Do environmental concerns affect commuting choices?: hybrid choice modelling with household survey data

Jennifer Roberts and Gurleen Popli

University of Sheffield, UK

and Rosemary J. Harris

Queen Mary University of London, UK

Summary. To meet ambitious climate change goals governments must encourage behavioural change alongside technological progress. Designing effective policy requires a thorough understanding of the factors that drive behaviours. In an effort to understand the role of environmental attitudes better we estimate a hybrid choice model (HCM) for commuting mode choice by using a large household survey data set. HCMs combine traditional discrete choice models with a structural equation model to integrate latent variables, such as attitudes, into the choice process. To date HCMs have utilized small bespoke data sets, beset with problems of selection and limited generalizability. To overcome these problems we demonstrate the feasibility of using this valuable modelling approach with nationally representative data. Our results suggest that environmental attitudes have an important influence on commute mode choice, and this can be exploited by governments looking to add to their climate change policy toolbox in an effort to change travel behaviours.

Keywords: Environment; Hybrid choice model; Structural equation modelling

1. Introduction

Tackling climate change is one of the most important challenges faced by governments around the world. The USA has committed to reducing greenhouse gas emissions by 17% below 2005 levels by 2020, and in the UK the Climate Change Act 2008 commits the government to cut emissions by at least 80% by 2050. Achieving these targets will not be possible via technical progress alone; it will also require a substantial behavioural change on the part of individuals and households. A prerequisite to designing effective policy interventions is a thorough understanding of the factors that drive behaviours and ultimately decisions. One topic that has been the subject of much discussion is the extent to which individual environmental concerns can motivate behaviour change. The majority of the literature is pessimistic in this regard; although the public express concern about climate change, this is rarely sufficiently strong to bring about change towards more sustainable behaviours, especially when these changes require personal sacrifice (Gifford, 2011). Nowhere is this more apparent than in our seemingly unshakeable attachment to the private car.

Address for correspondence: Jennifer Roberts, Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, UK.
E-mail j.r.roberts@shef.ac.uk

In this paper we use hybrid choice modelling to explore the effects of environmental concerns on choice of commuting transport mode in England. Hybrid choice models (HCMs) combine traditional discrete choice models (DCMs) with a structural equation model (SEM) to integrate latent variables, such as attitudes and other psychological constructs, into the choice process. Our overall aim is to improve our understanding of the way that people make travel choices; and we make three main contributions to the literature. Firstly, our study is a rare example of a study which attempts to evaluate the importance of environmental beliefs for travel behaviours. Secondly, a major innovation is the use of a large nationally representative household survey data set for model estimation; we also replicate the modelling with a second such data set, as a robustness check on the results. To date HCMs have been estimated by using relatively small data sets constructed to tackle the question in hand. These bespoke data have limited generalizability, are prone to substantial selection problems and focusing effects, and include little information on individual characteristics with which to control for heterogeneity. Thirdly, HCM studies of mode choice generally devote little or no attention to the theoretical model of decision making that underlies the empirical work; in contrast we explain how the *attitude–behaviour–context* (ABC) model (Stern, 2000; Stern and Oskamp, 1987) is an appropriate framework for our HCM of commuting mode choice incorporating latent environmentalism.

Domestic transport accounts for 25% of the UK's carbon dioxide emissions, more than half of which are from the private car (Department for Transport, 2008). Thus meeting climate change goals necessitates a shift away from the car and towards more sustainable modes such as public transport, walking and cycling. The regular commuting journey is a key arena in which to study these choices; 57% of all commute trips are by car (Department for Transport, 2012). A recent report for the UK Department for Transport reveals that the implications of climate change are not widely understood and that most people are unaware of their own contribution to the problem (King *et al.*, 2009). However, knowledge alone is not an adequate antecedent to behavioural change; in their systematic review of interventions to reduce car use Graham-Rowe *et al.* (2011) found no evidence that providing environmental information is effective.

Traditionally transport choice modelling has employed DCMs; these are based in random-utility theory, which is an economic framework in which time and cost are the key variables (see for example Train (1980)). Random-utility theory has been criticized for its fundamental assumption that consumers are a rational 'optimizing black box' (Morikawa *et al.*, 2002). HCMs were first proposed as an extension to random-utility theory in the 1980s, as a way of better understanding consumer behaviour by incorporating latent variables, such as preferences and attitudes, in the choice process (McFadden, 1986; Ben-Akiva and Boccara, 1987). More broadly HCMs can be seen as a reflection of the growing popularity of behavioural economics, which incorporate psychological concepts into economic analysis to improve our understanding of decision making under uncertainty (Tversky and Kahneman, 1974).

Empirical applications of HCMs have developed largely from 2000 onwards, and Golob (2003) provided coverage within his excellent review of SEMs in transport research. Morikawa *et al.* (2002) found a significant influence of latent variables for comfort and convenience on the decision to use rail or car for intercity travel between cities in the Netherlands. Temme *et al.* (2008) found that latent preferences for comfort, convenience, flexibility and safety affect the travel mode choices of a market research survey panel in Germany. Yáñez *et al.* (2010) considered the effects of three latent variables (accessibility, reliability and comfort) on commute mode following the introduction of a new urban transport system in Santiago, Chile. Córdoba and Jaramillo (2012) demonstrated the importance of a 'personality measure' on the commute mode choices of staff at the National University of Columbia.

Johansson *et al.* (2006) is the only HCM study that we know of that has incorporated any measure of environmentalism in a mode choice model. The data were from a postal survey of commuters in Sweden, and the environmentalism variable was inferred from measures of the frequency with which the respondents recycle glass, paper, batteries and metal. This variable was found not to be significant in the choice of car *versus* bus but has marginal significance in the train–bus choice. This neglect of environmental variables is a serious shortcoming given the key role of personal travel choices in climate change. Environmental attitudes are likely to influence the utility that an individual derives from different travel modes and hence ultimately may affect mode choice. The relative importance of environmental attitudes alongside other influences such as fiscal incentives is of key interest to policy makers.

2. Decision-making model

Existing empirical applications of HCMs tend not to be based in clear theoretical frameworks and thus it is often difficult to interpret the results, especially in relation to inferring causal relationships. This is particularly problematic for SEMs, because these models do not provide a means of establishing causality but rather can only confirm relationships that the researcher must impose from external knowledge (Sánchez *et al.*, 2005; Bollen and Pearl, 2013). Environmental concerns reflect how we feel about the environment and the way that we are predicated to behave with regard to it. These are complex phenomena that combine elements of prosocial preferences, risk and time preference, selfish regard for one's own (and one's children's) future, social pressures and norms. 'Environmentalism' is mediated by knowledge and institutions, which influence the immediate costs to individuals; it also involves interaction between attitudes and behaviours.

In our HCM we propose that 'environmentalism' is a latent construct that we cannot observe directly. Instead it is represented by a set of observable indicators that measure both *attitudes* towards the environment and climate change, as well as certain environmental *behaviours*; these behaviours are not directly related to commuting but relate to other areas of life such as recycling, and use of carrier bags and home energy. These indicators are used in an SEM, which is combined with a DCM to integrate the latent variable(s) for environmentalism into a model for commuting mode choice.

