statistically significant negative constant for the no choice alternative suggests that all else being equal, respondents are more likely to select a hypothetical vehicle than not. The constant associated with the first of the hypothetical alternatives is significant and negative indicating that respondents are more likely to select the middle two alternatives than they are the first alternative. The parameters related to the attributes describing the various vehicles are all statistically significant and of the expected signs. Respondents prefer battery electric vehicles over hybrid models, and vehicles that take less time to charge over vehicles which require longer periods of time to recharge. Vehicles with greater driving ranges are preferred to those with lower ranges, and respondents prefer both a lower vehicle purchase price and lower weekly running costs.

The two latent variables derived from the structural models enter the discrete choice model as separate main effects within the utility function of the no choice alternative. The parameters for both constructs are negative, suggesting that positive increases in either of the latent constructs will result in a lower probability of selecting the no choice alternative, and hence a higher probability of selecting one of the three vehicles offered, all else being equal. Again, the signs of the latent constructs as well as the parameters in the discrete choice model need to be considered when interpreting this result. Based on the first structural model, on average, the first latent construct will be positive (i.e., ignoring the stochastic error term) with women, higher income earners and older respondents having more positive values for the latent variable relative to other respondents. Respondents with more pro-environmental attitudes (i.e., females, higher income earners and older respondents) are predicted to have a lower utility for the no choice alternative and hence be more inclined to select one of the vehicles on offer relative to the no choice alternative, all else being equal. On the other hand, based on the deterministic component of the second structural model, the second latent variable will always be negative, with females, higher income individuals and older respondents displaying more negative values for this construct relative to other respondents. Given this outcome, and based on the fact that the structural parameter in the discrete choice model associated with the second latent variable is negative, the model suggests that women, higher income earners and older respondents have a higher degree

Table 8
Formative Composite measures HCM results.

Variable	Par.	(rob *t*-rat.)
Discrete choice model		
No choice ASC	−13.139	(-12.07)
Female	0.845	(2.15)
Income	$-1x10^{-04}$	(-1.38)
Age	0.108	(8.68)
Treatment group 1 (Negative information message)	0.116	(0.26)
Treatment group 2 (Positive information message)	−0.750	(-2.61)
In my opinion, new technology is less reliable than well-established one	0.189	(2.24)
Newly developed products are always too expensive for the benefit they produce	0.295	(2.81)
I fear that newly developed technology may be risky	0.051	(0.29)
It is extremely hard to get used to the idea of new technology replacing the current one	−0.134	(-0.86)
In my opinion, new technology is often too complicated to be useful	0.403	(2.53)
ASC1	−0.109	(-2.64)
ASC2	0.016	(0.43)
Electric vehicle	0.002	(0.04)
Vehicle recharging time	−0.002	(-13.04)
Vehicle range	0.003	(21.55)
Weekly running cost	−0.164	(-9.99)
Vehicle purchase price	0.000	(-21.50)
Error component (SP alternatives)	4.877	(15.02)
Model fit statistics		
LL(0)	−8 650.477	
LL(B)	−6 596.37	
ρ^2	0.237	
Adj. ρ^2	0.234	
AIC	13242.740	

of disutility for the no choice alternative (i.e., a negative parameter multiplied by a larger negative latent variable), meaning these respondents have a higher probability of selecting one of the vehicles on offer relative to the no choice, similar to the positive environmental latent construct. From the above findings we can conclude that the two latent variables appear to be working in same direction with respect to vehicle purchase behaviour insofar as individuals who display more pro-environmental attitudes are less likely not to purchase a vehicle, as are those who display more negative attitudes.

4.2. Formative composite measure results

Table 8 outlines the results of the composite measure approach based on the segment of the sample who completed the formative MIR scale survey version. Within the model, the indicator variables enter the utility function for the no choice alternative directly as individual main effects, as per the composite model approach outlined in Section 2.2.1. For completeness, income, age, and a dummy for female, as well as dummies for the positive and negative information states (relative to the base of no information) are also incorporated directly into the utility function for the no choice alternative. As with the reflective HCM, three ASCs are estimated, one each for the first two hypothetical vehicle alternatives, and one for the no choice alternative. The ASCs from the model suggest that respondents prefer to purchase a vehicle, all else being equal, but are less likely to select the first alternative relative to the middle two vehicles shown.

Overall, females and older respondents are more likely to select a vehicle whilst non female identifying individuals and younger respondents are more likely to select the no choice option over one of the vehicle offerings, ceteris paribus. The income parameter is negative but not statistically significant. Examining the parameters for the MIR indicator variables, two indicators are not statistically significant, with the remaining three being significant and positive. It is worth noting that the wording of the items in the survey is negatively framed suggesting that higher scores on these items reflect a greater preponderance to innovation resistance. As such, the model points out that respondents who are more resistant to innovativeness, are less likely to select one of the car options, ceteris paribus. We discuss this further in Section 5. Examining the information treatment groups it emerges that respondents who were tasked with reading a positive review of electric vehicles are less likely to select the no choice alternative, whereas those who observed the negative information frame have the same likelihood as the control group of selecting the no choice option.

