Transportation Research Part B 56 (2013) 234–253

Contents lists available at ScienceDirect

Transportation Research Part B

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

Specification issues in a generalised random parameters attribute nonattendance model

Andrew T. Collins *, John M. Rose [1], David A. Hensher [2]

[a] Institute of Transport and Logistics Studies, The University of Sydney Business School, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 30 November 2012
Received in revised form 16 July 2013
Accepted 1 August 2013

Keywords:
Attribute nonattendance
Random parameters attribute nonattendance model
Latent class model
Random parameters logit

ABSTRACT

An extensive literature has recognised that when travel choices are made, only a subset of the attributes of the choice alternatives may be considered or attended to by each decision maker. Numerous econometric approaches have been employed to identify attribute nonattendance (ANA), with the most prevalent in the literature being an adaptation of the latent class model. However, the two latent class structures so far employed either incur a potentially very high parametric cost, or rely on an assumption that nonattendance is independent across all attributes. We present a generalised model that allows for an arbitrary degree of correlation of nonattendance across attributes. In the presented stated choice study investigating short haul flights, this generalised model outperforms the existing approaches. Like two recent papers, the model handles both ANA and preference heterogeneity by combining continuously distributed random parameters with latent classes. However, we present recommendations regarding a number of identification issues stemming from the combination of these two forms of random parameters not covered in those papers. Further, covariates can be introduced into our generalised model to allow insights to be gained into ANA behaviour. We investigate stated ANA as a covariate, and find inferred ANA rates to be more aligned with stated ANA responses than alternative methods.

1. Introduction

A stream of literature has developed in recent years that recognises that when travel choices are made, each attribute of each choice alternative may only be considered by a subset of those making the choice. This has been described as the ignoring of attributes (Hensher et al., 2005), and, as used in this paper, attribute nonattendance (ANA; Scarpa et al., 2009). Much of the research has taken place in the transportation literature, although significant contributions have taken place in environmental economics (e.g. Scarpa et al., 2009) and health economics (e.g. Hole, 2011).

One interpretation of ANA is that it is a valid phenomenon reflecting the preferences of the individual making a choice. Cirillo and Axhausen (2006) suggest that some automobile drivers might legitimately have a zero, rather than negative, valuation of time in the vehicle, where an inflated mass at zero exists alongside some distribution of negative valuations. Gilbride et al. (2006) recognise that consumers choosing a product might have no intrinsic value for some of the attributes of the products on offer. This is plausible for travel choices such as the selection of long haul flights, where the choice alternatives may have many features (in-flight entertainment, seat pitch, stopover duration, etc.), and each individual may only

* Corresponding author. Tel.: +61 2 9114 1832.
E-mail addresses: andrew.collins@sydney.edu.au (A.T. Collins), john.rose@sydney.edu.au (J.M. Rose), david.hensher@sydney.edu.au (D.A. Hensher).
[1] +61 2 9114 1882.
[2] +61 2 9114 1824.

0191-2615/$ - see front matter
http://dx.doi.org/10.1016/j.trb.2013.08.001

value some of these features (Collins et al., 2012). In contexts such as these, ANA is not inherently a problem, but a valid behavioural phenomenon that may be of interest to the analyst. However, failure to capture ANA may lead to biases in model outputs such as willingness to pay measures (Hensher et al., 2005), and it may contribute to implausibly signed random parameter coefficients (Hensher, 2007). Another interpretation is that, for various reasons, what we believe to be ANA is only a partial reflection of the preferences of an individual. Hensher et al. (2012a) note that ANA might be a result of presenting respondents with behaviourally questionable trade-offs. Alemu et al. (2013) asked respondents who stated that they ignored an attribute why they did so, and found that respondents did so not only because they did not find the attribute important, but to make it easier to choose between the alternatives, because the attribute levels were unrealistically high or low, and because they believed that an attribute should not be traded against the others. One explanation for ANA is that it is a form of bounded rationality. In the face of choice complexity, an individual may selectively attend to the available information (DeShazo and Fermo, 2004; Cameron and DeShazo, 2011).

One approach to handling ANA is to ask participants in a stated choice experiment which attributes they ignored and then modify the utility expression of these individuals accordingly (Hensher et al., 2005), although various studies have questioned the reliability of such statements (e.g. Hess and Rose, 2007; Carlsson et al., 2010; Hess and Hensher, 2010). Lexicographic choice (Sæaelensminde, 2006) can be considered as an extreme form of ANA, whereby all attributes bar one are ignored. Such choice behaviour can be uncovered through inspection of multiple choices, although the shorter the panel, the less certain the analyst can be that the respondent truly is choosing based on one attribute only. An alternative is to identify ANA econometrically, such that no extra information such as stated ANA responses is required. The ANA literature has paid little attention to the censored normal random parameter distribution (Train and Sonnier, 2005) as a means of capturing ANA. However, since the estimated moments of the distribution capture both ANA and preference heterogeneity, the two phenomena could be confounded. Hess and Hensher (2010) proposed a technique that is informed by the conditional parameters estimates, however it is reliant on the selection of an arbitrary threshold value, and Mariel et al. (2011) have shown that the most accurate such value is dependent on the true ANA rate, which is latent. Several papers have focused on the relationship between ANA and scale (Campbell et al., 2008; Kragt, 2012), although scale will not be considered in this paper.

Another econometric approach now widely employed in the literature, referred to herein as the attribute nonattendance (ANA) model,[3] utilises the latent class model, where some classes have coefficients constrained to zero to represent ANA (Hess and Rose, 2007; Scarpa et al., 2009). A potential pitfall of the ANA model is that as the number of attributes for which nonattendance is modelled increases, the number of possible classes and class assignment parameters increases exponentially. One solution is to model attendance to only some of the available combinations. The decision as to which combinations to retain has been found to have a strong impact not just on model fit, but also the ANA rates inferred (Scarpa et al., 2009). Another solution proposed by Hole (2011) utilises an alternative class assignment model, in which the number of class assignment parameters rises linearly with respect to the number of attributes for which nonattendance is modelled. This approach relies on the assumption that the nonattendance rates are independent across attributes, and as such is referred to herein as the independent attribute nonattendance (IANA) model. Use of the conventional latent class model allows nonattendance rates to be correlated across attributes, and so is referred to as the correlated attribute nonattendance (CANA) model.

A concerning feature of the ANA model is the very high ANA rates reported across a number of studies. Cost has been found to be particularly susceptible, and particularly concerning, due to the key role that the attribute plays in the formulation of willingness to pay (WTP) values. For example, in a number of studies, ANA to cost ranged from over 60% to a high of 90.9% (Campbell et al., 2011; Hole, 2011; Campbell et al., 2010; Scarpa et al., 2009). Campbell et al. (2012) and Hess et al., 2013) have shown that such high ANA rates may be a consequence of not adequately handling taste heterogeneity. Campbell et al. (2012) estimated additional point masses for cost only, and noted a decrease in the ANA rate and a corresponding improvement in model fit. Whilst an important finding, this approach would prove even more parametrically challenging than the CANA model when ANA is modelled for more than one attribute. Hess et al. (2013) employed the Hole (2011) latent class structure, but in representing those individuals that attend to an attribute, estimated continuously distributed random parameters in place of a mean sensitivity. They observed a notable decrease in ANA rates and increase in model fit, and concluded that ANA and preference heterogeneity may be confounded under the ANA model. Hensher et al. (2012b) also estimated an ANA model with random parameters, which additionally handled the aggregation of common-metric attributes, and found an improvement in model fit over conventional latent class (LC) and random parameters logit (RPL) models. However, this model was only a small improvement over an ANA model that included multiple classes representing full attribute attendance. This alternative way of combining preference heterogeneity and ANA might be appealing in some contexts, as estimation times are very fast. Herein, the ANA model with random parameters is referred to as the random parameters attribute nonattendance (RPANA) model. Again there are two variants, depending on the assumption of independence of nonattendance across attributes: the random parameters independent attribute nonattendance (RPIANA) model and the random parameters correlated attribute nonattendance (RPCANA) model.

This paper makes a number of key contributions in the area of ANA generally, and the various forms of the ANA model in particular. Full correlation of nonattendance across attributes is parametrically expensive, and yet an assumption of full independence across all attributes may be too strong. This paper proposes a generalised latent class structure that can rely on an independence assumption across subsets of attributes, whilst allowing ANA to be correlated within those subsets. The

[3] Appendix A summarises this paper's nomenclature.

Table 1
Summary of ANA models in the literature.

Paper	Model	$K^{\star}$ [a]	Random parameters	Correlation structure	ANA covariates
Hess and Rose (2007)	ANA	1	No	N/A	Sociodem.
Scarpa et al. (2009)	CANA	5	No	Fully correlated	–
Hole (2011)	IANA	5	No	Independent	–
Hess et al. (2013)	RPIANA	2/5/6	Yes	Independent	–
Hensher et al. (2012b)	RPCANA	3	Yes	Fully correlated	–
Hole et al. (2012)	IANA	5	No	Independent	Stated ANA
This paper	RPANA generalised	4	Yes	Partially independent, flexible	Stated ANA

[a] The number of attributes for which ANA is modelled.

performance of this model is investigated in the context of a stated choice study investigating short haul flights, and differences in the ANA inferences are noted. The RPANA model is susceptible to various identification problems that have not been explored. This paper details some necessary conditions for the identification of the RPANA model. Only two papers utilising the ANA model have allowed the probability of estimated nonattendance to vary across respondents or observations in a systematic manner (Hess and Rose, 2007; Hole et al., 2012), and no papers utilising the RPANA model. Also, while it has been shown that stated ANA rates may be inaccurate, such information may still prove useful in an econometric specification (Hole et al., 2012). The generalised RPANA model proposed allows covariates to be specified in the latent class component of the RPANA model. In this application, the stated ANA responses are used, both to improve model performance, and to gain insight into the alignment of the stated ANA responses and the inferred ANA rates. Stated and inferred ANA is found to be more aligned under the RPANA model than with the conditional parameter estimate approach of Hess and Hensher (2010). Therefore, the contribution comes through the specification of a highly flexible ANA model that can handle random preference heterogeneity and both random and systematic ANA heterogeneity, an exploration of identification issues around this model, and a test of the model in an empirical application. Table 1 positions this paper against key contributions in the ANA literature, on a number of key dimensions discussed above.

2. Methodology

2.1. Attribute nonattendance (ANA) model

The model presented in this section generalises the latent class approach to modelling ANA (Hess and Rose, 2007; Hole, 2011). The two existing approaches, the IANA and CANA models, are first broadly outlined, and then a generalised ANA model is introduced in detail. Since the two existing approaches are special cases of the generalised model, the latter will be used to precisely define the former.

Consider a choice task wherein the choice alternatives are described by K attributes. The analyst wishes to model nonattendance to $K^{\star}$ of these attributes, which may represent all attributes ($K^{\star} = K$), or a lesser number ($1 \leqslant K^{\star} < K$). Choice of a lesser number may be behaviourally motivated, if some attributes are always attended to, or econometrically motivated, to lessen the number of parameters that must be estimated.