The psychological literature explains that attitudes and behaviours are related but theoretically distinct. A *behaviour* is an observable action, e.g. switching a tap off rather than letting it drip, or putting on an extra jumper rather than turning the heating up. *Attitudes* are the subjective importance that is attached to different issues, e.g. the extent to which a person believes that climate change is a cause for concern, or the extent to which they believe that the environmental crisis has been exaggerated. The attitude–behaviour relationship is a core topic in psychology (Kraus, 1995); in general it is understood that behaviours are driven by intention, and intention is, in turn, a function of attitudes (Ajzen and Fishbein, 1977). For example, people with pro-environmental attitudes might have a strong intention not to use the car for short trips and hence act in a way that is consistent with this, choosing instead to use public transport or to walk or cycle. However, there may also be discrepancies between attitudes and behaviours (Ajzen and Fishbein, 1970) and this has gained empirical support within the environmental context (Oskamp *et al.*, 1991; Gardner and Abraham, 2008).

Kline (1988) stressed the importance of contextual factors that can weaken the attitude–behaviour connection, arguing that people will be less willing to act in a proenvironmental way when this is costly or inconvenient, or when they do not feel that their personal contribution can make much difference and when they perceive that others are not behaving that way (Oskamp *et al.*, 1991). Also, many people believe that the government is responsible for solving

environmental problems and rely on this to justify their own behaviours (Stern *et al.*, 1985). Although it is usually thought that attitudes precede behaviour, behaviours can change attitudes; for example, individuals with proenvironmental beliefs who nevertheless use the car for short trips might change their attitudes in an attempt to rationalize their mode choice and to reduce cognitive dissonance (Festinger, 1962).

There are various theoretical models of decision making that support the integration of psychological variables into transport mode choice models. Gardner and Abraham (2010) tested Ajzen's (1991) *theory of planned behaviours* for local car use in a small UK city (Brighton and Hove). Bamberg and Schmidt (2003) compared the theory of planned behaviours with the *theory of interpersonal behaviour* (Triandis, 1977) and the *norm activation model* (Schwartz, 1977) for car use in a sample of students. Neither of these studies found much support for the influence of environmental attitudes; perceived personal benefits such as convenience outweigh environmental opinions and car use is so habitualized that there is little or no moral dimension to the choice.

The ABC model (Stern, 2000; Stern and Oskamp, 1987) was developed specifically to explain environmentally significant behaviours and provides an ideal theoretical framework for our HCM of commuting mode choice. Behaviour in this model is an interactive product of 'internal' attitudes, such as concerns over climate change, and 'external' contextual factors such as the transport costs, and institutional constraints such as the local availability of transport choices. Hence external factors (like time and cost) will moderate the effect of environmental beliefs, and the relative importance of psychological and contextual factors will depend on the behaviour in question. Attitudes have been found to have stronger effects for low constraint behaviours that are cheap or easy to change, such as curbside recycling or the use of low energy light bulbs (Stern and Oskamp, 1987; Guagnano *et al.*, 1995). We would expect them to have less influence on behaviours like car use, which have high personal benefits, are habitualized and are seen as difficult to change (Collins and Chambers, 2005). Nevertheless the relative influence of these different sets of factors is a key issue for designing policies to change behaviours, and this is where our study can make a clear contribution.

A schematic diagram of our HCM is shown in Fig. 1; this illustrates how a traditional DCM is combined with a latent variable model for environmentalism. The unobservable latent variable(s) for environmentalism are identified via observed indicators that reflect environmental *attitudes* or *behaviours*; the number of latent variables and the classification of indicators are determined via both exploratory and confirmatory factor analyses, as explained in the next section. Environmentalism is also determined by observed sociodemographic characteristics such as age, sex and household income; these are important *context* variables; for example whether or not the individual has young children will influence how much personal (in)convenience they might experience from using public transport rather than a car. Latent utility from commuting mode is determined by environmentalism and also directly by sociodemographic variables and the key mode attributes of time and cost, which are again important measures of *context*. We observe the final mode choice decision as a manifestation of the underlying latent utility. The statistical basis of this model and its estimation are explained in the next section.

The programs that were used to analyse the data can be obtained from

```
http://wileyonlinelibary.com/journal/rss-datasets
```

3.   Specification and estimation of hybrid choice model

3.1.   Structural model

u_{ij}^* is the unobserved (latent) conditional indirect utility for mode j ($j = 0, \ldots, J - 1$) for individual i ($i = 1, \ldots, n$). In the empirical analysis we assume a linear-in-parameters specification:
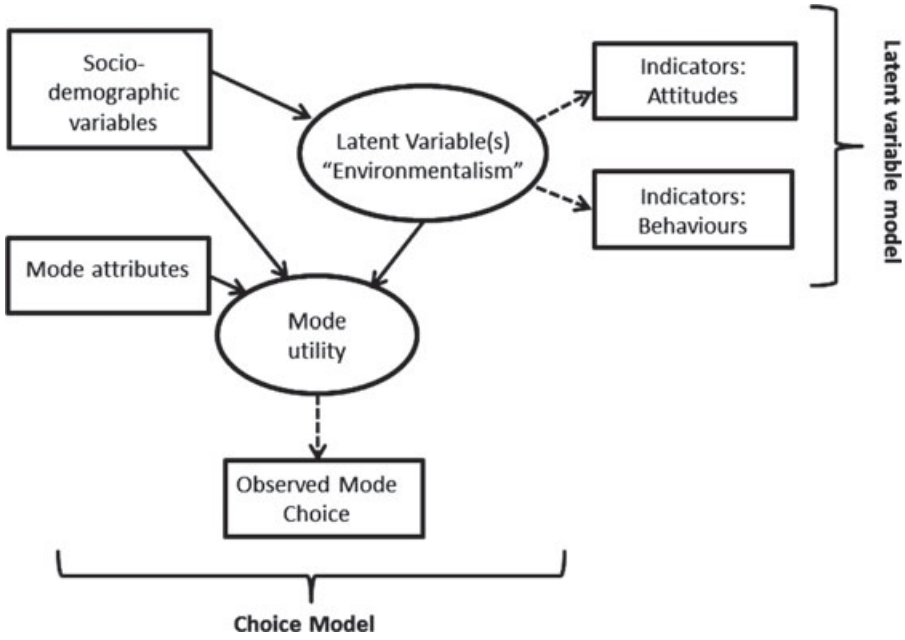
Fig. 1. HCM for commuting mode (adapted from Ben-Akiva *et al.* (2002)): □, observed variables; ○, unobservable variables; ⟶, cause-and-effect relationships; - - ⟩, measurement equations

$$\mathbf{u}_j^* = \mathbf{x}' \beta_{jx} + \mathbf{z}' \beta_{jz} + \boldsymbol{\eta}' \beta_{j\eta} + \nu_j \tag{1}$$