The parameters of the remaining attributes are as expected, however unlike the sample exposed to the reflective attitude scale, the

Table 9
Formative causal measures HCM results.

Variable	Par.	(rob *t*-rat.)
Factor analysis		
In my opinion, new technology is less reliable than well-established one	0.729	(26.40)
Newly developed products are always too expensive for the benefit they produce	0.634	(21.91)
I fear that newly developed technology may be risky	0.769	(28.37)
It is extremely hard to get used to the idea of new technology replacing the current one	0.796	(33.32)
In my opinion, new technology is often too complicated to be useful	0.829	(36.08)
Sigma1	0.694	(32.12)
Sigma2	0.779	(39.61)
Sigma3	0.652	(30.96)
Sigma4	0.621	(26.85)
Sigma5	0.575	(25.35)
Discrete choice model		
No choice ASC	−10.015	(-15.15)
Lambda	0.077	(4.96)
Female	1.108	(3.45)
Income	-2.12×10^{-04}	(-3.10)
Age	0.124	(11.68)
Treatment group 1 (Negative information message)	−0.592	(-1.57)
Treatment group 2 (Positive information message)	1.098	(2.84)
ASC1	−0.108	(-2.62)
ASC2	0.017	(0.45)
Electric vehicle	0.001	(0.03)
Vehicle recharging time	−0.002	(-13.05)
Vehicle range	0.003	(21.55)
Weekly running cost	−0.164	(-9.98)
Vehicle purchase price	-4.45×10^{-05}	(-21.50)
Error component (SP alternatives)	5.144	(14.75)
Model fit statistics		
LL(0)	−17502.153	
LL(B)	−14144.566	
ρ^2	0.192	
Adj. ρ^2	0.189	
AIC	28339.132	

sample appear to be indifferent between battery electric vehicles and hybrid models. As anticipated, respondents prefer vehicles with longer driving ranges, lower recharging times, and lower purchase and running costs.

4.3. Formative factor analytic casual measure results

The final model reported is estimated on the data for which respondents were asked to complete the MIR formative attitude scale. This is the same data used to estimate the model reported in Section 4.2. For the model reported in Table 9, we factor analyse the five items assuming a single factor in order to derive the factor loadings, from which we construct a non-refined weighted sum score to form the latent construct for innovative resistance. This model approach is discussed in Section 2.2.2. The latent score then enters the utility function for the no choice alternative in the discrete choice model as a main effect. Gender, age, and income also enter the utility function for the no choice alternative, as do dummies for the positive and negative information treatment groups. The factor analysis and discrete choice model are optimized jointly using maximum likelihood simulation techniques.

The top section of Table 9 presents the factor loadings for the five MIR indicator variables obtained from the factor analysis. The factor loadings for all five indictors are statistically significant. The sigma parameters reported within the table are associated with the random normal terms used to estimate the factor analysis component of the model. The effect of the composite weighted score generated from the factor analysis is measured by the parameter lambda. From the table we can conclude that lambda is statistically significant and positive, suggesting that higher scores on the MIR scale increase the probability of selecting the no choice alternative. Hence, individuals who are more resistant to innovation are less likely to purchase an electric vehicle all else being equal. This finding aligns with the composite score approach discussed in Section 4.2. The remaining estimates are consistent with the model reported in Table 9. Respondents are largely indifferent between battery electric vehicles and hybrid models, prefer vehicles with longer driving ranges, lower recharging times, and lower purchase and running costs.

5. Application of the models

Whilst it is not an objective of the paper to directly compare how formative versus reflective measurement scales impact the results of discrete choice experiments, in this section, we demonstrate how each model can be operationalized so as to exhibit the differences in outputs each type of measure produces, and hence policy impacts that can be explored using each approach. To do so, let us consider three hypothetical individuals selecting between purchasing one of three vehicles, or no vehicle at all. Persons 1 and 3 identify as females, whilst person 2 identifies as not female. Person 1 has a weekly gross income of $800 and is 20 years of age, person 2 $400 and is 50 years old, and person 3 $1000 and is 40. The sociodemographic and economic characteristics of each person are presented in Table 10. The vehicle attributes for the assumed scenario are given in Table 11, instead.

5.1. Reflective measures model outputs

Given the scenario as outlined, it is possible to compute the choice probabilities for the individuals with respect to selecting one of the assumed vehicles. Table 12 displays the predicted choice probabilities based on the reflective model estimates for each of the three hypothetical decision makers under each of the three information states. At the base of the table are the average latent variable estimates for each individual. The choice probabilities are generated using 5000 Sobol draws to simulate the various error terms within the model. Based on Tables 10 and 11, the model predicts that the same individual receiving either negative or positive information about the use of electric vehicles is less likely to select the no choice alternative relative to not receiving any information. This represents a somewhat counterintuitive result, although it is worth noting that those who receiving positive information are predicted to have a slightly lower probability of no choice alternative relative to those who were exposed to the negative information state. It is possible that respondents who receive no information in a DCE are more inclined than those who do, whether such information is framed positively or negatively, to stick with the status quo. Interestingly, differences across the socio-demographic characteristics are predicted to have only limited impact on the choice probabilities, with differences observed only after the third decimal point. Hence, whilst the model itself suggests that attitudes play a role in vehicle choice, application of the model suggests that their influence is only mild.