Under the latent class approach, the unconditional probability of respondent n choosing an alternative (or sequence of alternatives across multiple choice tasks) can be decomposed into the probability of that respondent exhibiting a certain pattern of attendance and nonattendance across attributes, and the probability of choosing the alternative or sequence of alternatives, conditional on belonging to a specific class of ANA behaviour. These two components are described in more detail:

Final ANA assignment probabilities. These are the probabilities of the respondent imposing specific combinations of attendance and nonattendance over $K^{\star}$ attributes, where there are up to $2^{K^{\star}}$ possible combinations. Each combination is represented by a class in the latent class model. Define M as the number of classes, where $M = 2^{K^{\star}}$ if all ANA combinations are to be modelled, or $1 < M < 2^{K^{\star}}$ if some specific ANA combinations are to be omitted. The probability of each respondent n belonging to class m is denoted P_{nm}, and will be referred to as the ANA assignment probability, in recognition of the behavioural interpretation of each class[4]. In some cases herein, P_{nm} will be referred to as the final ANA assignment probability, since this probability may be a function of two or more further probabilities, each of which also controls ANA assignment in some way. Alternate methods for generating P_{nm} will be detailed below.

Choice probabilities conditional on final ANA assignment. These are the probabilities of choosing an alternative, or sequence of alternatives across multiple choice tasks, conditional on assignment to a specific combination of ANA. Most examples in the literature employ an MNL model to calculate these probabilities. The choice alternatives are described by K attributes, $K^{\star}$ of which we model attendance or nonattendance to. For each ANA assignment class m, a unique combination of the taste coefficients associated with the $K^{\star}$ attributes will be constrained to zero, to reflect the specific com-

[4] This is distinct from the conventional latent class model, which has no such behavioural interpretation, with each class merely representing some combination of preference weights for the attributes of the choice alternatives.

bination of ANA that the class represents. When not constrained to zero, these coefficients are either constrained to be equal across classes (Scarpa et al., 2009), or unique coefficients are estimated for each class (Hensher and Greene, 2010). The former approach is the most common in the literature. While it requires less parameters to be estimated, it does not capture preference heterogeneity amongst those who attend to the attribute. The latter approach can capture preference heterogeneity which is systematically associated with the ANA pattern imposed.

The unconditional probabilities can be obtained by multiplying the final ANA assignment probabilities by the choice probabilities that are conditioned on the ANA assignment, and integrating over the M ANA assignment classes.

The most common approach in the literature for generating the final ANA assignment probabilities is to use the conventional latent class approach, with a single MNL model employed to calculate each of the M ANA assignment probabilities (Hess and Rose, 2007; Scarpa et al., 2009; Hensher and Greene, 2010; Campbell et al., 2011). If all combinations of ANA across $K^{\star}$ attributes are to be modelled, then the number of parameters required for ANA assignment increases exponentially as $K^{\star}$ increases. For even a trivial value of $K^{\star}$, the number of parameters might be prohibitive. However, specific ANA combinations may be omitted at the discretion of the analyst (Scarpa et al., 2009). This decision may be based either on an assumption that the combination of ANA is unreasonable or unlikely, or ex-post evidence that the combination does not occur.

An alternative, more parsimonious approach for generating the final ANA assignment probabilities has been proposed by Hole (2011). The conventional approach estimates a single MNL model that generates the probability of each combination of ANA across the $K^{\star}$ attributes. This approach estimates a binary logit model for each of the $K^{\star}$ attributes, each of which generates the probability of whether a single attribute is attended to or not. These will be referred to as ANA assignment probabilities, as distinct from the final ANA assignment probabilities, which are the probabilities of *combinations* of ANA across the $K^{\star}$ attributes. The final ANA assignment probability for each ANA combination, P_{nm}, is then the product of $K^{\star}$ ANA assignment probabilities, each obtained from the binary logit models. The selection of probability (attendance or nonattendance) to include in each element of this product is informed by whether m represents attendance or nonattendance to the attribute in question. There are $2^{K^{\star}}$ classes in the final ANA assignment model, but as few as $K^{\star}$ parameters controlling the assignment.[5] However, such parsimony relies on the assumption that the probability of not attending to any one of the $K^{\star}$ attributes is independent of the nonattendance probabilities of each of the other attributes. If the assumption holds, then the ANA assignment can be estimated more parsimoniously, and the conventional latent class approach will be an overparameterisation. If, however, some combination of attributes has a disproportionately high or low probability, then the independence assumption does not hold, and the approach might result in biased parameter estimates and a poorer model fit than the conventional latent class approach.

Whether the attribute nonattendance probabilities are independent is likely to vary from one empirical context to the next. Currently, the analyst could test both specifications, and see which best fits the data, using a measure such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC).[6] However, these two approaches represent two extremes of what can actually be considered a continuum. The approach proposed by Hole assumes that nonattendance is independent across all combinations of attributes. The conventional latent class approach makes no such assumption, and can handle any correlation structure over the $K^{\star}$ attributes. The conventional latent class approach can replicate the final ANA assignment probabilities obtained under the approach proposed by Hole, however it does so at the cost of more parameters, where these parameters may be superfluous, assuming independence holds. Crucially to the development of the generalised approach, it may be that the independence assumption is violated within some subsets of the $K^{\star}$ attributes, but not between these subsets. The most appropriate model then would be some intermediate point between the two extremes. Such a generalised model is now introduced.

Rather than have $K^{\star}$ ANA assignment models, each with two classes (Hole, 2011), or a single ANA assignment model, with up to $2^{K^{\star}}$ classes (Hess and Rose, 2007), we may have A ANA assignment models, with $1 \leqslant A \leqslant K^{\star}$. Each ANA assignment model a controls the nonattendance associated with $K_a^{\star}$ attributes. If all combinations of nonattendance to the $K_a^{\star}$ attributes are to be modelled, ANA assignment model a will have $2^{K_a^{\star}}$ classes. Specific combinations of attendance can be excluded by the analyst, resulting in fewer classes. Define C_a as the realised number of classes for each a. The final ANA assignment model will have $M = \prod_a^A C_a$ classes.

Define P_{nac} as the probability of respondent n belonging to class c in ANA assignment model a. The probability is calculated with an MNL model, such that

$$P_{nac} = \frac{e^{(\gamma_{ac} + \theta_{nac} z_n)}}{\sum_d^{C_a} e^{(\gamma_{ad} + \theta_{nad} z_n)}}. \tag{1}$$

A parameter, γ_{ac}, serves as a constant term, capturing the assignment to class c that cannot be explained by other factors. A vector of parameters, θ_{nac}, captures socio-demographic and other influences on assignment to class c in ANA assignment model a, for respondent n. To ensure identification, γ_{ac} is constrained to zero for one class. Given that most of the discussion in the literature is around attribute *non*attendance, constraining γ_{ac} to zero for the class that represents full attendance to the

[5] More than $K^{\star}$ parameters may control the ANA assignment, if covariates are introduced into the binary logit models. The fully notated generalised model presented below will allow covariates to influence ANA assignment.

[6] A likelihood ratio test is not possible, since the two models are not nested.

attributes is the most convenient such constraint to impose. It is also likely that in many empirical contexts, full attendance across the $K_a^\star$ attributes will have the highest probability of all possible ANA combinations, although this is not necessarily the case (e.g. Hensher et al., 2012a).

Recall that each of the C_a classes represents a unique pattern of ANA over the $K_a^\star$ attributes which have their ANA state determined by ANA assignment model a. In the final ANA assignment model, each class m will represent a unique pattern of ANA over all $K^\star$ attributes for which ANA is modelled. This pattern of ANA will be represented by a unique set of ANA assignment model classes, $\{c_1, \ldots, c_A\}$. The probability of respondent n belonging to class m is

$$P_{nm} = P_{n\{c_1,\ldots,c_A\}} = \prod_a^A P_{nac_a}. \tag{2}$$

Substituting in Eq. (1), this becomes

$$P_{n\{c_1,\ldots,c_A\}} = \prod_a^A \frac{e^{(\gamma_{ac_a}+\theta_{nac_a}z_n)}}{\sum_d^{C_a} e^{(\gamma_{ad}+\theta_{nad}z_n)}}. \tag{3}$$

Consider now the choice probabilities conditional on assignment to a class in the final ANA assignment model. While these probabilities can be derived using any form of choice model, the vast majority of latent class ANA models have utilised the multinomial logit model with fixed taste coefficients, which assumes that the unobserved component of utility is independently and identically extreme value type 1 distributed over alternatives and respondents. The formulation here also employs the MNL model, before being substituted by the RPL model in the next section.

The MNL model, without any constraints imposed, will first be defined. Then, the MNL model conditional on assignment to a class in the final ANA assignment model will be introduced, including the specific constraints that will be imposed to reflect ANA. Consider first the total utility of alternative i for respondent n, U_{ni}, which is composed of the representative utility V_{ni}, and the unobserved component of utility, ϵ_{ni}. The representative component is associated with a vector of observed variables, x_{ni}. The utility associated with these variables is estimated with a vector of taste coefficients β, such that the representative utility is $V_{ni} = \beta x_{ni}$. For the MNL model, the probability that alternative i will be chosen is

$$P_{ni} = \frac{e^{\beta x_{ni}}}{\sum_j^J e^{\beta x_{nj}}}. \tag{4}$$

The variables that enter into the representative utility contain the K attributes that describe the choice alternatives. Each attribute k may have more than one variable enter into the representative utility, for example if the attribute is dummy or effects coded. The taste coefficients in the β vector represent the sensitivities to the associated variables. For any choice model that is conditioned on a combination of ANA over $K^\star$ attributes, some elements of β may be constrained to zero to represent ANA to one or more attributes. Notably, if an attribute is coded such that more than one variable enters into the representative utility, as with dummy or effects coding, then nonattendance to that attribute is handled by constraining to zero *all* taste coefficients associated with all variables that represent the attribute (see Scarpa et al., 2009).

Next, we need to partition the full set of taste coefficients β into one or more subsets. First, β_0 is composed of the taste coefficients for the $K - K^\star$ attributes for which ANA is not modelled. This will be an empty set if ANA is modelled for all attributes. Then introduce A subsets, each denoted β_a, which are composed of the taste coefficients associated with the $K_a^\star$ attributes for which ANA is controlled by ANA assignment model a. Each a controls assignment to C_a classes, each representing a unique combination of ANA over $K_a^\star$ attributes. Each combination will represent a unique pattern of censoring of β_a. For each a, introduce C_a sets, each denoted β_{ac}. The elements of β_{ac} are either zero, representing ANA, or the taste coefficients drawn from the same position in β_a, representing attendance to the attribute. That is, the taste coefficients that are not censored are constrained to be equal across the C_a sets. Alternatively, unique coefficients could be estimated when censoring does not take place (as with Hensher and Greene, 2010), however an equality constraint will be imposed in this paper. The variables to enter into the representative utility, x_{nj}, are similarly partitioned into $A + 1$ subsets. Variables associated with attributes for which ANA is not modelled are in set x_{nj0}, while the variables associated with attributes for which ANA is modelled are partitioned into A subsets x_{nja}.

Conditional on assignment to classes $\{c_1, \ldots, c_A\}$ in each of the A ANA assignment models, the representative utility of alternative j for respondent n now becomes

$$V_{nj|c_1,\ldots,c_A} = \beta_0 x_{nj0} + \sum_a^A \beta_{ac_a} x_{nja}. \tag{5}$$

This censors the taste coefficients associated with the attributes that are ignored in the class of the final ANA assignment model upon which the representative utility is conditioned.