where $\mathbf{u}_j^*$ is an $n \times 1$ column vector of individual utilities u_{ij}^*; x_i is a $K_x \times 1$ vector of K_x individual-specific observed variables, such as age and income, and $\mathbf{x}$ is a $K_x \times n$ matrix obtained by horizontal concatenation of x_i; z_i is a $K_z \times 1$ vector of observed *context* variables for individual i, relating to mode, such as availability of local public transport, and $\mathbf{z}$ is a $K_z \times n$ matrix obtained by horizontal concatenation of z_i; η_i is a $Q \times 1$ vector of Q individual-specific latent variables, which represent the unobservable environmentalism of the individual (environmental *attitudes* and *behaviours*), and $\boldsymbol{\eta}$ is a $Q \times n$ matrix obtained by horizontal concatenation of η_i; β_{jx}, β_{jz} and $\beta_{j\eta}$ are respectively $K_x \times 1$, $K_z \times 1$ and $Q \times 1$ vectors of parameters to be estimated; ν_j is an $n \times 1$ column vector of random mean 0 error terms. The latent variables in matrix $\boldsymbol{\eta}$ are also assumed to depend linearly on the vector of individual-specific observed variables:

$$\boldsymbol{\eta} = \boldsymbol{\gamma}' \mathbf{x} + \boldsymbol{\xi} \tag{2}$$

where $\boldsymbol{\gamma}'$ is a $Q \times K_x$ matrix of parameters to be estimated and $\boldsymbol{\xi}$ is a matrix of $Q \times n$ random mean 0 error terms.

3.2. Measurement model

We do not observe u_{ij}^*; what we observe is the choice that is made by the decision maker whether to use mode j or an alternative l ($j, l \in J$). The observed decision variable is defined as

$$d_i = \begin{cases} j & \text{if } u_{ij}^* > u_{il}^*, \quad \forall \, l \in J, \\ 0 & \text{otherwise.} \end{cases}$$

In our empirical analysis $J = 2$; the individual either commutes by car ($j = 1$) or by public transport ($j = 0$), so without loss of generality we can assume that $u_{i0}^* = 0$. In subsequent discussion

we suppress the indexation for the mode; thus u_{i1}^* is denoted as u_i^*. The decision of the individual is modelled as

$$P(d_i = 1|x_i, z_i, \eta_i) = P(u_i^* > 0|x_i, z_i, \eta_i) = P(\nu_i > -x_i'\beta_x - z_i'\beta_z - \eta_i'\beta_\eta)$$
$$= 1 - F(-x_i'\beta_x - z_i'\beta_z - \eta_i'\beta_\eta) \tag{3}$$

where $F(\cdot)$ is the cumulative distribution function for the measurement error ν. We treat $F(\cdot)$ as a normal distribution function, and therefore estimate a probit model.

Let $\boldsymbol{\eta}' = (\eta^1, \ldots, \eta^Q)$, where η^q is an $n \times 1$ vector which contains as components the latent variables for all the individuals. We do not directly observe η^q ($q \in \{1, \ldots, Q\}$); what we do observe are different indicators for η^q. So, for example, we do not observe directly how 'green' people are; what we do observe are their responses to questions reflecting their environmental *attitudes* and *behaviours*, and these can be considered indicators of their underlying environmentalism. Let Y_s^q be an $n \times 1$ vector of the indicators for η^q, where $s = 1, \ldots, m^q$, such that $m^q \geqslant 2$, i.e. we need a minimum of two indicators for each latent variable. The observed indicators are related to the unobserved latent variable as

$$Y_s^q = \mu_s^q + \alpha_s^q \eta^q + \varepsilon_s^q \qquad \forall q \in \{1, \ldots, Q\}, s \in \{1, \ldots, m^q\} \tag{4}$$

where the number α_s^q is the factor loading from the factor analysis, which can be interpreted as the amount of information that the indicator Y_s^q contains about η^q; ε_s^q is an $n \times 1$ vector of zero-mean measurement errors, which captures the difference between the observed indicators and the unobserved variable; the intercept μ_s^q is an $n \times 1$ vector with all elements equal, i.e. no dependence on individual.

We have a set of measurement equations, similar to equation (4) for each of the latent variables in the matrix $\boldsymbol{\eta}$. In matrix notation:

$$\mathbf{Y} = \boldsymbol{\mu} + \boldsymbol{\eta}'\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$$

where $\mathbf{Y}$ is a matrix of $n \times M$ indicators, such that $M = m^1 + m^2 + \ldots + m^q$; $\boldsymbol{\alpha}$ is a matrix of $Q \times M$ factor loadings; $\boldsymbol{\varepsilon}$ is an $n \times M$ matrix of measurement errors; $\boldsymbol{\mu}$ is an $n \times M$ matrix containing the M intercepts. For example, if we assume $Q = 2$, $m^1 = m^2 = 2$, then:

$$(Y_1^1 \quad Y_2^1 \quad Y_1^2 \quad Y_2^2) = (\mu_1^1 \quad \mu_2^1 \quad \mu_1^2 \quad \mu_2^2) + (\eta^1 \quad \eta^2)\begin{pmatrix} \alpha_1^1 & \alpha_2^1 & 0 & 0 \\ 0 & 0 & \alpha_1^2 & \alpha_2^2 \end{pmatrix} + (\varepsilon_1^1 \quad \varepsilon_2^1 \quad \varepsilon_1^2 \quad \varepsilon_2^2)$$

where, as above, Y_s^q, η^q, ε_s^q and μ_s^q are themselves $n \times 1$ vectors.

The factor loadings in equation (4) can be identified only up to a scale, so we normalize them according to $\alpha_1^q = 1 \, \forall q \in \{1, \ldots, Q\}$. Further we cannot separately identify the mean of the latent variables, $E(\boldsymbol{\eta})$, and intercepts $\boldsymbol{\mu}$: we need to normalize one of them; we assume that $E(\boldsymbol{\eta}) = 0$ and identify $\boldsymbol{\mu}$.

We employ both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to explore and verify the latent variables that are to be used in our HCM. All model estimation and factor analyses are carried out by using Mplus version 7.11 (see below). In the EFA no preconceived structure is imposed, and the indicators are allowed to load freely, thus determining the dimension of matrix $\boldsymbol{\eta}$ (i.e. the value of Q). Factor extraction in the EFA is done via varimax rotation (Kaiser, 1958) and model selection criteria and diagnostic statistics are discussed in Section 5. In contrast, in the CFA we constrain the model to comply with prior beliefs based on evidence from the psychological literature that attitudes and behaviours are theoretically distinct both generally (Ajzen and Fishbein, 1977, 2005; Kraus, 1995) and in the context of environmental behaviours (Stern, 2000). Thus the indicators are split into two vectors ($Q = 2$)