The traditional HCM based on a reflective measurement scale results in the estimation of one or more latent variables which directionally feed into the discrete choice model, but are also used as explanatory variables to predict the responses given to the various indicator variables used as part of the measure. As such, whilst data on the indicators are used in model estimation, they also represent dependent variables to be predicted within the overall model framework. Table 13 presents the predicted scores for each indicator for the three hypothetical individuals under the three different information treatment states for two statements only (one each from the positive and anti-environmental statements). The first point to note is that for both selected statements, respondents are predicted to

Table 10
Assumed characteristics of individual decision makers.

Characteristic	Person 1	Person 2	Person 3
Female	1	0	1
Income	800	400	1 000
Age	20	50	40

Table 11
Vehicle attributes for scenarios 1 and 2.

Vehicle	A	B	C
Electric/hybrid (EV = 1/Hybrid = 0)	0	1	0
Recharge Time (mins)	30	60	15
Range (kms)	400	400	500
Purchase price ($)	48,000	50,000	52,000
Weekly running cost ($)	8	6.5	8.5

Table 12
Predicted choices by person under different information states reflective model.

		Vehicle A	Vehicle B	Vehicle C	None
Person 1	No information group	0.2572	0.3079	0.2642	0.1707
	+ve information group	0.2584	0.3093	0.2654	0.1669
	-ve information group	0.2582	0.3091	0.2653	0.1674
Person 2	No information group	0.2571	0.3077	0.2641	0.1711
	+ve information group	0.2583	0.3092	0.2653	0.1672
	-ve information group	0.2581	0.3090	0.2652	0.1677
Person 3	No information group	0.2578	0.3085	0.2648	0.1689
	+ve information group	0.2590	0.3100	0.2660	0.1651
	-ve information group	0.2588	0.3098	0.2658	0.1656

select higher score values than lower suggesting that for each of the three hypothetical respondents, they are likely to hold both positive and negative sentiments towards the environment simultaneously (i.e., they are ambivalent). The second thing to note from the table is that the model predicts that the provision of information, either positive or negative, results in respondents selecting a higher score for the positive statement relative to the no information state, but lower scores for the negative statement relative to the no information state. As with the vehicle choice prediction, this result is somewhat counterintuitive insofar one would expect respondents receiving negative information to provide less positive scores for the positive statements and higher scores for the negative statement, all else being equal. The third point to note is that the impact of the information state on the predicted score provided is marginal, particularly with respect to the negative score. That is to say, the choice probabilities for the positive information state and negative information state only changes the indicator values for the second respondent with respect to the negative statement shown (other

Table 13
Predicted attitudinal responses under different information states reflective model.

Person 1	Humans are severely abusing the environment	1	2	3	4	5	6	7
	No information group	0.084	0.059	0.090	0.189	0.201	0.167	0.210
	+ve information group	0.059	0.046	0.074	0.170	0.200	0.183	0.268
	-ve information group	0.061	0.047	0.075	0.172	0.200	0.182	0.263
	The earth has plenty of natural resources if we just learn how to develop them	1	2	3	4	5	6	7
	No information group	0.023	0.042	0.084	0.224	0.251	0.192	0.184
	+ve information group	0.025	0.046	0.089	0.232	0.251	0.186	0.172
	-ve information group	0.025	0.046	0.089	0.232	0.251	0.186	0.172
Person 2	Humans are severely abusing the environment	1	2	3	4	5	6	7
	No information group	0.088	0.061	0.092	0.191	0.201	0.165	0.202
	+ve information group	0.062	0.048	0.076	0.173	0.200	0.181	0.259
	-ve information group	0.064	0.049	0.078	0.175	0.201	0.180	0.254
	The earth has plenty of natural resources if we just learn how to develop them	1	2	3	4	5	6	7
	No information group	0.022	0.042	0.083	0.222	0.251	0.194	0.187
	+ve information group	0.024	0.045	0.088	0.229	0.251	0.188	0.176
	-ve information group	0.024	0.045	0.088	0.230	0.251	0.187	0.175
Person 3	Humans are severely abusing the environment	1	2	3	4	5	6	7
	No information group	0.052	0.042	0.069	0.162	0.197	0.187	0.291
	+ve information group	0.035	0.031	0.054	0.139	0.187	0.195	0.358
	-ve information group	0.037	0.032	0.055	0.141	0.188	0.195	0.352
	The earth has plenty of natural resources if we just learn how to develop them	1	2	3	4	5	6	7
	No information group	0.031	0.055	0.103	0.249	0.247	0.170	0.145
	+ve information group	0.034	0.059	0.109	0.255	0.245	0.163	0.135
	-ve information group	0.034	0.059	0.109	0.256	0.245	0.163	0.135

respondents' probabilities are observed to change in the fourth decimal point and are hence not reflected in the table). Although differences are also observed with respect to the first selected positive attitude statement, again, the choice probability changes are marginal. Where differences are observed, is across the different respondents, suggesting that the rating score predictions are being driven mainly by socio-demographic differences than anything else.