For the MNL model, the probability that respondent n will choose alternative i, conditional on assignment to classes $\{c_1, \ldots, c_A\}$, is

$$P_{ni|c_1,\ldots,c_A} = \frac{e^{\beta_0 x_{ni0} + \sum_a^A \beta_{ac_a} x_{nia}}}{\sum_j^J e^{\beta_0 x_{nj0} + \sum_a^A \beta_{ac_a} x_{nja}}}. \tag{6}$$

For panel data, we can specify the probability with respect to a sequence of choices of alternatives over T time periods, $\{i_1, \ldots, i_T\}$. Assuming that the unobserved component of utility is now independently and identically extreme value type 1 distributed over alternatives, respondents, *and* time, the probability of a sequence of choices of alternatives, conditional on assignment to classes $\{c_1, \ldots, c_A\}$, is

$$P_{n|c_1,\ldots,c_A} = \prod_t^T \left[\frac{e^{\beta_0 x_{ni_t t0} + \sum_a^A \beta_{ac_a} x_{ni_t ta}}}{\sum_j^J e^{\beta_0 x_{njt0} + \sum_a^A \beta_{ac_a} x_{njta}}} \right]. \tag{7}$$

Use of panel data as formulated limits the ability of the model to handle task specific ANA.

The unconditional probability of a sequence of choices for respondent n is obtained by taking the product of two probabilities: the probability of a combination of ANA, and the probability of the sequence of choices, conditional on assignment to that combination of ANA; then integrating over all analyst specified combinations of ANA. This can be expressed as

$$P_n = \sum_{c_1}^{C_1} \cdots \sum_{c_A}^{C_A} P_{n\{c_1,\ldots,c_A\}} P_{n|c_1,\ldots,c_A}. \tag{8}$$

Substituting in Eqs. (3 and 7), Eq. (8) becomes

$$P_n = \sum_{c_1}^{C_1} \cdots \sum_{c_A}^{C_A} \prod_a^A \left[\frac{e^{(\gamma_{ac_a} + \theta_{nac_a} z_n)}}{\sum_d^{C_a} e^{(\gamma_{ad} + \theta_{nad} z_n)}} \right] \prod_t^T \left[\frac{e^{\beta_0 x_{ni_t t0} + \sum_a^A \beta_{ac_a} x_{ni_t ta}}}{\sum_j^J e^{\beta_0 x_{njt0} + \sum_a^A \beta_{ac_a} x_{njta}}} \right]. \tag{9}$$

Certain specifications of A allow the model to represent the two latent class approaches in the literature. If there is only one ANA assignment model, i.e. $A = 1$, then this is a conventional latent class model, with specific constraints on the taste coefficients across classes, reflecting ANA. Since this can capture correlation in ANA across all attributes, this extreme will be referred to as the correlated attribute nonattendance (CANA) model. If there is one ANA assignment model for every attribute for which ANA is modelled, i.e. $A = K^\star$, then this is the endogenous attribute attendance model from Hole (2011). This extreme will be referred to as the independent attribute nonattendance (IANA) model. If $1 < A < K^\star$, then this is an ANA model that assumes that independence of ANA holds only between some subsets of the $K^\star$ attributes. This ANA model has not been presented in the literature, and represents one of the contributions of this paper. In the interest of brevity, the ANA acronyms may be appended by $K^\star$, which represents the number of attributes to which nonattendance is modelled. If $K^\star = 1$, then the single attribute for which ANA is modelled may follow the acronym when referencing the model (e.g. ANA1 fare model).

2.2. Random parameters attribute nonattendance (RPANA) model

To capture preference heterogeneity amongst decision makers that attend to the attributes, we now introduce random parameters, such that the taste coefficients β vary over decision makers with density $f(\beta)$. A distribution is specified for each taste coefficient, and the moments of these distributions are estimated with structural parameters. Most commonly used distributions are described by two moments, however this paper employs several distributions for which a single moment is estimated; notably, the constrained triangular and uniform distributions, wherein the spread is constrained to equal the mean, and the Rayleigh distribution. All of these distributions are constrained in sign, the motivation for which will be discussed in Section 4.3.1.

Eq. (7) now becomes

$$P_{n|c_1,\ldots,c_A} = \int \prod_t^T \left[\frac{e^{\beta_0 x_{ni_t t0} + \sum_a^A \beta_{ac_a} x_{ni_t ta}}}{\sum_j^J e^{\beta_0 x_{njt0} + \sum_a^A \beta_{ac_a} x_{njta}}} \right] f(\beta) d\beta. \tag{10}$$

Substituting Eq. (10) into Eq. (8), we obtain an unconditional probability of a sequence of choices for respondent n of

$$P_n = \sum_{c_1}^{C_1} \cdots \sum_{c_A}^{C_A} \prod_a^A \left[\frac{e^{(\gamma_{ac_a} + \theta_{nac_a} z_n)}}{\sum_d^{C_a} e^{(\gamma_{ad} + \theta_{nad} z_n)}} \right] \int \prod_t^T \left[\frac{e^{\beta_0 x_{ni_t t0} + \sum_a^A \beta_{ac_a} x_{ni_t ta}}}{\sum_j^J e^{\beta_0 x_{njt0} + \sum_a^A \beta_{ac_a} x_{njta}}} \right] f(\beta) d\beta. \tag{11}$$

This choice probability underpins the RPANA model. Practical issues with estimating the RPANA model are discussed in Section 4.3.1. All models were estimated using proprietary code.

3. Empirical setting

The empirical setting for this paper is a stated choice experiment conducted in early 2004, that was based on a short haul flight between Sydney and Melbourne, Australia. Respondents were asked to imagine that they were making the flight for holiday travel. Each choice task contained three labelled flight alternatives. A choice was made between one flight each from three airlines: Qantas, the dominant Australian carrier; Virgin Blue, then a relatively young airline with four years of operations; and Air New Zealand (Air NZ), a foreign carrier that does not operate the Sydney–Melbourne route. A fourth, no-choice option was presented, which signalled that the respondent would not want to make any of the three flights. Two choices were obtained: one that included the no-choice alternative, and a forced choice over the three airlines. The analysis contained herein makes use only of the forced choice.

Each alternative was described by four attributes: fare, flight time, departure time, and flight time variability. Fare assumed one of four levels in Australian dollars: $79, $99, $119 and $139. Flight time was either 40, 50, 60 or 70 min. Departure time was either 6am, 10am, 2pm or 6pm. Flight time variability was used to convey the range of likely flight times. However, the attribute was not well received, with 69% of respondents stating in a subsequent question that they ignored the attribute. Tests with random parameters imply an even distribution of respondents for and against flight time variability, suggesting that many did not understand the attribute. Hess and Hensher (2013) also note problems with a travel time variability attribute. More recent experiments have looked at alternative expressions of variability (e.g. Hensher et al., 2011), and the approach used for this data set, which was state of the art at the time, has been found to be ambiguous to a large number of respondents. It will be omitted from subsequent analysis as the ambiguity just adds unnecessary heterogeneity.

Each airline alternative was described by the same set of attribute levels that were varied via an orthogonal experimental design. That is, no airline had a disproportionate number of each of the attribute levels, despite, for example, a tendency for Virgin Blue to offer cheaper tickets than Qantas in the market. The orthogonal design contained 40 choice tasks in total, all of which were completed by 213 respondents, in one of three ways. As a part of a broader research agenda investigating multiple survey sessions per respondent spread over time, respondents either completed all 40 choice tasks in one sitting; 20 choice tasks each in two sessions, with one week of separation; or 10 choice tasks per session over four sessions each separated by a week. Regardless of the configuration employed, this paper utilises the first 20 choice tasks completed by each respondent. Given the length of the panel, it is possible that ANA is varying to some extent across choice tasks. However, the RPANA model as formulated cannot simultaneously handle preferences that are invariant across choice tasks, and ANA that varies across these tasks. The sample consisted of students, with an average age of 21. Fifty-nine percent of the sample was female, and 41% had made a holiday trip to Melbourne prior to the study. No other socio-demographic or experience information was gathered.

An examination of the 213 respondents revealed two who chose the Qantas alternative for all 20 choice tasks, and one who always chose the Virgin Blue alternative. Since one flight from each airline was presented in each choice task, if this is a true representation of lexicographic choice behaviour, then no trading is occurring between the attributes describing the airlines. It may be that trading is taking place across the airline label and the attributes, with the attributes in the other two alternatives just failing to compensate in each of 20 successive choices. Nonetheless, the length of the panel suggests this is unlikely,[7] and so these three observations are dropped. Interestingly, this is an extreme case of attribute nonattendance, where all attributes are ignored, and only the airline labels are attended to. The final sample size is 4200 observations across 210 respondents. Six point nine percent of these respondents stated that they ignored fare, 18.1% flight time, and 15.95% departure time. Respondents may also have ignored the airline label, however they were not asked if this was the case. Whilst the literature has called into question the reliability of the responses to these questions (Hess and Rose, 2007), and cautions against using them deterministically (Hensher et al., 2007), they nonetheless provide a broad sense of what the nonattendance rates might be in aggregate across respondents. Further, by suggesting that there is likely to be at least some incidence of ANA in this dataset, they motivate the analyst to find a way to adequately accommodate ANA econometrically.

4. Results

The specifications and notable features of the 13 reported models are summarised in Table 2. Models 1 and 4 are MNL and RPL models, respectively, and serve as a baseline. Models 2 and 3 are two specifications of the ANA models widely reported in the literature, that do not utilise random parameters. Comparisons are made between the two extreme specifications of correlation in ANA, and with stated ANA rates. Models 5 through 8 introduce random parameters, with ANA modelled for only one attribute per model. This allows insights into the distributional form and ANA rates to be investigated in isolation. Models 9 through 11 include random parameters and nonattendance to all four attributes. The two extremes of ANA correlation are specified, as is a hybrid specification that selectively relaxes the independence assumption. Finally, Models 12 and 13 nest Models 6 and 7 respectively, and introduce stated ANA responses as covariates. The motivation for estimating each model is more comprehensively elucidated in each relevant section.

[7] Panel lengths in stated choice experiments are typically shorter. Bliemer and Rose (2011) examined top tier transportation journals from January 2000 to August 2009, and found an average panel length of 9.4 and a median length of nine. The shorter the panel, the less confidence can be placed on any interpretation of lexicographic behaviour.

Table 2
Overview of models and stated ANA rates.

Model number	Model type	ANA correlation structure	Probability ignored				AIC/n	BIC/n
			Fare	Flight time	Dep. time	Airline		
1	MNL	–	–	–	–	–	1.3255	1.3361
2	CANA	Correlated	13.63%	16.33%	53.42%	78.98%	1.2140	1.2469
3	IANA	Independent	13.26%	11.78%	52.17%	80.72%	1.2175	1.2341
4	RPL	–	–	–	–	–	1.1054	1.1266
5	RPANA1	–	2.12%	–	–	–	1.1056	1.1283
6	RPANA1	–	–	10.78%	–	–	1.1050	1.1261
7	RPANA1	–	–	–	29.22%	–	1.0958	1.1184
8	RPANA1	–	–	–	–	73.34%	1.1017	1.1243
9	RPANA4	Hybrid	4.05%	8.55%	27.28%	44.30%	1.0916	1.1173
10	RPCANA4	Correlated	4.97%	9.36%	26.24%	48.92%	1.0920	1.1237
11	RPIANA4	Independent	1.50%	8.38%	26.83%	45.36%	1.0939	1.1195
12	RPANA1	–	5.21%/51.81%	–	–	–	1.1019	1.1245
13	RPANA1	–	–	20.86%/73.06%	–	–	1.0914	1.1156
–	Stated ANA rates		6.90%	18.10%	15.95%	–	–	–

Table 3
Model 1: MNL.