Table 1. Definitions of the variables

Variable	Definition
Commute mode choice	Dichotomous: car, 1; public transport, 0
Commute time	In minutes, one way for regular commuting journey
Gender	Female, 1; male, 0
Age	In years
Education	Set of dichotomous variables for highest level of attainment: degree or higher, A level, O level; base category, no qualifications
Household income	log(annual equivalized household income) in pounds sterling
Children	Set of dichotomous variables for whether there are any children in the household aged 0–2, 3–4, 5–11 or 12–15 years; base category, no school age or preschool children
Married	Dichotomous: married or living as a couple, 1; single, 0
Health	Ordinal self-assessed health indicator: 1, poor; 2, fair; 3, good; 4, very good; 5, excellent
Public transport quality	Public transport quality for each local authority: travel time to nearest town centre, by car or by public transport (source, Department for Transport accessibility data for 2009, Table ACS0408-2009)
Average traffic speed	Average vehicle speed (flow if weighted) during the weekday morning peak on locally managed 'A' roads for each local authority, miles per hour (source, Department for Transport data on journey reliability, Table CGN0206a); mean of the monthly figures from 2009
Indicators—behaviours	All behaviours indicators are increasing in environmentalism
TV	Does not leave television on standby overnight: frequency, 5-point scale
Lights	Switches off lights in empty rooms: frequency, 5-point scale
Tap	Does not leave tap running when brushing teeth: frequency, 5-point scale
Heating	Wears extra clothes rather than turn heating up: frequency, 5-point scale
Packing	Does not buy goods with excess packaging: frequency, 5-point scale
Energy	Has solar or wind energy or green tariff: frequency, 3-point scale
Produce	Buys recycled products: frequency, 5-point scale
Bags	Takes own bags for shopping: frequency, 5-point scale
Recycle	Separates rubbish for recycling: frequency, 4-point scale
Flights	Takes fewer flights where possible: frequency, 5-point scale
Indicators—attitudes	All attitudes indicators are increasing in environmentalism
Own life	Leads an environmentally sympathetic life: extent of agreement, 5-point scale
Alternative	Being green is 'alternative lifestyle': extent of agreement, 4-point scale
Own resp	Own behaviour contributes to climate change: agree or disagree
Pay more	Prepared to pay more for environmentaly sympathetic products: agree or disagree
Disaster	World on course for major environmental disaster: agree or disagree
Exaggerate	The environmental crisis has been exaggerated: agree or disagree
Control	Climate change is beyond our control: agree or disagree
Future	The effects of climate change are too far in the future: agree or disagree
Lifestyle	Changes made have to fit in with current lifestyle: agree or disagree
Others	Not worth doing anything unless others do the same: agree or disagree
Britain	Not worth UK trying to do anything about climate change: agree or disagree
30 years	Climate change will affect UK in next 30 years: agree or disagree

†Source: unless otherwise stated variables are obtained from the UK Household Longitudinal Study, wave 1.

according to the question wording, with one set forced to load onto an attitudes factor, and the other onto a behaviours factor. The definitions and split are provided in Table 1.

Equations (1) and (3) give the standard DCM, and equations (2) and (4) give the latent variable model; together these equations define the HCM. It is worth pointing out here that an alternative specification to the HCM would be to include the indicator variables Y directly in the DCM instead of including the latent variables η; this is analogous to treating the indicators as direct measures of environmentalism rather than as functions of it. This is inappropriate for three main reasons. Firstly, the indicator variables may be correlated with the errors from the DCM due to omitted (unobservable) effects and this would lead to endogeneity bias. Secondly,

the latent variables that the indicators represent are measured with error; thus their direct inclusion in the DCM can lead to inconsistent estimates (Ashok *et al*., 2002). Thirdly, the HCM specification is a closer representation of the psychological decision-making framework. Attitudes are inherently latent and strong agreement with a proenvironmental statement does not necessarily translate into a causal relationship with choice. Attitudes and related behaviours are therefore not direct antecedents of mode choice but are indirectly related via, *inter alia*, latent environmental concerns (Daly *et al*., 2012).

A further advantage of our approach arises from simultaneous estimation of the DCM and latent variable parts of the model. It is also possible to use an SEM for environmentalism and a separate choice model, which includes latent variables from the SEM, but to estimate these models sequentially (this is the approach that was taken by for example Johansson *et al*. (2006) and Choo and Mokhtarian (2004)). However, in this case the SEM would not use information on the observed choices to inform the latent variable part of the model, whereas simultaneous estimation of the choice part and latent variable part makes fuller use of the information and hence is more efficient than a sequential approach (Morikawi *et al*., 2002; Daly *et al*., 2012).

3.3. Identification, estimation and diagnostic statistics

To be able to identify the parameters in the system of equations (1)–(4) we need to make the following assumptions.

Assumption 1. The error term ν is independent of $\mathbf{x}$, $\mathbf{z}$ and $\boldsymbol{\eta}$, $\boldsymbol{\xi}$ is independent of $\mathbf{x}$ and ε is independent of $\boldsymbol{\eta}$.

Assumption 2. The errors terms $\nu, \boldsymbol{\xi}$ and ε are not correlated with each other.

Assumption 3. The variance–covariance $E(\nu\nu')$ is diagonal with off-diagonal elements 0 and, similarly, there are no correlations between the different matrix elements of $\boldsymbol{\xi}$ and ε.

The system of equations (1)–(4) that comprise the HCM are estimated simultaneously, using the asymptotically distribution-free weighted least squares (WLS) estimator (Browne, 1984; Muthén, 1983, 1984). WLS is chosen over the more commonly used maximum likelihood approach because the latter requires the indicator variables to be continuous and this is not so in our application, as we have a number of dichotomous and ordinal variables (see the next section). The WLS estimator for categorical indicator variables works in three steps. In the first step, a set of probit regressions are run for each categorical indicator (i.e. all the observed indicators that are given by equation (4) and the observed mode choice dummy d_i), followed by a set of bivariate probit regressions for each pair of categorical indicators. The thresholds, for the measurement equations (4) and the mode choice equation (3), are obtained from these probit regressions. In the second step the estimated thresholds, conditional mean of the indicators and the conditional variance–covariance matrix are used to form weighting matrix $\mathbf{W}$. In the third step $\mathbf{W}$ is used to estimate the parameters by using the WLS method. The WLS function is optimized numerically by using an iterative quasi-Newton technique. The estimator is distribution free in the sense that the final estimated parameters of the model (other than the thresholds) do not require normality.

The identification of the model and the asymptotic properties of the WLS estimator are discussed in Muthén and Satorra (1995); and the exact form of the WLS function that is used is given in Muthén and Muthén (2016). Estimation is carried out by using the `WLSMV` estimator in Mplus version 7.11. As a check of consistency, we also replicated estimates for one of our data sets (the British Household Panel Survey—see below) using the `gsem` estimator in Stata

version 13 (StataCorp., 2013). We cannot rely on the asymptotic properties of WLS when the sample size is small but (as described in the next section) our estimation samples are over 6000, which is unusually large for HCM applications. Given the structure of our data (see Section 3) the estimation takes account of clustering of individuals in households.

As is common in the SEM literature we rely on a number of diagnostic statistics to determine the adequacy of model fit: firstly, the root-mean-square error of approximation RMSEA, which shows the amount of unexplained variance (Steiger and Lind, 1980) and ranges from 0 to 1 with smaller values indicating better fit; secondly, the comparative fit index CFI, which considers the discrepancy between the data and the hypothesized model, while adjusting for sample size (Bentler, 1990); thirdly, the Tucker–Lewis reliability index TLI, which is an adjusted version of the normed fit index of discrepancy between the χ^2-value of the hypothesized model and the χ^2-value of the null model (Tucker and Lewis, 1973). Both CFI and TLI range from 0 to 1 with larger values indicating better model fit. Hu and Bentler (1999) suggested that acceptable model fit requires RMSEA < 0.06 and both CFI and TLI greater than 0.90. We also provide the χ^2-test of model fit for the baseline model, which tests the null hypothesis that all slope parameters in the structural part of the model are 0 and the factor loadings in the measurement part of the model are all 1; for good model fit we would wish to reject this null hypothesis. It is worth noting here that a standard SEM would normally report the χ^2-test for model fit, which tests for differences between the observed and expected covariance matrices. This test is not valid for the WLS estimator, because the distributional assumptions are violated. In addition it is not appropriate for our large sample, as the probability of rejecting the null hypothesis increases with sample size (Jöreskog, 1969). In addition to these formal tests model validity is also judged on the basis of the parameter estimates; specifically whether the estimates pass the 'sense test' in that they accord with expectations from theory and previous empirical findings. We also replicate the modelling with two different data sets as a further check on the robustness of our results.