In short, as with the vehicle choice prediction model, the model results suggest that the latent attitudes from the model predict the indicator scores provided, however the information states provide counterintuitive outcomes. This highlights an important lesson for the use of HCMs and how they are reported within the literature. That is, if one were simply to report the parameter estimates from the model (as in Table 7), one would conclude that attitudes play a statistically significant role in the choices observed in the DCE. However, in applying the model (albeit in such a limited manner) reveals that the actual inclusion of attitudes in the model can play a limited, and actual counter-intuitive role in the predicted outcomes of the model.

5.2. Formative measures composite model outputs

With the formative measures, the directional relationship of the indicators and latent constructs is reversed. As such to operationalise the model, one requires advanced knowledge of the values the indicator variables will take, so as to derive a measure for the latent variable, and subsequently choice behaviour. Table 14 presents two sets of scores for the indicator variables, one based on the data median values and the second a scenario where some intervention is assumed to decrease the values given by respondents (i.e., respondents become less resistant to innovation) by a constant factor of 2.

In the composite score model formulation, the indicator variables enter directly as main effects into the no choice alternative. In addition to the MIR indicator variables, the no choice utility function also includes socio-demographic characteristics related to the gender, age and income of the respondent. Assuming the same individuals and vehicles from Section 5.1, Table 15 presents the predicted choice behaviour of all three individuals assuming the median indicator variable values and the new values under the low resistance scenario. The results shown in Table 15 demonstrate the relationship between the MIR measurement scale and choice behaviour. As shown in the table, any intervention that makes a respondent less resistant to innovativeness is predicted to increase their probability of purchasing an electric vehicle, independent of any information condition to which they belong.

Care needs to be given however in interpreting the information state outcomes. Unlike the reflective model, the information state conditions and socio-demographic variables enter into the choice model directly as opposed to describing the respondent's attitudes. This represents a different decision process than was previously assumed, and hence any direct comparison with the previous model's outputs should be taken with extreme care. It is worth noting, however, that although not explored herein, it is possible to estimate models to predict the indicator responses and use the predicted outcomes to enter the choice behaviour. Under such a model framework, the information state treatment dummies and socio-demographic variables could be used as explanatory variables explaining the observed indicator variable outcomes, similar to the MIMIC approach employed in the traditional HCM framework.

5.3. Formative measures factor analytical causal model outputs

Table 16 provides the predicted choice behaviour for the factor analytic formative measurement model under the same set of assumptions described in Sections 5.1 and 5.2. As with the composite measures model, the indicator values are inputs rather than estimated outputs of the model process. Although the choice shares appear to differ from those obtained from the composite measures model, the direction of outcomes is similar. That is, any intervention resulting in a respondent becoming less resistant to innovativeness will directly result in an increase in the probability that they will purchase an electric vehicle, holding all else equal. As with the composite measures model, the treatment group dummy variables and socio-demographic variables enter the utility directly as opposed to via the modelled attitudes. As noted in Section 5.3, such an approach is not necessary insofar as it is possible to predict the indicator variables and replace these in the model application, thus allowing for a greater exploration of the policy impacts on attitudes and choice.

6. Discusion and concluding comments

Both the composite measures and factor analytic models demonstrate the key difference between the outputs derived from the use of formative versus reflective attitude measures. The use of reflective indicators and the modelling of such data allows the analyst to assess how attitudes impact both preferences and the response provided to indicator variables. As noted by Chorus and Kroesen (2014) however, unless respondents are assigned to different treatment groups as done here, the typical use of cross-sectional data alongside

Table 14
Modified Innovation Resistance Theory formative indicators assumed values.

Statement	Median	Scenario
In my opinion, new technology is less reliable than well-established one	4	2
Newly developed products are always too expensive for the benefit they produce	5	3
I fear that newly developed technology may be risky	4	2
It is extremely hard to get used to the idea of new technology replacing the current one	4	2
In my opinion, new technology is often too complicated to be useful	4	2

Table 15
Predicted choice shares composite model under different information states and indicator variable scores.