	Param.	*t*-ratio	WTP	*t*-ratio
Fare	−0.0729	−47.16	–	–
Flight time	−0.0407	−19.24	$0.56[b]	18.94
Depart 6am	−0.7398	−9.60	$10.15[b]	9.22
Depart 10am	0.6638	7.85	$9.11[a]	8.11
Depart 2pm	0.0723	0.92	$0.99[a]	0.92
Virgin Blue	0.0065	0.13	$0.09[a]	0.13
Air NZ	−0.4201	−7.84	$5.77[b]	7.93
Model fits				
LL (0)	−4614.17			
LL (MNL)	−2776.56			
Parameters	7			
ρ^2	0.3983			
Adjusted ρ^2	0.3972			
AIC/n	1.3255			
BIC/n	1.3361			
Observations	4200			
Respondents	210			

[a] WTP to obtain the attribute level, or a one unit increase in the attribute.
[b] WTP to avoid the attribute level, or a one unit increase in the attribute.

4.1. MNL model

The first model estimated is an MNL model, which is reported in Table 3. Fare and time are both highly significant and of expected sign, with respondents preferring cheaper fares and shorter flights. Willingness to pay measures are split into two categories: WTP to obtain a desirable attribute level, or a one unit increase in a desirable attribute (such measures will be suffixed by [a] in Table 3 and henceforth); and WTP to avoid an undesirable attribute level, or a one unit increase in an undesirable attribute ([b]). On average, respondents are willing to pay 56 cents to avoid one minute of flight time, or equivalently, $33.50 to avoid one hour of flight time. The departure time levels are dummy coded, with 6pm forming the base level. Significant parameters and WTPs are obtained for 6am and 10am, with the WTP values suggesting that respondents are, on average, willing to pay $10.15 to depart at 6pm instead of 6am, and $9.11 to depart at 10am instead of 6pm. The parameter and WTP for 2pm departure is not significant, suggesting an indifference between 2pm and 6pm departure, ceteris paribus.

Alternative specific constants (ASCs) were estimated for travel with Virgin Blue and Air NZ, with estimates being relative to travel with Qantas. An insignificant parameter for Virgin Blue suggests that, on average, respondents are indifferent to whether they fly with Qantas or Virgin Blue, ceteris paribus. There is a mean sensitivity against Air NZ however, with a willingness to pay to avoid the airline of $5.77. This measure captures preferences that are not accounted for via attributes in the choice experiment. It is worth noting that the same levels of fare, flight time and departure time were applied to all three airline alternatives. That is, no airline was presented as operating flights that tended to be cheaper or shorter than that of the competitor, or operating disproportionately at certain times of the day, as may be the case in the market. Thus differences in the ASCs are unlikely to be the consequence of different attribute ranges across the choice alternatives. One possible

Table 4
ANA models accommodating ANA for all attributes.

	Model 2: CANA			Model 3: IANA		
	Param.	*t*-ratio		Param.	*t*-ratio	
Fare	−0.1084	−49.78		−0.1059	−53.78	
Flight time	−0.0606	−19.03		−0.0568	−19.05	
Depart 6am	−2.4617	−16.16		−2.3533	−17.72	
Depart 10am	0.9870	6.35		0.9347	7.75	
Depart noon	0.6691	4.23		0.6213	4.98	
Virgin Blue	−0.6165	−4.85		−0.6855	−6.54	
Air NZ	−2.0201	−11.37		−2.0892	−13.98	
	WTP	WTP *t*-ratio	Diff.[a] *t*-ratio	WTP	WTP *t*-ratio	Diff.[a] *t*-ratio
Flight time	$0.56[d]	20.68	0.00	$0.54[d]	21.74	0.09
Depart 6am	$22.72[d]	16.10	7.92	$22.23[d]	18.20	7.92
Depart 10am	$9.11[c]	6.64	0.00	$8.83[c]	7.92	0.19
Depart noon	$6.17[c]	4.34	3.28	$5.87[c]	4.99	3.25
Virgin Blue	$5.69[d]	4.79	4.25	$6.47[d]	6.50	5.09
Air NZ	$18.64[d]	11.38	8.37	$19.73[d]	13.83	9.52
	Ignored	*t*-ratio		Ignored	*t*-ratio[b]	
Fare	13.63%	–		13.26%	26.20	
Flight time	16.33%	–		11.78%	11.81	
Departure time	53.42%	–		52.17%	40.15	
Airline	78.98%	–		80.72%	17.51	
Model fits						
LL	−2526.75			−2545.82		
Parameters	22			11		
ρ^2	0.4524			0.4483		
Adjusted ρ^2	0.4495			0.4468		
AIC/*n*	1.2140			1.2175		
BIC/*n*	1.2469			1.2341		

[a] *t*-ratio of difference between this model's WTP and MNL WTP.
[b] *t*-ratio for difference to 0% ANA.
[c] WTP to obtain the attribute level, or a one unit increase in the attribute.
[d] WTP to avoid the attribute level, or a one unit increase in the attribute.

influence on the ASCs is a left-to-right bias, whereby respondents are more likely to choose the first alternative of the three, which were presented side by side. The order of the alternatives was not varied, where such variation would help mitigate such a bias. While the possibility of some degree of left-to-right bias cannot be dismissed, it is believed to be minimal in this setting. Consequently, when modelling ANA, a censoring of the ASCs to zero will be interpreted as nonattendance to the airline, even though there may be some degree of confounding with the other unobserved effects. In the interest of brevity, any reference in the text to nonattendance to attributes also includes nonattendance to the choice alternative labels.

Model fit statistics appear reasonable, and serve foremost as a baseline for subsequent models. All models estimated in this paper utilise the 4200 observations obtained from the 210 respondents that were retained after data cleaning, and so these numbers will not be presented in subsequent tables.

4.2. ANA model

This section presents the results from CANA and IANA models, with nonattendance modelled for fare, flight time, departure time, and airline. In the CANA model (Model 2), any pattern of correlation of ANA can be captured through the estimation of 16 classes in a single ANA assignment model, while the IANA model (Model 3) assumes independence of ANA, and contains four separate ANA assignment models, each with two classes. The plausibility of the ANA rates will be assessed. Model outputs and fit will be compared between the two, to gauge whether the assumption of independence holds. Finally, the models will serve as one benchmark for the RPANA models. As discussed in Section 2, nonattendance to dummy coded attributes (departure time) and ASCs (airline) is best modelled by setting to zero the coefficients for *all* dummy coded levels or ASCs. If not all coefficients are set to zero, then what is captured is not nonattendance, but rather an alternative expression of preference.

Table 4 presents the model results[8]. There is a strong alignment in nonattendance rates between the two ANA models, with the exception of flight time, with a rate of 16.33% for the CANA model and 11.78% for the IANA model. However, the estimated rates themselves do not always have face validity, which in part can be established by a comparison with the stated ANA rates.[9] The estimated ANA rates of fare, at 13.63 and 13.26, are somewhat higher than the stated ANA rate of 6.9%. To put these rates in

[8] The ANA assignment model parameters are not reported in the interest of brevity.
[9] While stated ANA is not reliable, it provides a ballpark figure.

a broader context, Hess et al. (2013) estimated an ANA rate of 25% for fare, using the ANA model. The ANA rate of flight time, estimated by the CANA model at 16.33%, is close to the stated rate of 18.1%. However, the rate of 11.78% estimated by the IANA model is somewhat lower. The estimated and stated rates are wildly divergent for departure time, with estimated rates of 53.42% and 52.17% being far higher than the stated rate of 15.95%. The stated rate of attendance to the airline alternative labels was not collected in the survey, however the estimated rates of 78.98% and 80.72% appear implausibly high. In sum, the ANA rates appear questionable.

Since a single parameter controls the ANA rate for each attribute in the IANA model, the associated standard error can be used to provide a measure of statistical reliability of the ANA rate. Table 4 presents, for each IANA nonattendance parameter, a *t*-ratio which represents whether the ANA rate is different from zero. The difference is significant for all attributes. In contrast, the CANA model determines ANA rates by summing the class assignment probabilities of all classes that treat the attribute as ignored, and no measure of statistical confidence can be calculated at the attribute level.

Comparing model fits, both models offer a significant improvement on the MNL model, with drastically lower AIC values. The CANA model has a better log likelihood value than the IANA model, and outperforms it on the AIC despite costing an additional 11 parameters. However, on the BIC, which more strongly penalises additional parameters, the IANA model outperforms the CANA model. No definitive conclusions can therefore be drawn about the appropriateness of the independence assumption that facilitates the more parsimonious IANA model. The generalised ANA model introduced in this paper allows the assumption to be made only for subsets of attributes, and such a specification may lead to a more appropriate ANA model. Given the shortcomings of the ANA model, testing of the many possible independence configurations made available by this generalised model will be reserved for the RPANA model, which additionally introduces random parameters. Indeed, the parametric cost of employing random coefficients will likely make a parsimonious model specification particularly desirable.

4.3. RPANA model

4.3.1. Model specification and identification

For the RPANA models, 5000 Halton draws are employed. Two types of estimation problems are encountered. The less problematic of these are cases whereby the model converges on a local maxima, which is plausible in the context of such a highly nonlinear model. This was found to be more common when attendance to multiple attributes was being modelled, and typically manifested itself as nonattendance rates tending to zero. In most cases, such problems were overcome by first estimating nonattendance to one attribute at a time, then using the recovered parameter values as start values for the RPANA model that models attendance to multiple attributes. Therefore, caution must be warranted before concluding that ANA rates for an attribute are indeed zero.

A more fundamental problem is concerned with the choice of random parameter distribution, and what is believed to be a fundamental incompatibility between the RPANA model and parameter distributions that can span both the positive and negative domain. In this empirical context, any attempt to include such distributions led to a multitude of estimation problems, including flat log likelihoods and singular covariance matrices. Problematic distributions include the normal, which is unbounded and by definition will always have support over both the positive and negative domain; the triangular, which is bounded but can freely span zero; and the uniform, also bounded but free to span zero.

Interestingly, the censored normal also exhibits the same problems. With its point mass at zero, the censored normal can already capture ANA. The motivation for the RPANA model over simply using the censored normal is that the latter is likely more prone to confounding ANA with preference heterogeneity, since ANA is captured through the same parameters that capture preference heterogeneity. The unbounded nature of the underlying normal distribution suggests that the ANA rate implied by the censored normal distribution is always greater than zero. If it is very close to zero, through some appropriate combination of μ and σ, then the RPANA model could capture the vast majority of ANA, and the censored normal distribution would primarily capture the continuous component of utility. What appears to be happening in practice, at least in the context of this dataset, is that the potential to capture ANA through both the ANA parameter, and the censoring of the normal distribution, leads to an identification problem whereby some arbitrary combination of the two sources of ANA can approximate the 'true' ANA. This in turn leads to the problems with estimation.

The same phenomenon may be occurring with the normal, triangular and uniform distributions. Now, however, a certain proportion of coefficients close to zero, including those of implausible sign, is approximating ANA. This in turn leads to an identification problem, with the ANA parameter and the continuous distribution's support near zero both 'competing' for the share of attribute nonattenders. By limiting the support of the continuous distribution near zero through the application of a distribution that is bounded on one side at zero, this identification problem can potentially be overcome. Distributions that do not encounter problems in this dataset include the constrained triangular, lognormal, and Rayleigh. The two other papers that have employed the RPANA model use the lognormal (Hess et al., 2013) and the constrained triangular distributions (Hensher et al., 2012b), where Hensher et al. (2012b) also report instability with unconstrained distributions, citing an earlier version of this paper as guidance on choosing a distribution that is constrained in sign.

However, the use of a zero bound distribution appears to be a necessary but not sufficient condition. Problems were encountered in this dataset with the constrained uniform distribution, in which the spread is constrained to be equal to the mean. This results in an equal share of coefficients over a domain spanning between zero and two times the mean. It may be that by not tapering towards zero, the continuous distribution has enough support near zero to suitably approximate ANA, leading to an identification problem. The consequence of this is that care must be taken when choosing distributions, and the specifics of any empirical application may have an impact on what can be identified. To some extent, this also calls into question the confidence the analyst can place on an inferred ANA rate. Indeed, it may not be possible to completely unentangle attribute nonattendance and low attribute sensitivity.