4.   Data

Our main data come from the first wave of the UK Household Longitudinal Study (UKHLS), which is otherwise known as 'Understanding society'; a nationally representative survey of approximately 40 000 households (University of Essex, 2012). Data are obtained from face-to-face interviews with all adults in each household and cover various topics including personal background, economic circumstances, family relationships, health and wellbeing, as well as expectations, aspirations and opinions on a variety of issues. Wave 1 interviews were carried out in 2009–2010 and include a module on environmental attitudes and behaviours. Our analysis sample is restricted to people who commute to work regularly in England. It is necessary to restrict the sample to England because we cannot obtain comparable area level transport data (see below) for Scotland, Wales and Northern Ireland. We also restrict the sample to respondents who live in an urban area (as defined by the Office for National Statistics classification), have access to a car and who commute for up to 120 min each way by car or public transport. These restrictions ensure that it is reasonable to assume that the respondents have some choice over their commuting mode. The resulting sample size is $n = 13\,139$; 6883 women and 6256 men (Table 2). This is contrasted with the relatively small bespoke data sets that have been used in previous HCM studies; for example Johansson *et al.* (2006) analysed 811 responses to a postal survey on mode choice for one specific route in Sweden and Yáñez *et al.* (2010) used data for 303 individuals working at university campuses in Santiago, Chile.

Table 2.   Construction of the analysis sample from UKHLS wave 1

Inclusion and exclusion criteria	Number of individuals
Total adults interviewed in the UKHLS	50994
... with full face-to-face interviews†	47732
... who commute to work regularly	22277
Do not live in an urban area	4364
Have no access to a car	2620
Commute for > 120 min each way	54
Do not commute by car or public transport	2075
Missing values on key variables	25
Analysis sample	13139
Women	6883
Men	6256

†Other interviews (by telephone and proxy) do not cover the full set of interview topics.

Several advantages follow from our use of household survey data. Firstly, larger sample sizes give us more statistical power to detect effects; this is particularly valuable for the factor analysis, resulting in more stable estimates (MacCallum *et al.*, 1996). Secondly, our results are generalizable to the population of commuters in England and not specific to a particular journey setting; thirdly, the data contain a rich set of individual and household characteristics for use as control variables. Finally, both the potential endogeneity between attitudes and choices, and the influence of focusing effects are minimized because the UKHLS is a general household survey, rather than a survey that is focused on commuting or the environment; the questions on mode choice and those on environmental attitudes and behaviours occur in separate sections of the survey with no apparent links between the two. Focusing effects mean that questions can elicit misleading responses; it is highly likely that, when people are asked about their environmental attitudes and commuting choices in a survey that is designed to explore the link between the two, they will overstate the importance of the influence of the environment and offer consistent answers in an effort to rationalize their behaviours. Further, individuals' attitudes can be affected by their mode choices since they may modify their attitudes to reduce the cognitive dissonance arising from inconsistent attitudes and behaviours; attitudes can be altered *ex post* whereas behaviours cannot be. If this was so then the latent construct for environmentalism would be endogenously determined, but this is unlikely in our data because of the nature of the household survey. It is also worth stressing here that, in one respect, endogeneity is an inherent part of the modelling framework that we use. The HCM framework explicitly recognizes that both behaviour (mode choice) and attitudes (the responses to the indicator questions) are driven by the same underlying latent variable (that we term environmentalism).

These advantages come with one shortcoming; whereas we have extremely rich information about individuals and households, we have only limited information about the journeys in question (for example we have time and mode but not monetary cost) and in particular we do not have information about the characteristics of the mode that is not chosen; so for example if someone chooses to commute by car we do not know how long that specific journey would take by public transport. Given that journey time and cost are key variables in any mode choice model, we overcome this by matching in area level data on local transport context to construct proxies for journey time and cost; specifically we include measures of the availability of local public transport and the amount of local traffic congestion (see below).

A list of all variables with definitions is provided in Table 1. Our choice outcome variable is usual commuting mode for the regular journey to work; a binary variable where 1 represents car and 0 represents public transport. We also have an average one-way travel time for this journey, in minutes, which we include in some of our models. Our indicator variables Y_s^q are a set of responses to questions on environmental behaviours and attitudes. There are 10 questions on behaviours, which ask things like 'do you leave the television on standby overnight?', 'do you wear extra clothes rather than turning the heating up?' and 'do you buy recycled products?'. The responses reflect frequency of engaging in that behaviour and most of the questions have a five-point scale that ranges from 'never' to 'always'; these indicators are all coded to be increasing in environmentalism. These questions were selected for inclusion in the UKHLS because they cover

> '... several issues which, collectively, influence a considerable proportion of greenhouse gas emissions and other resource use resulting from individual activity'

(Lynn and Longhi (2011), page 2). Longhi (2013), for example, used these data and found that women have higher proenvironmental behaviour than men, and having a university degree also has a positive correlation with proenvironmental behaviour. MacPherson and Lange (2013) used one of the indicators, exploring the *determinants of green electricity tariff uptake in the UK*. Thomas *et al*. (2016) again used only one indicator to evaluate the effects of the introduction charges for carrier bags in Wales on *own bag use*. There are also 12 questions on environmental attitudes. The majority of these ask whether the respondent agrees or disagrees with statements like 'climate change is beyond our control', 'it's not worth me doing things to help the environment if others do not do the same' and 'any changes I make to help the environment need to fit in with my lifestyle'. Two of the attitudes questions have ordinal responses: 'How would you best describe your current lifestyle?' has a five-point scale, where 1 represents not really doing anything sympathetic to the environment and 5 represents being sympathetic to the environment in everything that they do; 'being green is an alternative lifestyle', has a four-point scale where 1 represents disagree strongly and 4 represents agree strongly. These indicators are all coded to be increasing in environmentalism. The *attitudes* questions were chosen for inclusion in the UKHLS to replicate, as far as a possible, questions that were used in the 2007 and 2009 Department for Environment, Food and Rural Affairs surveys of public attitudes to the environment (Department for Environment, Food and Rural Affairs, 2007; Thornton *et al*., 2010). The Department for Environment, Food and Rural Affairs largely chose questions from the revised 'New ecological paradigm' battery of statements. (Dunlap *et al*., 2000). The 'New ecological paradigm' was developed to tap into '... primitive beliefs about the nature of the earth and humanity's relationship with it' (Dunlap *et al*. (2000), page 427) and has become a widely used measure of proenvironmental orientation. The developers have carried out extensive validity and reliability testing, and concluded that the items can be treated as an internally consistent summary rating scale which strongly discriminates between known environmentalists and the general public.