Median indicator values assumed									
	No information			Neg. information			Pos. information		
Vehicle	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3
A	0.2631	0.2193	0.2256	0.2613	0.2169	0.2233	0.2736	0.2343	0.2400
B	0.2909	0.2425	0.2495	0.2889	0.2398	0.2469	0.3025	0.2590	0.2654
C	0.2723	0.2271	0.2336	0.2705	0.2246	0.2311	0.2832	0.2425	0.2485
None	0.1737	0.3111	0.2914	0.1792	0.3187	0.2987	0.1407	0.2642	0.2460
New indicator values assumed									
	No information			**Neg. information**			**Pos. information**		
Vehicle	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3
A	0.2886	0.2573	0.2621	0.2952	0.2686	0.2728	0.2874	0.2555	0.2604
B	0.3191	0.2845	0.2899	0.3264	0.2970	0.3016	0.3178	0.2825	0.2879
C	0.2987	0.2664	0.2714	0.3056	0.2780	0.2824	0.2975	0.2645	0.2695
None	0.0937	0.1917	0.1766	0.0727	0.1564	0.1432	0.0973	0.1976	0.1822

Table 16
Predicted choice shares factor analytic model under different information states and indicator variable scores.

Median indicator values assumed									
	No information			Neg. information			Pos. information		
Vehicle	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3
A	0.2596	0.2315	0.2345	0.2649	0.2382	0.2410	0.2489	0.2186	0.2217
B	0.2869	0.2559	0.2592	0.2928	0.2633	0.2664	0.2751	0.2416	0.2451
C	0.2687	0.2397	0.2428	0.2742	0.2466	0.2495	0.2577	0.2263	0.2295
None	0.1848	0.2728	0.2635	0.1681	0.2519	0.2430	0.2183	0.3136	0.3037
New indicator values assumed									
	No information			**Neg. information**			**Pos. information**		
Vehicle	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3	Person 1	Person 2	Person 3
A	0.2842	0.2542	0.2577	0.2892	0.2619	0.2650	0.2735	0.2386	0.2424
B	0.3142	0.2810	0.2848	0.3196	0.2894	0.2929	0.3023	0.2637	0.2680
C	0.2942	0.2632	0.2667	0.2994	0.2711	0.2744	0.2831	0.2470	0.2510
None	0.1074	0.2016	0.1908	0.0918	0.1776	0.1677	0.1411	0.2508	0.2386

the use of socio-demographic variables as explanatory variables means that such models are likely to be limited in terms of policy relevancy. Outside of predicting how attitudes are likely to change as population demographic characteristics change over time, the traditional use of HCMs does not allow for an understanding of how attitudes may change temporally. Note that this is not a criticism of the HCM itself, nor the use of reflective measures, but rather the data used in such analysis. Two solutions are possible to correct this issue. Firstly, as noted by Chorus and Kroesen (2014), the use of time series data where changes in attitudes can be observed over time, will allow for models that can predict attitudinal changes that may provide policy relevant insights. The collection of time-series data may be prohibitively expensive however, and the amount of panel waves required to detect attitudinal changes extensive. The second approach explored herein refers to utilise cross-sectional data but assign respondents to different treatment conditions. In the current paper, respondents were assigned to surveys in which they were exposed to either positive, negative or no information about electric vehicles.

The assigning of respondents to different treatment conditions may however be problematic in and of itself. In the current paper, the experimental conditions to which respondents were assigned occurred at the commencement of the survey, meaning that the information (or lack thereof) was provided prior to both the discrete choice experiment and the battery of attitudinal indicators. This represents a significant limitation of the current paper, as one cannot disentangle whether the information content impacted attitudes, vehicle preference or both simultaneously. Thus, whilst the reflective model assumes that the information state directly impacts attitudes and indirectly choice behaviour, and the formative models that the information state influences choice behaviour and not attitudes, it is possible that neither assumption is correct. Within this context of application, we are able to ignore such an issue given the papers' main set of objectives, which are to first introduce the concept of formative versus reflective measurement scales to the choice modelling community, and secondly to simply demonstrate how such data can be modelled. The inclusion of different information states within the data reflects a secondary consideration of the paper designed to demonstrate that treatment conditions can be experimentally introduced into cross sectional data dealing with attitudes and choice. As such, we advocate future studies that apply this approach also vary where such information is introduced within the survey.

As for the criticisms levelled at research including latent variables with a choice modelling framework, a further discussion is needed. Whilst we agree with the three principal criticisms levelled by Chorus and Kroesen (2014) towards the inclusion of latent variables within discrete choice models, it is worth noting that examples they provide are somewhat unhelpful, particularly to the first issue of endogeneity between attitudes and choice. Firstly, Chorus and Kroesen seem to consider attitudes and perceptions to be interchangeable. As pointed out by Bahamonde-Birke et al. (2017) and Borriello and Rose (2021), and described by McFadden (1986), perceptions and attitudes are very different concepts with the former appearing to be more closely linked to choices than the later. As such, we suggest that the issue of endogeneity is more likely to occur when perceptions are being measured as opposed to attitudes, although one cannot completely rule out endogeneity issues between attitudes and choice. In any case, the examples provided by Chorus and Kroesen all relate to perceptions rather than attitudes (e.g., perceptions around travel times). Chorus and Kroesen (2014) also argue that the issue exists in stated preference data sets, and suggest that latent attitudes may be influenced by the survey order, relative to the stated preference experiment, they are presented in. We would argue that whilst it is plausible, attitudes, as opposed to perceptions, should be more robust to such influences, and any impact on a latent construct caused by the survey is more an example of what psychologist's term demand induced artefacts (Orne 1959, 1969), as opposed to an actual impact on a person's underlying latent attitude.