The problem with distributions spanning zero poses a challenge in this empirical setting for the dummy coded departure time parameters and alternative specific constants, all of which span zero when RPL models are estimated, and have behaviourally sound justifications for doing so. For example, most respondents have a preference against 6am departures, but some do not. Consequently, the identification problem presented above might apply here, although it appears to be overcome by the joint censoring of each of the coefficients associated with an attribute, when that attribute is ignored. The optimism that the model can be identified stems from the likelihood that the ANA condition is harder to approximate with a number of independently varying random parameters that are each associated with an attribute level or alternative label. Dummy coding the departure time and airline, and introducing ANA jointly across all related parameters, is found to provide large improvements in model fit. However, model estimation is not stable, with singular covariance matrices commonly occurring, suggesting that an identification problem may remain. One potential source of the problem is the normalisation of the dummy parameters and ASCs, where one coefficient is fixed to zero. It may be that the base level of all zero coefficients under dummy coding is confounded with the ANA condition, which is represented in exactly the same way. An alternative normalisation can be achieved with effects coding. An attribute with L levels is coded into $L-1$ variables. A utility coefficient, β_l, is estimated for each of these variables. The base level of utility is not zero, as with dummy coding, but $\sum_l^{L-1} -\beta_l$. Crucially, with effects coding employed, no estimation problems are encountered. Effects coding the ASCs is unusual, however Train (2009) notes that the ASCs need not be normalised to zero, and that doing so is merely easier. With the RPANA model, we have sufficient motivation to deviate from convention.

4.3.2. Single attribute nonattendance

The first set of RPANA models presented (Models 5 through 8) will only model nonattendance to a single attribute. This allows the impact of ANA to each attribute to be examined in isolation, with no assumption about the independence or correlation of ANA necessary. Table 5 first details Model 4, the RPL model that serves as the base specification for the RPANA1 models, then the four RPANA1 models.

The best fitting RPANA1 fare model, Model 5, which uses the lognormal distribution, is presented in Table 5. The ANA rate is low, at 2.12%, and statistically different from 0% (whereupon it would collapse to an RPL model). Confidence intervals provide another useful way to assess the precision of the estimated ANA rate. A 95% confidence interval ranges from 0.32% to 12.67%. The stated ANA rate of 6.9% is higher than the estimated rate, but comfortably within the confidence interval. The ANA rate estimated with the RPANA1 model is much lower than the ANA rates of 13.63% and 13.26% inferred from the CANA and IANA models reported earlier. This is in line with the drop in ANA rate for fare observed by Hess et al. (2013) when moving from the ANA to RPANA model, from 25% to 9%. The present drop is plausible, as the ANA models were found to estimate suspiciously high ANA rates, and the model has recently been criticised for confounding ANA with preference heterogeneity (Hess et al., 2013).

To assess whether modelling ANA with a RPANA1 model leads to an improved model fit, each of the reported RPANA1 models estimated are compared with their RPL counterparts that utilise the same distributions, using likelihood ratio tests. The null hypothesis that Model 5 is equivalent to a pure RPL specification cannot be rejected. With one degree of freedom, the test statistic of 1.14 does not exceed the chi-squared critical value of 3.84 at the 95% confidence level. Furthermore, testing of various RPL models with alternative fare distributions, not reported here in detail, resulted in a RPL model with uniformly distributed fare that outperformed Model 5 on the AIC (1.1044, 1.1056) and BIC (1.1255, 1.1283).

We consider two reasons why the RPANA model may not be a significant improvement over the RPL model, even when non-zero ANA rates are identified. First, a relatively small ANA rate may be adequately approximated by the parameter distribution in the RPL model. For example, a RPL model with uniformly distributed fare has a mean fare parameter of -0.1586 and a spread of 0.1493, meaning that there is some support near zero. This model has an AIC of 1.1044 and BIC of 1.1255, outperforming Model 5. As the ANA rate increases, a conventional continuous parameter distribution would be less able to approximate the ANA condition, and it is likely that model fit would be diminished. It may be that for low nonattendance rates to an attribute, it may be sufficient to not model nonattendance, and simply allow the continuous distribution to approximate the nonattendance condition, especially if the analyst is more interested in the model predictions than the ANA rates per se.

Second, the ANA may interact not just with the low sensitivities, but with the entire distribution, including the extreme sensitivities. Consider the case of the lognormal distribution. As the mean of the lognormal distribution decreases, the tail will become fatter. Consequently, the lognormal may perform well when there is a mass at or close to zero, and a mass of high sensitivities to the attribute, captured by a long, fat tail (see Hess et al., 2013). If, however, the ANA is largely captured through the estimated discrete mass point with the RPANA model, then this may limit the ability of the lognormal to capture the high sensitivities with the long tail. In effect, a long tail would increase the mass close to zero, which would compete with the mass point at zero. Indeed, the ANA rate with the lognormal distribution in Model 5 (2.12%) is lower than the 3.47% observed in another, unreported model, that employed the constrained triangular distribution for fare, which suggests that such confounding may be occurring. So, whereas Hess et al. (2013) suggest caution be taken when using the lognormal for cost, because ANA might lead to very high WTP values, we additionally suggest that care be taken if using the lognormal in the RPANA model if it is believed that there truly may be high sensitivities to cost, as the resulting mass near zero in the

Table 5
RPL and RPANA1 models.

	Model 4 RPL		Model 5 RPANA1 fare		Model 6 RPANA1 flight time		Model 7 RPANA1 departure time		Model 8 RPANA1 airline	
	Param.	*t*-ratio	Param.	*t*-ratio	Param.	*t*-ratio	Param.	*t*-ratio	Param.	*t*-ratio
Fare (lognormal)										
μ	−2.0250	−30.92	−1.9795	−33.53	−2.0257	−29.19	−2.0162	−25.72	−2.0007	−30.03
σ	0.8618	13.56	0.7906	10.09	0.8586	13.57	0.8982	12.89	0.8534	13.06
$\gamma_{Ignored}$	–	–	−3.8333	3.17[a]	–	–	–	–	–	–
$P_{Ignored}$	–		2.12%		–		–		–	
$P_{Ignored}$ 95% C.I.	–	–	0.32%	12.67%	–	–	–	–	–	–
Flight time (varies)										
μ	−0.0717	−15.73	−0.0722	−16.25	−0.0833	−16.30	−0.0758	−15.05	−0.0608	−13.08
σ	0.0407	8.19	0.0409	8.07	–	–	0.0407	7.57	−1.3928	−10.30
Distribution	Normal		Normal		Const. triangular		Normal		Normal	
$\gamma_{Ignored}$	–	–	–	–	−2.1139	8.82[a]	–	–	–	–
$P_{Ignored}$	–		–		10.78%		–		–	
$P_{Ignored}$ 95% C.I.	–	–	–	–	4.00%	25.94%	–	–	–	–
Depart 6am (normal)										
μ	−1.3466	−11.30	−1.3723	−10.52	−1.3320	−10.66	−1.9953	−8.18	1.4780	14.35
σ	1.4825	14.98	1.4735	14.64	1.4730	15.34	2.0484	12.25	0.9959	10.38
Depart 10am (normal)										
μ	0.9778	10.51	0.9734	10.29	0.9944	10.45	1.5333	11.57	0.8365	7.63
σ	0.8369	7.69	0.8428	7.73	0.8334	7.43	0.9425	6.48	0.1142	1.15
Depart 2pm (normal)										
μ	0.1410	1.59	0.1286	1.35	0.1257	1.32	0.4343	2.78	1.0019	9.42
σ	0.9761	9.64	0.9768	9.51	0.9893	10.07	1.3703	9.25	0.3462	3.61
$\gamma_{Ignored}$	–	–	–	–	–	–	−0.8846	12.85[a]	–	–
$P_{Ignored}$	–		–		–		29.22%		–	
$P_{Ignored}$ 95% C.I.	–	–	–	–	–	–	14.14%	50.85%	–	–
Virgin Blue (normal)										
μ	0.1829	3.85	0.1827	3.84	0.1829	3.81	0.1551	3.28	0.5252	5.06
σ	0.3458	6.48	0.3392	6.36	0.3460	6.22	0.3443	6.40	−0.8839	−6.30
Air NZ (normal)										
μ	−0.4524	−8.44	−0.4501	−8.39	−0.4568	−8.46	−0.4643	−8.41	0.5110	6.53
σ	0.4372	9.08	0.4329	8.98	0.4360	9.03	0.4590	8.52	−3.2343	2.99
$\gamma_{Ignored}$	–	–	–	–	–	–	–	–	0.1173	15.39[a]
$P_{Ignored}$	–		–		–		–		52.93%	
$P_{Ignored}$ 95% C.I.	–	–	–	–	–	–	–	–	31.49%	73.34%
Model fits										
LL	−2307.35		−2306.78		−2306.40		−2286.10		−2298.53	
Parameters	14		15		14		15		15	
ρ^2	0.4999		0.5001		0.5001		0.5045		0.5019	
Adjusted ρ^2	0.4983		0.4983		0.4985		0.5028		0.5001	
AIC/n	1.1054		1.1056		1.1050		1.0958		1.1017	
BIC/n	1.1266		1.1283		1.1261		1.1184		1.1243	
$\chi^2_{1,.05}$ w.r.t. RPL model = 3.84 [b]	–		1.14		9.02		42.50		17.63	

[a] *t*-ratio for difference to 0% ANA.
[b] χ^2 test is with respect to the RPL model with the same distributions, which is not necessarily Model 4.

continuous distribution may compete with the ANA mass point, and lead to an underreporting of the nonattendance condition.

Model 6, with a constrained triangular distribution for and nonattendance modelled to flight time, is presented in Table 5. The ANA rate of 10.78% is significantly different to zero, close to the IANA model rate of 11.78%, and a little less than the CANA model rate of 16.33%. However, again, the RPANA model strongly outperforms the ANA models. The RPANA model rate of 10.78% is somewhat lower than the stated ANA rate of 18.1%, but again the stated rate lies comfortably within the confidence interval, which spans from 4.00% to 25.94%. Model 6 represents a statistically significant improvement in model fit over the RPL model with the same distributions. Nonetheless, Model 6, with an AIC of 1.1050 and BIC of 1.1261, is weakly outperformed by the RPL model with Rayleigh (1.1048, 1.1244), censored normal (1.1049, 1.1260) and uniform distributions (1.1049, 1.1261). These differences are slight, but it is notable that for both fare and flight time, the RPANA1 model does not have *improved* model fit over the RPL model.

Model 7 models nonattendance to departure time only. Normally distributed departure time parameters give the best performance for both the RPL and RPANA models. Compared to the equivalent RPL model, the RPANA model represents a large, statistically significant improvement in model fit. The ANA rate is 29.22%, with a confidence interval of 14.14–50.85%. The stated ANA rate of 15.95% lies towards the low end of this range, while the rates recovered by the ANA models exceed the top end of the range, at 53.42% and 52.17%.

Model 8 models nonattendance to airline only. This model fit is a considerable improvement on the equivalent RPL model. The ANA rate is sizeable, at 52.93%, with a confidence interval of 31.49–73.34%. There was no stated ANA collected, with which the estimated rate can be compared. The implausible ANA rates under the ANA models of 78.98% and 80.72% exceed the upper end of the confidence interval.