To capture information on local transport context we use geographical identifiers that are available under the special licence for the 'Understanding society' survey; these show the local authority that each household is in. We use these identifiers to match in information about local transport conditions provided by the Department for Transport. We derive two variables from the information that is available. The first is the average traffic speed during the rush hour, which is a proxy for the amount of traffic congestion in the local area. The second is the journey time to the nearest town centre by car relative to public transport; this is a proxy for the availability of public transport locally. We would expect the utility that is derived from choosing the car for commuting, relative to public transport, to be higher if there is less congestion and lower

the better the availability of public transport. Although we do not include monetary cost in our model because of a lack of available data, these relative journey time variables can be considered as proxy variables for the economic concept of opportunity cost, or cost in terms of time.

Other control variables include age (in years), household income, highest educational attainment, whether or not the household contains children (in various age groups), self-reported health and marital status. All estimation is carried out for men and women separately given the evidence from previous literature that men and women differ in their commuting behaviour (Roberts *et al*., 2011) and their environmental beliefs (Anable *et al*., 2006); thus we expect sex differences to impact significantly on the overall HCM once other explanatory variables have been taken into account.

Given the novelty of using household survey data to estimate an HCM, replication is an important step in the model validation process. All of our modelling is replicated by using the BHPS. The BHPS was an annual longitudinal household panel that ran from 1991 to 2008 and had a very similar design to the UKHLS, with a similar set of interview questions, including, in 2008, a module on environmental attitudes and behaviours. The main difference between the two data sets is that the BHPS is smaller; in total 13 454 face-to-face interviews were carried out in 2008, and our analysis sample (given the selection described above) is 830 women and 900 men.

5. Results

Descriptive statistics are presented in Table 3. 70% of the women in our sample commute by car, and 73% of men; men's average one-way journey time is slightly longer at 28 min compared with just under 24 min for women. Highest educational achievement is similar for both sexes, as is household income; household incomes are highly skewed and are used in log-equivalized form in the models below. 70% of women are married or living as a couple compared with 76% of men. Average health for both men and women is 3.65 on a scale where 1 is poor and 5 is excellent. Average traffic speed on main roads during rush hour is just over 24 m.p.h.; given that the maximum speed on these roads is between 40 and 70 m.p.h., the rush hour averages are relatively slow and represent high levels of congestion; the range is wide from 9.4 to 39.2 m.p.h. Public transport quality (measured as the time that it takes to travel to the nearest town centre by car relative to public transport) suggests that it is on average three times faster by car, with a range between two and seven times faster; this reflects extremely variable public transport quality across the local authorities.

Men and women appear to be very similar in terms of their environmental behaviours; this may be because many of the behaviours are determined at the household level and the majority of our sample is living as a couple. The biggest difference is in taking own bags for shopping, which women are more likely to do than men. Switching lights off in empty rooms is very common, as is taking own bags for shopping, not leaving the television on standby overnight and separating rubbish for recycling. In contrast not buying goods with excessive packing, having green energy or a green tariff and taking fewer flights where possible are much less prevalent behaviours. There are more differences between the sexes in attitudes than behaviours. Women are more likely to think that the world is on course for environmental disaster. However, they are also more likely to think that the environmental crisis has been exaggerated, that it is not worth doing anything about climate change unless others do the same, and similarly that it is not worth the UK doing anything. It is not common for men and women to think that they lead an environmentally sympathetic life; the average score is around 2.6, on a scale where 1 represents not really doing anything environmentally sympathetic and 5 represents being environmentally sympathetic in

Table 3. Descriptive statistics

Variable	Results for women			Results for men		
	Mean	*Minimum*	*Maximum*	*Mean*	*Minimum*	*Maximum*
Commute by car (not public transport)	0.70	0	1	0.73	0	1
Journey time, one way (min)†	23.88	1	120	28.27	1	120
Age (years)	40.67	16	83	41.07	16	81
Highest education						
Degree or above	0.45	0	1	0.43	0	1
A level	0.20	0	1	0.21	0	1
O level	0.28	0	1	0.25	0	1
Household income (£ previous month)‡	4207	116	20000	4332	35	20000
Children aged						
0–2 years	0.10	0	1	0.15	0	1
3–4 years	0.08	0	1	0.10	0	1
5–11 years	0.23	0	1	0.23	0	1
Married or couple	0.70	0	1	0.76	0	1
Health	3.65	1	5	3.65	1	5
Public transport quality	0.35	0.15	0.55	0.36	0.15	0.55
Average traffic speed (m.p.h., peak time)	24.53	9.4	39.2	24.12	9.4	39.2
Behaviours						
Not leave television on standby overnight	3.61	1	5	3.56	1	5
Switches off lights in empty rooms	4.35	1	5	4.37	1	5
Not leave tap running when brushing teeth	3.26	1	5	3.25	1	5
Wears extra clothes rather than turn heating up	3.37	1	5	3.29	1	5
Does not buy goods with excess packaging	1.81	1	5	1.65	1	5
Has solar or wind energy or green tariff	1.18	1	3	1.18	1	3
Buys recycled products	2.54	1	5	2.45	1	5
Takes own bags shopping	3.75	1	5	3.15	1	5
Separates rubbish for recycling	3.61	1	4	3.60	1	4
Takes fewer flights where possible	1.90	1	5	1.89	1	5
Attitudes						
Leads an environmentally friendly life	2.62	1	5	2.54	1	5
Being green is 'alternative lifestyle'	2.55	1	4	2.48	1	4
Own behaviour contributes to climate change	0.54	0	1	0.49	0	1
Prepared to pay more for environmentally friendly product	0.43	0	1	0.40	0	1
World on course for major environmental disaster	0.52	0	1	0.46	0	1
Environmental crisis has been exaggerated	0.54	0	1	0.45	0	1
Climate change is beyond our control	0.70	0	1	0.65	0	1
Effects of climate change are too far in the future	0.68	0	1	0.62	0	1
Changes made have to fit in with current lifestyle	0.31	0	1	0.33	0	1
Not worth doing anything unless others do same	0.66	0	1	0.56	0	1
Not worth UK trying to do anything	0.67	0	1	0.58	0	1
Climate change will affect UK in next 30 years	0.64	0	1	0.61	0	1

†One-way regular commute time truncated at 120 min omitting 54/13139 observations (0.4%).
‡The monthly household income figure of £20000 is the recorded maximum in the data. However, this is truncated by the data owners to preserve the anonymity of respondents. There are 111/13139 observations (0.8%) where monthly household income is set to £20000.

everything that they do. The majority of men and women think that climate change is beyond our control, and its effects are too far in the future to worry about. However, over 60% believe that climate change will affect the UK in the next 30 years. Generally environmental attitudes show a large degree of confusion and inconsistency, which accords with the findings of the review work that was carried out for the Department for Transport (Anable *et al.*, 2006); it also contributes to the observed inconsistencies between attitudes and behaviour.