The second criticism levelled at the current HCM framework by Chorus and Kroesen is related not so much at the technique itself as opposed to the data to which it is often estimated on. Currently, most if not nearly all models combining latent variables proxying attitudes and choice have been estimated on cross-sectional data. The use of cross-sectional data means that only between person comparisons are possible. The inability of such data to allow for within person comparisons precludes such analysis from understanding how a person, whose attitudes change over time, may also change their choice behaviour. Given the lack of covariation between the latent construct and choice data, Chorus and Kroesen argue that no causal inferences can be made. That is to say, the framework does not allow for an exploration of how attitudes can change over time within a person (other than linked to changes in any covariates used to describe the individual included within the model, such as age or income), and what impact such changes are likely to have on choice. This in turn makes any examination of policy impacts difficult as depending on the policy implemented, attitudes within a population might be affected. Chorus and Kroesen support this argument by citing Borsboom et al. (2003) who used the example that "If Lucy grows ten inches taller, she will become able to grab the book from the upper shelf" is equivalent to a statement of the type "If we replace Lucy by someone who is ten inches taller, that person would be able to grab the book from the upper shelf". We would argue that whilst valid, this is an example of a formative as opposed to reflective measure however.

The current paper deals with this issue in two ways. We do so firstly via the use of formative as opposed to reflective measures. With formative measures, the direction of the relationship from underlying construct to indicators is such that the indicators form the construct and not the other way around. Hence, even if one were to base their analysis on cross-sectional data, it is possible to explore how changes in the scores provided by respondents change assumed underlying latent constructs, and how such changes in turn influence observed choices. This can be done without any reference to covariates of individuals. That is, HCMs based on formative scales allow one to explore what the impact is of changing one or more indicator variables, not only on attitudes, but also choice. If one has external data on what exogenous impacts are likely to do to the indicator variables used, then one can explore the effects different policy settings will have on choice behaviour, even when the modelling has been done using cross-sectional data.

The second approach to allowing for a more nuanced exploration of policy settings on latent attitudes and choice given the presence of cross-sectional data introduced within this paper is the between subject exposure of respondents to different experiment conditions. In the current paper, respondents were assigned to one of three experimental conditions, one in which they exposed to positive information about electric vehicles, a second where negative information was provided, and a third where no information was provided (forming a control group). Such a process, even though applied to a cross-sectional survey, allows one to explore how different information states may influence attitudes. Of course, such a process may also suffer from endogeneity concerns similar to those raised by Chorus and Kroesen insofar as the treatment condition to which an individual is assigned may influence not just a person's attitude, but also more directly their choices. For this reason, additional research is required as to how and where such experimentally varied conditions should be introduced within the survey.

The third issue raised by Chorus and Kroesen is that respondents completing surveys do so not in isolation, and may interact with the survey in unexpected ways. That is, individuals may guess what research questions are being addressed by the researcher and provide answers that reflect what they expect is required of them, rather than what they actually think or feel about the topic. Alternatively, respondents may seek to align the answers they give to questions with the answers provided earlier in the survey. Whilst not our intention to diminish such concerns, although not addressed within the current paper, we argue that such issues relate to demand induced artefacts, and are much broader than attitude and choice studies. Indeed, we would argue that better survey design is required across many applied economic fields dealing with choice data, and that greater attention be given to literature on best survey practices.

As part of this paper, we hope to educate discrete choice modellers that not all attitude scales and indicators are reflective in nature. Indeed, as pointed out in the introduction, researchers have used formative measures in the past, but mistakenly treated them as if they were reflective. Such a mistake ignores the true directional causality of the scale and latent variable rendering the modelled outputs to be questionable. Aside from this, we note difficulties in how reflective measures may be used for prediction, and hence limit the outputs from such models in terms of policy outcomes derived. Whilst this issue can be overcome either via the use of time-series data or the use of experimental conditions within cross-sectional data, the problem may be less of an issue when formative scale measures are used. This is because one can explore how changes in the attitude indictor variables directly impact preferences, even if the sources of why the attitude changes are not known. In other words, it becomes possible to explore the impacts of different policy scenarios, even if the exact impact of the policy on indicators is known precisely at the time of the choice study. Thus, for example, in the current study, it was found that decreasing resistance to innovation increases the probability of individuals purchasing an electric vehicle. Armed with this knowledge, subsequent small-scale studies could be designed to test how innovative resistance can be reduced, with such studies undertaken independent of the original questionnaire.