From the evidence presented thus far, two key conclusions can be drawn. The ANA rates under the RPANA model are much lower than under the ANA model, and it appears as if the RPANA model is limited in this study to handling ANA to departure time and airline. However, two techniques will be introduced that capture nonattendance to fare and flight time *and* lead to the RPL model being outperformed. The first approach, detailed in the next section, is to allow for nonattendance to multiple attributes, and, crucially, to allow ANA to fare and flight time to be correlated, by including both in the same ANA assignment model. The second approach, detailed in Section 4.3.4, introduces stated ANA as covariates in the ANA assignment models.

4.3.3. Multiple attribute nonattendance

A possible reason for the lack of improvement in model fit when moving from the RPL to the RPANA1 fare and flight time models is that nonattendance to the two attributes may not be independent, and thus an assumption of the RPIANA model is violated. This assumption can be relaxed by combining fare and flight time into the one ANA assignment model, with the incidence rates of all *combinations* of fare and flight time nonattendance estimated. A series of models that handled nonattendance to two attributes (i.e. RPANA2 models) were first estimated to gain further insight into whether ANA is independent across the attributes. For each pair of attributes, a RPIANA2 and a RPCANA2 model were estimated. Comparisons were performed on model fits and consistency of ANA rates for each attendance pattern across the two attributes. The only decisive evidence for correlation in ANA was found between fare and flight time. On the AIC, this RPCANA2 model (1.1030) outperforms the equivalent RPIANA model (1.1050), and the best RPL model tested (uniform fare and normal flight time at 1.1044).

Next, three models are presented in Table 6 that model nonattendance to all four attributes. The first, Model 9, is a hybrid RPANA4 model that leverages the independence assumptions, but not for fare and flight time. Model 9 is motivated by the poor performance of the RPANA1 fare and flight time models, and the promising results of the RPANA2 model that allows for correlation in ANA between these two attributes only. The next two models are the two possible extremes of the model. Model 10 is a RPCANA4 model, that allows any degree of correlation in ANA across attributes to be captured. Motivation for Model 10 comes from the possibility that ANA is not independent across *any* attributes, and that failure to capture such correlation will likely be detrimental to model fit and the model outputs. Model 11 is a RPIANA4 model, that assumes full independence of ANA across all attributes. The parsimony of the RPIANA4 is desirable. Of key interest is whether the added complexity of selective relaxation of the independence assumption is worthwhile.

Whilst Model 11 provided only one possible model specification with respect to the ANA assignment model, Models 9 and 10 required that specific combinations of ANA be dropped. During estimation of Model 9, the incidence rate of the combination of ignored fare and attended flight time approached zero ($\gamma_{ac} \rightarrow -\infty$), and so this class was dropped from the ANA assignment model. For Model 10, the RPCANA model, all 16 ANA combinations were initially modelled. However, it was apparent that not all combinations could be supported, and classes were removed in a stepwise fashion. The most obvious problem in the first model estimated lay in the four classes representing fare nonattendance and flight time attendance, consistent with Model 9. The log likelihood became flat, and the standard errors for the ANA assignment parameters associated with these combinations became extremely large. Three more ANA combinations were dropped, because their incidence rate approached zero. The final specification modelled nine combinations of ANA, requiring eight parameters in the ANA assignment model.

Models 9 to 11 all outperform Model 4, the baseline RPL model, on the AIC and BIC. Model 9 nests the RPANA1 departure time and airline models (Models 7 and 8), as well as the aforementioned RPANA2 model with correlated ANA to fare and flight time. Log likelihood ratio tests reveal that all three previous models are outperformed by Model 9. Models 9, 10 and 11 are now compared to each other in terms of model fit, using the AIC and BIC, since they do not nest. Comparing Models 9 and 10, the former has a slight advantage on the AIC, and a stronger advantage when the four additional parameters of Model 10 are more greatly penalised by the BIC. Model 11 has the same number of parameters as Model 9, but a worse log likelihood, AIC and BIC. This suggests that an assumption of full independence of ANA across attributes in this instance may be detrimental to model fit. A comparison of Models 10 and 11 draws mixed results, and conclusions as to which model fits better depends on whether the AIC or BIC is used for the comparison.

Specific ANA rates for Models 9–11 are reported in Table 6. The rates are reported in several ways, to reflect the different ways that the RPIANA, RPCANA and hybrid RPANA models estimate ANA, and to allow standard errors to be reported where appropriate. The first set of rates reflect the ANA rates and associated standard errors for the 12 retained classes of Model 10, the RPCANA model. The second set of rates are estimated directly by Model 9, the hybrid RPANA model. The third set of rates rely on the ANA independence assumption, and are estimated directly by both Models 9 and 11. All rates reported can be inferred trivially through multiplication or addition, but only those rates estimated directly by a model have associated standard errors.

Table 6
RPANA4, RPCANA4 and RPIANA4 models.

				Model 9: RPANA4			Model 10: RPCANA4			Model 11: RPIANA4		
				Param.	*t*-ratio		Param.	*t*-ratio		Param.	*t*-ratio	
Fare (lognormal)		μ		−1.9112	−29.21		−1.8917	−26.58		−1.9584	−26.97	
		σ		0.7600	9.77		0.7293	9.86		0.8240	10.56	
Flight time (const. $\triangle$)		μ		−0.0834	−14.64		−0.0840	−15.05		−0.0826	−14.86	
		σ		–	–		–	–		–	–	
Depart 6am (normal)		μ		−1.9702	−9.73		−1.9548	−7.91		−1.9642	−7.86	
		σ		1.9985	12.39		1.9923	12.42		1.9748	12.31	
Depart 10am (normal)		μ		1.4931	9.48		1.4997	10.58		1.4642	10.84	
		σ		0.9475	6.16		0.9466	6.59		0.9303	6.55	
Depart 2pm (normal)		μ		0.3643	2.69		0.3439	2.25		0.3392	2.05	
		σ		1.3563	10.20		1.3587	8.75		1.3391	8.53	
Virgin Blue (normal)		μ		0.2730	3.64		0.2957	3.03		0.2832	3.15	
		σ		0.4666	6.41		0.4699	4.88		0.4735	4.95	
Air NZ (normal)		μ		−0.7926	−5.92		−0.8432	−6.28		−0.8001	−5.32	
		σ		0.5074	5.59		0.5218	6.19		0.5301	6.39	
Fare	Flight time	Dep. time	Airline	Param.	s.e.	Rate	Param.	s.e.	Rate	Param.	s.e.	Rate
Ignore	Ignore	Ignore	Ignore	–	–	0.49%	−3.0596	0.8919	1.60%	–	–	0.02%
Attend	Ignore	Ignore	Ignore	–	–	0.54%	–	–	–	–	–	1.00%
Attend	Attend	Ignore	Ignore	–	–	11.05%	−1.0069	0.4836	12.49%	–	–	10.98%
Ignore	Ignore	Attend	Ignore	–	–	1.30%	–	–	–	–	–	0.04%
Attend	Ignore	Attend	Ignore	–	–	1.45%	–	–	–	–	–	2.74%
Attend	Attend	Attend	Ignore	–	–	29.46%	0.0190	0.4677	34.83%	–	–	29.95%
Ignore	Ignore	Ignore	Attend	–	–	0.62%	−3.2006	0.9016	1.39%	–	–	0.02%
Attend	Ignore	Ignore	Attend	–	–	0.68%	−3.0610	0.9609	1.60%	–	–	1.21%
Attend	Attend	Ignore	Attend	–	–	13.90%	−1.3173	0.6390	9.15%	–	–	13.23%
Ignore	Ignore	Attend	Attend	–	–	1.64%	−2.8500	0.7140	1.98%	–	–	0.05%
Attend	Ignore	Attend	Attend	–	–	1.82%	−2.5082	1.6017	2.78%	–	–	3.30%
Attend	Attend	Attend	Attend	–	–	37.04%	–	–	34.18%	–	–	36.08%
Ignore	Ignore	–	–	−3.1174	0.5041	4.05%	–	–	4.97%	–	–	0.13%
Attend	Ignore	–	–	−3.0116	0.7031	4.50%	–	–	4.38%	–	–	8.25%
Attend	Attend	–	–	–	–	91.45%	–	–	90.64%	–	–	90.25%
Ignore	–	–	–	–	–	4.05%	–	–	4.97%	−4.1855	1.0731	1.50%
Attend	–	–	–	–	–	95.95%	–	–	95.03%	–	–	98.50%
–	Ignore	–	–	–	–	8.55%	–	–	9.36%	−2.3924	0.5744	8.38%
–	Attend	–	–	–	–	91.45%	–	–	90.64%	–	–	91.62%
–	–	Ignore	–	−0.9805	0.2614	27.28%	–	–	26.24%	−1.0032	0.2407	26.83%
–	–	Attend	–	–	–	72.72%	–	–	73.76%	–	–	73.17%
–	–	–	Ignore	−0.2290	0.5705	44.30%	–	–	48.92%	−0.1863	0.5809	45.36%
–	–	–	Attend	–	–	55.70%	–	–	51.08%	–	–	54.64%
Model fits												
LL				−2275.35			−2272.11			−2280.09		
Parameters				17			21			17		
ρ^2				0.5069			0.5076			0.5059		
Adjusted ρ^2				0.5049			0.5051			0.5038		
AIC/n				1.0916			1.0920			1.0939		
BIC/n				1.1173			1.1237			1.1195		

First consider differences in the ANA rates across the various ANA and RPANA models, and the stated ANA responses. For every attribute, the various RPANA models (5 through 11) always have lower ANA rates than in the two ANA models (2 and 3). This finding is consistent with the findings of Hess et al. (2013). For fare and departure time, this represents a move towards the stated ANA rates, but for flight time, the move is away.

Next, compare the ANA rates across Models 9–11, the three RPANA4 models with various assumptions of independence of ANA. For fare, it is notable that the rate for Model 11, at 1.50%, is lower than the other two models, and much closer to the rate of 2.12% in Model 5. Model 5 implicitly treats nonattendance to fare as independent across attributes, by not estimating nonattendance to other attributes, while Model 11 explicitly relies on an independence assumption. When ANA rates between fare and flight time are allowed to be correlated, in Models 9 and 10, no classes simultaneously representing nonattendance to fare and attendance to flight time can be supported, suggesting that the incidence rate for such a combination is zero. However, the independence assumption forces this specific combination to be non-zero, and it appears that the consequence is a downward bias in the ANA rate for fare. By contrast, the ANA rates are quite consistent across Models 9–11 for flight time and departure time, and relatively consistent for airline.

From these findings, we can see that assuming independence of ANA across all attributes leads to inferior model fit compared to a model that selectively allows the assumption to be relaxed. Further, the assumption reduces the ANA rate to fare dramatically. The selective relaxation is also parametrically less expensive than a full relaxation that allows for full correlation, and this is borne out in better model fits under both the AIC and BIC. Further, Model 9 allows ANA covariates such as stated ANA to be entered more directly against the attribute itself.

4.3.4. Covariates in ANA

All of the RPANA models estimated thus far have treated the probability of attribute nonattendance as being the same across respondents. However, introducing covariates into the ANA assignment models allows for a systematic examination of how different subsets of respondent types approach ANA. With such covariates, the ANA probabilities can vary across respondents.