We have carried out both EFA and CFA on our 22 observed indicator variables to explore and verify the latent structure of the data. In the EFA the indicators are allowed to load freely, and the appropriate number of factors is chosen by looking at several diagnostic statistics, including the eigenvalues for each factor, scree plots (Cattell, 1966) and χ^2-tests. The eigenvalue for a given factor reflects the variance in all the variables, which is accounted for by that factor. The Kaiser–Guttman criteria (Kaiser, 1960; Guttman, 1954) recommend retaining factors with an eigenvalue greater than 1. For both men and women, all these statistics suggest that two or three factors are superior to a one-factor model. In χ^2-tests the null hypothesis of one factor is rejected is favour of the alternative hypothesis of two factors, and a null of two factors is rejected in favour of three. Comparing the factor loadings in the two- and three-factor models, the two-factor model has a 'cleaner' structure where both factors have very distinct loadings, with all except one of the attitude indicators loading on the first factor and all the behaviour indicators loading on the second factor. This can be justified on the basis of psychological theory that treats attitudes and behaviours as separate constructs both generally (Ajzen and Fishbein, 2005) and in the context of environmental decision making (Stern, 2000). Thus, taking account of all this information, a two-factor model is preferred.

Given that EFA suggests that it is reasonable to view our indicators for attitudes and behaviours as two separate constructs, CFA is then employed as the first step in the estimation of the measurement model in the SEM and here the indicators are forced to load onto their two respective factors. The results are shown in Table 4; the second and third columns show the factor loadings for environmental behaviours and the last two columns for attitudes (these are

Table 4. CFA factor loadings†

Variable	Environmental behaviours		Variable	Environmental attitudes	
	Women	*Men*		*Women*	*Men*
Packing	2.781	2.080	Britain	4.529	4.555
Produce	2.522	2.111	Others	4.384	4.260
Flights	2.306	1.615	Future	4.370	4.404
Bags	1.975	1.927	Control	4.061	3.879
Heating	1.748	1.160	Exaggerate	3.823	3.854
Tap	1.731	1.458	30 years	3.803	4.011
Lights	1.550	1.125	Disaster	3.600	3.730
Energy	1.390	1.172	Pay more	2.844	3.162
Recycle	1.284	0.875	Own responsibility	2.778	2.993
TV‡	1.000	1.000	Lifestyle	2.538	2.764
			Alternative	1.444	1.409
			Own life‡	1.000	1.000
n	6883	6256		6883	6256

†Factor loadings are the α-coefficients from equation (4). All estimates are significant at $p < 0.0001$.
‡Loading fixed at 1 via normalization.

the α-coefficients from equation (4)); in both cases the results are presented in descending order of the factor loadings for women. The factor loadings show how each indicator is associated with the underlying latent construct. For behaviours the indicators are normalized so that the loading on 'not leaving the television on standby overnight', TV, is set to 1. For women all other indicators have a higher loading than TV; 'not buying goods with excess packaging', packing, has the largest loading onto the behaviours factor, followed by 'buying recycled products', produce, and 'taking fewer flights where possible', flights; the lowest loadings are for 'having green energy or tariff', energy, 'separating rubbish for recycling', recycling. The ranking of loadings onto behaviour is very similar for men. For attitudes the loading on belief that they 'lead an environmentally friendly life', own life, is set to 1 to normalize the scale. For both women and men the largest loadings are for 'it's not worth Britain trying to do anything about climate change', Britain, 'it's not worth doing anything unless others do the same', others, and 'the effects of climate change are too far in the future', future, and the lowest loadings are for own life, 'being green is an alternative lifestyle', alternative, and 'any changes I make have to fit in with my current lifestyle', lifestyle.

The results for the latent variable model are presented in Table 5; these show the associations between the two latent constructs and observable individual characteristics (the γ-coefficients from equation (2)). As is common in the SEM literature, standardized coefficients are reported for continuous variables, as these allow a comparison of the relative size of the effects within models. The standardized coefficients are $\beta^* = \beta\sigma_x/\sigma_q$, where σ_x and σ_q are the standard deviations of the continuous explanatory variable x and dependent variable q. Non-standardized

Table 5. Latent variable model results†

Latent dependent variable	Environmental behaviours		Environmental attitudes	
	Women	*Men*	*Women*	*Men*
Age	0.683‡	0.517‡	0.248‡	0.295‡
Age squared	−0.448‡	−0.312§	−0.238‡	−0.236‡
Children aged				
0–2 years	−0.008	0.023§§	−0.014	0.005
3–4 years	0.011	−0.004	−0.016	0.006
5–11 years	0.001	−0.005	−0.023‡	−0.015§
12–15 years	−0.003	0.001	−0.007	−0.005
Education				
Degree	0.121‡	0.157‡	0.127‡	0.127‡
A level	0.053‡	0.079‡	0.090‡	0.097‡
O level	0.021	0.023	0.058‡	0.068‡
Household income	−0.032§	−0.058‡	0.097‡	0.099‡
Married	0.031‡	0.064‡	0.003	0.010
Health status	0.009‡	0.018‡	−0.001	−0.001
n	6883	6256	6883	6256

†These are the γ-coefficients from equation (2). Standardized coefficients ($\gamma^* = \gamma\sigma_x/\sigma_q$) are reported for continuous variables and non-standardized for dichotomous and ordinal variables. Standard errors for the standardized coefficients are obtained by the delta method (see Davidson and MacKinnon (2004), section 5.6).
‡Significance at $p < 0.001$.
§Significance at $p < 0.05$.
§§Significance at $p < 0.1$.

coefficients are reported for dichotomous and ordinal variables, and these show the estimated change in the dependent variable for a discrete unit change in the explanatory variable. *p*-values for coefficient estimates are calculated under the assumption of asymptotic normality, where the asymptotic properties of the WLS estimator are discussed in Muthén and Satorra (1995). For both men and women proenvironmental behaviours and attitudes are non-linearly related to age. Having children does not seem to matter for behaviours for men or women; but for both men and women it seems that having primary-school-age children means that it is less likely that you will have proenvironmental attitudes. Behaviours and attitudes are increasing in education for both men and women. Income has a negative association with proenvironmental behaviours for men and women, but it has a positive association with proenvironmental attitudes. Being married and having better health have a positive association with behaviours for both men and women but have no effect on attitudes.

Table 6 shows the results for the mode choice model where the dependent variable is a dichotomous choice between commuting by car ($d_i = 1$) and by public transport ($d_i = 0$) (Table 6 reports the β-coefficients from equation (1)). Columns 1(a) and 1(b) present the results from a standard DCM, where the latent variables for greenness have been omitted. Columns 2(a) and 2(b) include the two latent variables, and columns 3(a) and 3(b) also include commuting time as an additional regressor; again standardized coefficient estimates are reported for continuous variables. Firstly, we see that omitting the latent variables makes virtually no difference to the coefficient estimates on the variables included. However, when included the behaviours latent variable is significant for both men and women and the attitudes latent variable is significant for men. In addition a Wald test for joint significance of the two latent variables shows them to be jointly significant ($p = 0.000$) in all models in which they are included. This suggests little collinearity between the latent variables and the other explanatory variables and that the former are independently important in explaining mode choice. The pseudo-R^2 statistic (McKelvey and Zavoina, 1975) and predictive ability as shown by the proportion of correctly predicted cases also suggest the superiority of models including the latent variables.