The paper also draws attention to how HCMs are often reported within the literature. Often, researchers present the results of such models simply by reporting the parameter estimates obtained and their respective level of statistical significance. In this paper, we demonstrate, albeit in a limited manner, that even where the estimates of the model appear to make some form of logical sense, application of the model may result in counterintuitive outcomes. For example, in the current paper, we demonstrate that the provision of distinct information to different treatment groups can produce statistically significant behavioural estimates that when applied to different respondent types, can actually lead to counter-intuitive prediction outcomes. As such, we call for papers utilising such models to not just report the estimates obtained from the modelling exercise, but also test, either via simulation or some other mechanism, what impact the model has on predicted outcomes.

Finally, by drawing attention to reflective versus formative measures, we hope that the choice modelling community pay more attention to scale development and consider using a broader range of scale types. Firstly, we ask choice modellers to consider using only properly validated scales rather than simply writing statements and assuming that each statement properly measures some underlying latent attitude. Along these lines, we note that proper scale development with respect to reflective measures requires multiple items be used to measure underlying latent constructs, and that the use of single item measures is problematic for a number of reasons. Secondly, we note that many of the examples used to discuss the theoretical limitations of attitude scales appear to be framed using formative measures applied to models that assume reflective measures. Whilst we do not refute the issues raised, this misunderstanding has the potential to create issues where they may not actually exist. For example, many of the examples used by Chorus and Kroesen appear to use formative measures related to perceptions as opposed to reflective measures linked to attitudes. This misconception is interesting in and of itself as it may indicate that measures of perceptions may be better made using formative measures whilst attitude measurement may be better served using reflective measures. Far more research is required to support this position however, as well as to confirm the extent that the issues raised span both formative and reflective measurement types. Finally, the concept of formative versus reflective measures has the potential to impact different sections of the choice modelling literature. For example, Hess and Hensher (2013) promote using the HCM to model attribute non-attendance (ANA) using stated ANA data as the indicator variables within the model. Given that there exists only a single indicator variable related to the ANA for each attribute, questions persist as to whether this is the correct approach, or whether answers to the ANA questions related to separate attributes truly reflect a single latent variable. Put simply, should stated ANA questions be treated as formative or reflective measures in terms of how they are modelled?

Credit author statement

John M Rose, Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition, Antonio Borriello, Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Andrea Pellegrini, Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Declaration of competing interest

There are nil conflicts of interest involved with this manuscript.

Acknowledgements

This research was undertaken whilst all authors were members of the University of Technology Sydney UTS. This study has been supported by the Australian Research Council Discovery Program Grant DP180103718 titled: ‘Improving external validity of stated choice experiments’.

References

Bahamonde-Birke, F.J., Kunert, U., Link, H., Ortúzar, J., 2017. About attitudes and perceptions: fnding the proper way to consider latent variables in discrete choice models. Transportation 44, 475–493.
Bierlaire, M., 2020. A Short Introduction to PandasBiogeme, *Technical Report TRANSP-OR 200605, Transport And Mobility Laboratory*. ENAC, EPFL.
Blalock, H.M., 1964. Causal Inferences in Nonexperimental Research. University of North Carolina Press, Chapel Hill, NC.
Bollen, K.A., Diamantopoulos, A., 2017. In defence of causal-formative indicators: a minority report. Psychol. Methods 22 (3), 581–596.
Bollen, K.A., 1984. Multiple indicators: internal consistency or no necessary relationship? Qual. Quantity 18, 377–385.
Bollen, K.A., Bauldry, S., 2011. Three Cs in measurement models: causal indicators, composite indicators, and covariates. Psychol. Methods 16 (3), 265–284.
Bollen, K.A., Lennox, R., 1991. Conventional wisdom on measurement: a structural equation perspective. Psychol. Bull. 110, 305–314.
Bollen, K.A., Ting, K.F., 2000. A tetrad test for causal indicators. Psychol. Methods 5 (1), 3–22.
Borriello, A., Rose, J.M., 2021. Global and localised attitudinal responses in discrete choice. Transportation 48 (1), 131–165.
Borsboom, D., Mellenbergh, G.J., Van Heerden, J., 2003. The theoretical status of latent variables. Psychol. Rev. 110 (2), 203–219.
Chorus, C., Kroesen, M., 2014. On the (im-) possibility of deriving transport policy implications from hybrid choice models. Transport Pol. 36, 217–222.
Coltman, T., Devinney, T.M., Midgley, D.F., Venaik, S., 2008. Formative versus reflective measurement models: two applications of formative measurement. J. Bus. Res. 61 (12), 1250–1262.
Daziano, R.A., 2012. Taking account of the role of safety on vehicle choice using a new generation of discrete choice models. Saf. Sci. 50, 103–112.
Daziano, R.A., Bolduc, D., 2012. Covariance, identification, and finite-sample performance of the MSL and Bayes estimators of a logit model with latent attributes. Transportation 40, 647–670.
Diamantopoulos, A., 1999. Export performance measurement: reflective versus formative indicators. Int. Market. Rev. 16, 444–457.
Diamantopoulos, A., Winklhofer, H., 2001. Index construction with formative indicators: an alternative to scale development. J. Market. Res. 37, 269–277.
Dunlap, R.E., Van Liere, K.D., 1978. The “new” environmental paradigm. J. Environ. Educ. 9 (4), 10–19.