We introduce stated ANA as a covariate in the ANA assignment models of the RPANA model. This approach was taken by Hole et al. (2012) in the context of the ANA model. Whilst they handled some taste heterogeneity through the specification of an additional class with full attendance, questions remain as to whether the nonattendance classes are in fact associated with nonattendance, or instead capture some taste heterogeneity (Hess et al., 2013). The motivation for including stated ANA in the RPANA model is threefold. First, introducing covariates may lead to an improvement in model fit. Whilst the stated responses are unlikely to be completely accurate, they might still contain useful information (Hess and Hensher, 2013), as those who state that they ignore an attribute are more likely to ignore it than those who state otherwise. There is an opportunity for those RPANA model specifications presented earlier that did not outperform the RPL model to now do so, by leveraging the stated ANA responses. Second, there are behavioural insights into the nature of stated ANA responses. For example, the percentage of the sample for which stated and inferred ANA do not align can be determined. This is more informative than simply comparing aggregate stated and inferred ANA rates. Finally, the use of stated ANA as a covariate allows for a more nuanced comparison of the RPANA model with alternative methods of inferring ANA, such as the conditional parameter estimate approach of Hess and Hensher (2010). It may be used to determine if, on the current dataset, the RPANA model has greater face validity. Each of these motivations will now be investigated in turn.

First though, it should be noted that Hess and Hensher (2013) criticise the use of stated ANA responses on two grounds. Stated ANA may not accurately align with actual attendance behaviour, and there may be an endogeneity problem as the stated ANA responses might be correlated with the error component of the model. Whilst the covariate models we report do not overcome this second problem, they do allow the stated ANA responses to be handled probabilistically, and so not be reliant on the very strong assumption that stated ANA is completely accurate and free from error.

Three RPANA1 models were estimated, each modelling ANA for a single attribute, with the stated ignoring response for that attribute, for respondent n, included as a dummy in the vector of covariates z_n. The three models correspond to the three attributes for which stated ANA was captured: fare, flight time, and departure time. The fare model failed to converge, perhaps due to the estimation of an additional parameter for the covariate, when the baseline model without the covariate only yielded an estimated ANA rate of *2.12%*. The flight time model could be estimated, and strongly outperformed Model 6, the baseline model without the covariate. This is Model 12 reported in Table 7. With only a single additional parameter, associated with the covariate, Model 12 represents a significant improvement in model fit over Model 6 ($15.06; \chi^2_{1,.05} = 3.84$). With an AIC of *1.1019* and BIC of *1.1245*, Model 12 also outperforms Model 4, the baseline RPL model (1.1054, 1.1266). Thus, if a RPANA1 model does not outperform the RPL model, as was the case in this dataset for fare and flight time, ANA covariates in a RPANA model offer another way to do, along with an appropriate specification of the ANA correlation structure in a RPANA model that handles nonattendance to multiple attributes. Model 13 in Table 7 is a RPANA1 model which leverages stated nonattendance to departure time, and strongly outperforms the baseline model, Model 7 ($20.34; \chi^2_{1,.05} = 3.84$). No stated ANA responses were collected for airline. In summary, it can be seen that stated ANA can be effectively utilised to improve model fit.

Next we consider the behavioural implications, by examining the implied ANA rates for both stated attenders and stated ignorers. The rates for flight time are telling, with only a small proportion of stated attenders, *5.21%*, ignoring flight time, or conversely, *94.79%* of stated attenders doing what they stated. This suggests that stated attendance is fairly accurate. In contrast, only *51.81%* of stated ignorers actually did so. For departure time, stated ignorers have an ANA rate of *73.06%*, while stated attenders have an ANA rate of *20.86%* (i.e. *79.14%* did what they said). Both of these ANA rates are higher than for flight time, implying that stated attendance responses are less accurate for departure time, but stated nonattendance responses are more accurate.

The final motivation for using stated ANA a covariate is to facilitate a comparison of the alignment between stated and inferred ANA rates, across alternative models which can infer ANA. This may determine, at least on the current dataset, whether the RPANA model has greater face validity than the alternative models. Where the stated and inferred approaches do not align, it is not possible to determine what is the truth. Both the stated responses and the method of inference may exhibit error. However, when comparing alternative models on the same dataset, more alignment does provide greater face validity, in a much more convincing way than a mere comparison of aggregate ANA rates.

A comparison on this dataset is drawn between three different models: the RPANA model, and the conditional parameter estimate methodology of Hess and Hensher (2010), using two different sets of thresholds. Additionally, the findings from Hess and Hensher (2010) are reported, to see if any patterns emerge across datasets. The RPANA model with stated ANA as a covariate allows, for each possible stated ANA response, the generation of the probabilities of attendance and nonattendance. By contrast, the conditional parameter estimate approach categorises each respondent as either attending or ignoring, and these can be compared to the stated ANA responses. Table 8 reports the alignment of the ignored and attended classi-

Table 7
RPANA1 models with stated ANA as a covariate for modelled ANA.

	Model 12 Flight time		Model 13 Departure time	
	Param.	*t*-ratio	Param.	*t*-ratio
Fare (lognormal)				
μ	−2.0208	−34.80	−2.0167	−25.85
σ	0.8651	13.55	0.8952	13.17
Flight time (varies)				
μ	−0.0858	−15.70	−0.0757	−14.99
σ	–	–	0.0403	7.56
Distribution	Const. triangular		Normal	
$\gamma_{Ignored}$	−2.9011	8.97	–	–
$P_{Ignored\vert StatedAttended}$	5.21%		–	
$\theta_{StatedIgnored}$	2.9735	10.44	–	–
$P_{Ignored\vert StatedIgnored}$	51.81%		–	
Depart 6am (normal)				
μ	−1.3533	−10.64	−1.9890	−8.31
σ	1.4771	15.37	2.0658	12.13
Depart 10am (normal)				
μ	1.0030	10.35	1.5288	11.73
σ	0.8438	7.55	0.9623	6.83
Depart 2 pm (normal)				
μ	0.1172	1.18	0.4489	3.14
σ	1.0121	10.26	1.3828	9.98
$\gamma_{Ignored}$	–	–	−1.3335	19.20
$P_{Ignored\vert StatedAttended}$	–		20.86%	
$\theta_{StatedIgnored}$	–	–	2.3311	13.40
$P_{Ignored\vert StatedIgnored}$	–		73.06%	
Virgin Blue (normal)				
μ	0.1822	3.81	0.1564	3.33
σ	0.3455	6.26	0.3434	6.42
Air NZ (normal)				
μ	−0.4567	−8.42	−0.4653	−8.50
σ	0.4321	8.97	0.4558	8.25
Model fits				
LL	−2298.87		−2275.93	
Parameters	15		16	
ρ^2	0.5018		0.5068	
Adjusted ρ^2	0.5000		0.5049	
AIC/n	1.1019		1.0914	
BIC/n	1.1245		1.1156	

fications at the sample level (columns 1 and 2), and the overall alignment as the sum of these (column 3). Further, for each of stated and inferred ANA, we report the percentage of classifications to the ignoring condition for which the other approach aligns (columns 4 and 5). These two columns are informative if two extreme assumptions are made, that either stated or inferred ANA is fully correct. Under either assumption, the extent to which the alternative method can recover the 'correct' classification is reported.

In implementing the Hess and Hensher (2010) approach, we calculated the coefficients of variation (CV) of the conditional parameter estimates of an RPL model with all normal distributions. The reader is referred to that paper for a full exposition of the methodology. Full model results are not reported here, for brevity. Applying the CV threshold to the departure time attribute is challenging, as the effects coding results in three parameters and CVs. Attribute nonattendance is most likely associated with high CVs for all attributes, just as ANA is represented in the RPANA model by constraining the coefficients of all associated parameters to zero. Thus, we applied the threshold to the minimum of the three CVs. Initially, a threshold of two was employed, where this value was suggested by Hess and Hensher (2010), and acknowledged as being somewhat arbitrary, but conservative. We can confirm the conservative nature of the threshold, as it led to very low inferred ANA rates: for flight time and departure time, inferred rates of 1.43% and 2.38% respectively, against stated rates of 18.1% and 15.95% respectively. Thus we additionally optimised the thresholds, a concept broadly proposed but not operationalised by Hess and Hensher (2010). For each attribute, a threshold was chosen that forced the aggregate inferred ANA rate to equal the stated ANA rate. Of course, this does not imply that all stated ignorers will be inferred as ignoring the attributes they stated they did.

Table 8
Alignment of stated and inferred attendance and nonattendance, and comparison of percentages of correct classification.

Paper, methodology	Attribute	Alignment (% of sample)			Of inferred ignored stated as such (%)	Of stated ignored inferred as such (%)
		Ignored	Attended	Overall		
(Data column)		(1)	(2)	(3)	(4)	(5)
This paper, RPANA model	Flight time	9.38	77.63	87.01	68.73	51.81
	Departure time	11.65	66.52	78.17	39.93	73.06
This paper, conditionals, threshold = 2	Flight time	0.48	80.95	81.43	33.33	2.63
	Departure time	0.71	82.38	83.10	30.00	4.48
This paper, conditionals, optimised threshold	Flight time	6.43	70.24	76.67	35.53	35.53
	Departure time	6.43	74.76	81.19	40.91	40.30
Hess and Hensher (2010), conditionals, threshold = 2	Free flow travel time	2.93	74.63	77.56	18.77	23.09
	Running costs	3.41	69.27	72.68	63.62	11.85
	Travel time variability	9.27	50.24	59.51	31.67	31.15
	Slowed down travel time	0.00	81.95	81.95	0.00	0.00
	Toll costs	0.49	89.76	90.25	25.13	5.58

The first comparisons drawn are between the alternative threshold specifications. Once the threshold is optimised, the alignment of stated and inferred ignoring (column 1) increases dramatically, from under 1% to 6.43% for both attributes. The percentage of stated ignoring inferred as such (column 5) increases dramatically, because very little is inferred as ignoring when a threshold of two is applied. In contrast, the percentage of inferred ignoring stated as such (column 4) increases only modestly. However, the alignment of stated and inferred attendance (column 2) drops moderately, such that total alignment (column 3) drops for each attribute. Thus, while the optimisation by its very nature increases the inferred ANA rate, it also leads to a fair degree of misclassification.

The next comparison is between the RPANA and conditional parameter estimate methodologies. For flight time, the overall alignment (column 3) is higher with the RPANA model than with both threshold specifications, but a little less for departure time. On the measures in columns 4 and 5, the alignment is noticeably higher for flight time, and generally higher for departure time, with one measure being on par. Comparing the RPANA results to those from Hess and Hensher (2010), the alignment is on average much higher, although caution is warranted, as these are different datasets, and accuracy of the stated ANA responses might be explaining some of this difference. Overall, the evidence suggests that the RPANA model is producing results that are more consistent with stated ANA, thus lending it greater face validity.

The only socio-demographic variables collected were gender and age. Introducing these as ANA covariates for each attribute in turn failed to lead to any improvement in model fit. Overall, it is found that introducing stated ANA as a covariate in the ANA assignment models has the potential to lead to key insights into the accuracy of stated ANA, and significant improvements in model fit. One disadvantage is that stated ANA responses need to be collected, where the RPANA model was motivated in part by a desire to be relieved of such a burden. Also, the endogeneity issue remains, however the hybrid model approach of Hess and Hensher (2013) shows promise as a way to overcome any potential bias. As in their model, a latent variable could explain stated ANA responses. This same latent variable could still moderate the marginal utilities, but crucially, could also moderate the propensity to attend to an attribute in the RPANA model, by being entered as a covariate in the ANA assignment model. Thus, the RPANA model with covariates proposed herein would serve as the underlying model, with extra layers added on top. Due to the scope of the undertaking, this will remain an area for future research.