Dunlap, R.E., Van Liere, K.D., Mertig, A.G., Jones, R.E., 2000. New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. J. Soc. Issues 56 (3), 425–442.
Fantechi, T., Contini, C., Scozzafava, G., Casini, L., 2022. Consumer preferences for wild game meat: evidence from a hybrid choice model on wild boar meat in Italy. Agric. Food Econ. 10 (1), 1–18.
Fornell, C., Bookstein, F.L., 1982. Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. J. Market. Res. 19 (4), 440–452.
Green, P., 1984. Hybrid models for conjoint analysis: an expository review. J. Market. Res. (21), 155–169.
Green, P., Goldberg, S., Montemayor, M., 1981. A hybrid utility estimation model for conjoint analysis. J. Market. 45 (1), 33–41.
Guevara, C.A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. Transport. Res. Part A 82, 240–254.
Guzman, L.A., Arellana, J., Camargo, J.P., 2021. A hybrid discrete choice model to understand the effect of public policy on fare evasion discouragement in Bogata's Bus Rapid Transit. Transport. Res. Part A 151, 140–153.
Hardin, A., 2017. A call for theory to support the use of causal-formative indicators: a commentary on Bollen and Diamantopoulos (2017). Psychol. Methods 22 (3), 597–604.
Heise, D.R., 1972. Employing nominal variables, induced variables, and block variables in path analyses. Socio. Methods Res. 1 (2), 147–173.
Hensher, D.A., Rose, J.M., Greene, W.H., 2015. Applied Choice Analysis: A Primer, second ed. Cambridge University Press, Cambridge.
Hess, S., Hensher, D.A., 2013. Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. Transportation 40, 397–412.
Jin, F., Yao, E., An, K., 2020. Understanding customers' battery electric vehicle sharing adoption based on hybrid choice model. J. Clean. Prod. 258, 120764.
Joreskog, K.G., Goldberger, A.S., 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. J. Am. Stat. Assoc. 70 (351), 631–639.
Kløjgaard, M.E., Hess, S., 2014. Understanding the formation and influence of attitudes in patients' treatment choices for lower back pain: testing the benefits of a hybrid choice model approach. Soc. Sci. Med. 114, 138–150.
MacCallum, R.C., Browne, M.W., 1993. The use of causal indicators in covariance structure models: some practical issues. Psychol. Bull. 114, 533–541.
MacKenzie, S.B., Podsakoff, P.M., Jarvis, C.B., 2005. The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. J. Appl. Psychol. 90 (4), 710.
McFadden, D., Heiss, F., Jun, B., Winter, J., 2006. On Testing for Independence in Weighted Contingency Tables, *Medium for Econometric Applications*, 14.2. Rotterdam Econometrics Institute, pp. 11–19.
McFadden, D.L., 1980. Econometric models for probabilistic choice among products. J. Bus. 53 (3–2), S13–S29.
McFadden, D.L., 1986. The choice theory approach to market research. Market. Sci. 5 (4), 275–297.
Orne, M.T., 1959. The Demand Characteristics of an Experimental Design and Their Implications. American Psychological Association, Cincinnati, September. Paper presented at the.
Orne, M.T., 1969. Demand characteristics and the concept of quasi-controls. In: Rosenthal, R., Rosnow, R.L., Kazdin, A.E. (Eds.), Artifact in Behavioral Research. Academic Press, New York, pp. 143–179.
Ram, S., Sheth, J.N., 1989. Consumer resistance to innovations: the marketing problem and its solutions. J. Consum. Market. 6 (2), 5–14.
Rose, J.M., Bliemer, M.C.J., 2009. Constructing efficient stated choice experimental designs. Transport Rev. 29 (5), 587–617.
Rossiter, J.R., 2002. The C-OAR-SE procedure for scale development in marketing. Int. J. Res. Market. 19, 305–335.
Santos, A.C., Roberts, J.A., Barreto, M.L., Cairncross, S., 2011. Demand for sanitation in Salvador, Brazil: a hybrid choice approach. Soc. Sci. Med. 72 (8), 1325–1332.
Train, K.E., McFadden, D.L., Goett, A.A., 1987. Consumer attitudes and voluntary rate schedules for public utilities. Rev. Econ. Stat. 69 (3), 383–391.
Venaik, S., Midgley, D.F., Devinney, T.M., 2005. Dual paths to performance: the impact of global pressures on MNC subsidiary conduct and performance. J. Int. Bus. Stud. 36 (6), 655–675.