5. Discussion and conclusion

Attribute nonattendance has now been an active area of research for nearly a decade. The search for an appropriate econometric method of identifying and accommodating ANA has been a central theme of this research. The latent class based ANA model has seen widespread use across a range of fields including transportation, environment economics and health economics. The most recent variant has been the addition of random parameters to the latent class structure (Hess et al., 2013; Hensher et al., 2012b), in what we refer to as the RPANA model. We extend on these two papers in several important ways. The RPANA model is susceptible to identification problems. It is important that an analyst employing the RPANA model is aware of these, and so we have explored them in detail. Forcing random parameter distributions associated with linearly coded attributes to be constrained in sign appears to be a necessary but not sufficient condition for identification. Distributions that span zero can be introduced when multiple parameters are specified for an attribute, however it is necessary to effects code the attribute. Of the distributions that are constrained in sign, the lognormal is the most widespread and probably has the most appeal. However, other distributions should also be tested. Beyond the obvious motivation to find the best distribution that fits the data, differences in the ANA rate recovered as the distribution varied in this paper's empirical application suggest that some degree of confounding between ANA and preference heterogeneity may remain. Which distribution leads to the least confounding is difficult to determine, but model fit is still likely a good guide.

As with the two previous applications, the RPANA model leads to an improved model fit over the RPL and ANA models. However, a lack of improvement when nonattendance was modelled to fare only suggests that correlation in nonattendance across attributes might be important in some contexts. Such correlation can be accommodated when the conventional latent class model is applied, as in the majority of ANA models estimated in the literature, although the parametric cost can be considerable. An alternative latent class structure that assumes full independence of ANA across attributes (Hole, 2011; Hess et al., 2013) is more parsimonious, but the assumption may be too strong in some contexts. We have proposed and tested a generalised latent class component to the ANA and RPANA models, that can apply the independence assumption between selected subsets of attributes. In the dataset tested herein, this model leads to an improvement in model fit, and different ANA inferences. This suggests that the correlation structure of the nonattendance may be important as the RPANA model is implemented. At the very least, alternative models should be tested before a model with full independence is accepted. The generalised model proposed provides the analyst with more flexibility when performing such a specification search.

The generalised latent class component of the RPANA model may make the inclusion of other heuristics more appealing. Common-metric aggregation (Hensher et al., 2012b) could be introduced selectively into appropriate ANA assignment models[10]. For example, two assignment models could be introduced, where one handles various time components, and the other various cost components. Within each assignment model, common-metric aggregation could coexist with various nonattendance patterns, and yet assignment to a time and cost class could remain independent. Alternatively, various time and cost nonattendance and aggregation patterns could be handled independently of nonattendance to all remaining attributes.

The proposed generalised RPANA model also allows covariates to be entered into the ANA assignment component of the model, allowing the ANA rate to vary across decision makers. Use of stated ANA responses as a covariate led to an improvement in model fit, including cases where it was not previously observed when ANA was modelled for fare in isolation. It also allowed for a more nuanced comparison of stated and inferred ANA. As expected, stated ANA did not completely align with inferred ANA, however the alignment was more promising than under the conditional parameter approach proposed by Hess and Hensher (2010). An endogeneity issue remains when stated ANA is used as a covariate, but the addition of latent variables to the RPANA model is a promising way forward and an interesting area for future research.

As noted in the introduction, Alemu et al. (2013) collected a number of stated reasons for not attending to an attribute. Against some of these reasons, they were still able to estimate significant coefficients for some attributes, suggesting that some respondents may still attend to these attributes even if they stated otherwise, at least on some of the choice occasions. The RPANA model could employ these responses, by entering each stated reason as a covariate in place of a single indicator variable. The specification proposed herein integrates the choice probabilities at the level of the respondent, such that significant covariates would likely suggest nonattendance by respondents across all choice occasions. Consequently, in addition to true indifference, the model is likely to capture protests against attributes being traded. Other reasons from Alemu et al. (2013), including unrealistic attribute levels, and ease of choosing between alternatives, may lead to different ANA behaviour by the same respondent across different choice tasks. There is a chance therefore that such ANA behaviour would be captured as taste heterogeneity. If the ANA behaviour is a consequence of an inappropriately designed choice task, then a stated response would help identify the problem, and to some extent, an appropriately specified RPANA model might isolate any biasing influence.

The RPANA model has appeal in the context of travel behaviour, as attribute nonattendance is plausible in a range of circumstances. This paper identifies considerable ANA in short haul flights to departure time and airline, moderate ANA to flight time, and low levels of ANA to fare. The extra product features of long haul flights, such as in-flight entertainment and differing seat pitches, are also likely to appeal to only part of the population. It is possible, though, that ANA is not actually capturing true preferences or processing rules. The attribute may have no influence on choice, and in effect not be attended to, because the attribute levels may not be sufficiently differentiated, or may be in an inappropriate range (Hensher et al., 2012a). If these levels and tradeoffs are reflective of choice scenarios that would be encountered in real life, either now or in some plausible future scenario, then ANA that is induced in this way is reasonable, and capturing it may indeed be preferable. For example, some drivers may not attend to a toll, cordon charge, or road user charge for any values that would plausibly be introduced, or variable time-of-day charging may be of insufficient range over time to have salience. However, if the ANA is the consequence of poor experimental design in a stated choice study, then while the RPANA model may do a good job of identifying and isolating ANA, the ANA might have limited behavioural meaning. Care must be taken when generating the experimental design. Careful piloting might help prevent such problems. Also, this paper only employed a single dataset. Further research with more empirical applications would reveal how frequently the flexible correlation structure proposed in this paper leads to an improvement over the extremes of full independence and full correlation, as well as the usefulness of a wider range of nonattendance covariates.

[10] These might then be considered information processing assignment models.

Appendix A. Glossary

ANA	Attribute nonattendance
ANA model	Attribute nonattendance model
ANA#/RPANA# model	An ANA/RPANA model where nonattendance is modelled for # attributes
ANA1/RPANA1 *name* model	An ANA/RPANA model where nonattendance is modelled for a single attribute, where that attribute is called *name*
ASC	Alternative specific constant
CANA/RPCANA model	An ANA/RPANA model in which nonattendance is fully correlated across attributes, i.e., $A = 1$
CV	Coefficient of variation
Hybrid ANA/RPANA model	An ANA/RPANA model in which nonattendance is assumed to be independent between some, but not all attributes, i.e., $1 < A < K^{\star}$
IANA/RPIANA model	An ANA/RPANA model in which nonattendance is assumed to be fully independent across attributes, i.e., $A = K^{\star}$
RPANA model	Random parameters attribute nonattendance model
RPL model	Random parameters logit model
WTP	Willingness to pay

References

Alemu, M.H., Mørkbak, M.R., Olsen, S.B., Jensen, C.L., 2013. Attending to the reasons for attribute non-attendance in choice experiments. Environmental and Resource Economics 54 (3), 333–359.

Bliemer, M.C.J., Rose, J.M., 2011. Experimental design influences on stated choice outputs: an empirical study in air travel choice. Transportation Research Part A 45 (1), 63–79.

Cameron, T.A., DeShazo, J.R., 2011. Differential attention to attributes in utility-theoretic choice models. Journal of Choice Modelling 3 (3), 73–115.

Campbell, D., Hensher, D.A., Scarpa, R., 2011. Non-attendance to attributes in environmental choice analysis: a latent class specification. Journal of Environmental Planning and Management 54 (8), 1061–1076.

Campbell, D., Hensher, D.A., Scarpa, R., 2012. Cost thresholds, cut-offs and sensitivities in stated choice analysis: identification and implications. Resource and Energy Economics 34 (3), 396–411.

Campbell, D., Hutchinson, W., Scarpa, R., 2008. Incorporating discontinuous preferences into the analysis of discrete choice experiments. Environmental and Resource Economics 41 (3), 401–417.

Campbell, D., Lorimer, V., Aravena, C., Hutchinson, G., 2010. Attribute processing in environmental choice analysis: implications for willingness to pay. In: Agricultural Economics Society Annual Conference. Edinburgh.

Carlsson, F., Kataria, M., Lampi, E., 2010. Dealing with ignored attributes in choice experiments on valuation of Sweden. Environmental and Resource Economics 47 (1), 65–89.

Cirillo, C., Axhausen, K.W., 2006. Evidence on the distribution of values of travel time savings from a six-week diary. Transportation Research Part A 40 (5), 444–457.

Collins, A.T., Rose, J.M., Hess, S., 2012. Interactive stated choice surveys: a study of air travel behaviour. Transportation 39 (1), 55–79.

DeShazo, J.R., Fermo, G., 2004. Implications of Rationally-Adaptive Pre-choice Behavior for the Design and Estimation of Choice Models.

Gilbride, T.J., Allenby, G.M., Brazell, J.D., 2006. Models for heterogeneous variable selection. Journal of Marketing Research 43 (3), 420–430.

Hensher, D., Rose, J., Greene, W., 2012a. Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. Transportation 39 (2), 235–245.

Hensher, D.A., 2007. Reducing sign violation for VTTS distributions through endogenous recognition of an individual's attribute processing strategy. International Journal of Transport Economics 34 (3), 333–349.

Hensher, D.A., Collins, A.T., Greene, W.H., 2013. Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: a warning on potential confounding. Transportation 40 (5), 1003–1020.

Hensher, D.A., Greene, W.H., 2010. Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. Empirical Economics 39 (2), 413–426.

Hensher, D.A., Greene, W.H., Li, Z., 2011. Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings. Transportation Research Part B 45 (7), 954–972.

Hensher, D.A., Rose, J., Bertoia, T., 2007. The implications on willingness to pay of a stochastic treatment of attribute processing in stated choice studies. Transportation Research Part E 43 (2), 73–89.

Hensher, D.A., Rose, J., Greene, W.H., 2005. The implications on willingness to pay of respondents ignoring specific attributes. Transportation 32 (3), 203–222.

Hess, S., Hensher, D.A., 2010. Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. Transportation Research Part B 44 (6), 781–790.

Hess, S., Hensher, D.A., 2013. Making use of respondent reported processing information to understand attribute importance: a latent variable scaling approach. Transportation 40 (2), 397–412.

Hess, S., Rose, J.M., 2007. A latent class approach to modelling heterogeneous information processing strategies in SP studies. In: Oslo Workshop on Valuation Methods in Transport Planning.

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V., Caussade, S., 2013. It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. Transportation 40 (3), 583–607.

Hole, A.R., 2011. A discrete choice model with endogenous attribute attendance. Economics Letters 110 (3), 203–205.

Hole, A.R., Kolstad, J.R., Gyrd-Hansen, D., 2012. Inferred vs stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour, Working Paper. Sheffield Economic Research Paper Series. Department of Economics, University of Sheffield. ISSN 1749-8368..

Kragt, M.E., 2012. Attribute attendance in choice experiments: exploring issues of scale and attribute framing. In: 19th Annual EAERE Conference. Prague.

Mariel, P., Meyerhoff, J., Hoyos, D., 2011. Stated or inferred attribute non-attendance? A simulation approach. In: 2nd International Choice Modelling Conference, Leeds, United Kingdom.
Sæaelensminde, K., 2006. Causes and consequences of lexicographic choices in stated choice studies. Ecological Economics 59 (3), 331–340.
Scarpa, R., Gilbride, T.J., Campbell, D., Hensher, D.A., 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. European Review of Agricultural Economics 36 (2), 151–174.
Train, K., 2009. Discrete Choice Methods with Simulation, second ed. Cambridge University Press, New York.
Train, K., Sonnier, G., 2005. Mixed logit with bounded distributions of correlated partworths. In: Scarpa, R., Alberini, A. (Eds.), Applications of Simulation Methods in Environmental and Resource Economics. Springer, pp. 117–134.