In general the results are very similar whether or not commute time is included (models 2(a) and 2(b) *versus* 3(a) and 3(b)). In terms of the latent variables, having a latent tendency to proenvironmental behaviours in other areas of life has a negative effect on the probability of commuting by car. Similarly, latent proenvironmental attitudes have a negative effect on the probability of car commutes for men but no significant effect for women. In terms of the conditioning variables, there is a similar non-linear age effect for men and women. Having preschool children means that women are more likely to commute by car, but the effect is not significant for men. The probability of commuting by car is increasing in education for women but this is not significant for men. Household income has a negative effect for men. Married women are less likely to commute by car but married men are more likely to. Health has no effect; however, it is worth stressing that this is a relatively healthy sample because by definition all respondents are working. The quality of public transport in the local area and the average traffic speed have the expected signs and are significant for both men and women; the better the public transport the less likely people are to commute by car and the higher the average traffic speed during the rush hour (i.e. the less congestion) the more likely.

The quantitative interpretation for the standardized coefficients in Table 6 is that, for any coefficient estimate $\hat{\beta}$, a 1-standard-deviation change in the associated continuous explanatory variable results in a $\hat{\beta}$-standard-deviations change in the underlying latent dependent variable (the utility derived from choosing to commute by car). Hence, the standardized coefficients on the continuous latent explanatory variables can be compared straightforwardly with those for other continuous variables. Here we see that for men the effects of environmental attitudes and

Table 6. Mode choice model results†

Variable	*Results for the following models:*					
	1(a), women	*1(b), men*	*2(a), women*	*2(b), men*	*3(a), women*	*3(b), men*
Environmental behaviours			−0.170‡	−0.182‡	−0.167‡	−0.181‡
Environmental attitudes			−0.004	−0.109‡	0.001	−0.105‡
Age	0.578‡	0.509‡	0.695‡	0.635‡	0.778‡	0.701‡
Age squared	−0.463‡	−0.418‡	−0.540‡	−0.501‡	−0.636‡	−0.563‡
Children						
0–2 years	0.163‡	0.051	0.156‡	0.070	0.152‡	0.071
3–4 years	0.176‡	0.089	0.185‡	0.092	0.211‡	0.094
5–11 years	0.053	−0.059	0.053	−0.071	0.013	−0.060
12–15 years	0.040	0.021	0.037	0.019	0.009	−0.006
Education						
Degree	0.268‡	−0.268‡	0.372‡	−0.083	0.459‡	−0.033
A level	0.324‡	−0.178‡	0.369‡	−0.067	0.408‡	−0.036
O level	0.234‡	−0.059	0.253‡	−0.002	0.265‡	0.018
Household income	0.029§	−0.112‡	0.024	−0.112‡	0.054‡	−0.088‡
Married	−0.190‡	0.149‡	−0.164‡	0.200‡	−0.199‡	0.195‡
Health	−0.025	−0.008	−0.017	0.005	−0.014	0.008
Public transport quality	−0.157‡	−0.180‡	−0.157‡	−0.180‡	−0.126‡	−0.160‡
Average traffic speed	0.168‡	0.106‡	0.168‡	0.106‡	0.147‡	0.100‡
Commute time					−0.233‡	−0.166‡
n	6883	6256	6883	6256	6883	6256
SEM diagnostic statistics						
Pseudo-R^2	0.106	0.124	0.132	0.164	0.177	0.185
Correct predictions	71.74	74.57	72.03	75.03	76.12	76.12
Chi-sq	295.5‡	218.2‡	59934.0‡	49040.8‡	60266.4‡	49211.9‡
CFI	—§§	—§§	0.897	0.906	0.899	0.907
TLI	—§§	—§§	0.885	0.895	0.887	0.896
RMSEA	—§§	—§§	0.042	0.038	0.041	0.037
90% confidence interval for RMSEA	—§§	—§§	(0.041, 0.042)	(0.037, 0.039)	(0.040, 0.041)	(0.036, 0.038)

†Dependent variable: car, 1; public transport, 0. These are the β-coefficients from equation (1). Standardized coefficients ($\beta^* = \beta\sigma_x/\sigma_y$) are reported for continuous variables and non-standardized for dichotomous and ordinal variables. Standard errors for the standardized coefficients are obtained by the delta method (see Davidson and MacKinnon (2004), section 5.6). Pseudo-R^2 is McKelvey and Zavoina's (1975). The correct predictions variable is the proportion of correctly predicted cases. Chi-sq—H_0: all slope parameters in the structural part of the model are 0, and the factor loadings in the measurement part of the model are all 1. CFI, comparative fit index. TLI, Tucker–Lewis reliability index; RMSEA, root-mean-square error of approximation.
‡Significance at $p < 0.001$.
§Significance at $p < 0.1$.
§§Not applicable.

behaviours are very similar in size to the effects of local public transport quality and average traffic speeds. For example, increasing proenvironmental attitudes by 1 standard deviation reduces the utility that is derived from car use by 0.109 standard deviations; this is almost identical to the increase in utility that arises from a 1-standard-deviation increase in local traffic speeds. Similarly for women, having a latent tendency to undertaking environmental behaviours in other areas of life has a similar effect on reducing the utility from car use to that from having better local public transport, or more road congestion. It is worth pointing out here that the relative sizes of the standardized coefficients for the latent *versus* transport variables could be affected by possible errors-in-variables problems for the latter, due to the use of local authority areawide

variables to represent the effects for individuals. This would probably bias the coefficients of the area level transport variables downwards, which may mean that the latent variables have a smaller relative effect. However, it is also possible that the exclusion of individual travel cost variables could cause the travel time variables to be biased in the opposite direction.

In the second pair of models where the respondents' usual commute time is included, this has a negative effect as expected, i.e. the longer your commute the less likely you are to use a car. Including commuting time means that household income is now significant and positive for women (it remains negative for men). This is unsurprising because there is a close positive correlation between household income and commute time for men in particular; this correlation is because the rational decision maker will choose to commute for longer only if they are compensated, and part of this compensation comes from the labour market in the form of higher wages. However, it is also likely that commute time is endogenous in this model, not least because there is a two-way relationship between length of commute and mode, and also because there may be a set of unobserved factors which influence both mode and time. One such factor is 'trip chaining', which arises where individuals make multiple stops on their commute, e.g. to take children to school or to pick up shopping; this information is not available in our data. Nevertheless the fact that inclusion of commute time does not substantively change our estimates of the relative importance of environmental behaviours and attitudes is a strong robustness check on our results.

Model fit statistics for the SEM are reported in the lower part of Table 6, and these are all supportive of our model specification. We can reject the null hypothesis of the χ^2-test that all slope parameters in the structural part of the model are 0, and the factor loadings in the measurement part of the model are all 1. The CFIs are all above (or very close to) the recommended cut-off of 0.9; similarly the TLIs are all very close to (but just below) 0.9. In addition the RMSEAs for all four models are below 0.06.

For conciseness we do not report the results of estimating these models with our alternative data set (the BHPS) here. In summary the story is essentially the same, although the smaller sample sizes result in larger standard errors. The factor analysis suggests two latent factors, and both of these (environmental attitudes and behaviours) are significant in determining mode choice for both men and women; this is slightly different from the UKHLS results, where only behaviours are significant for mode choice for women. As for the UKHLS, the BHPS estimates suggest that proenvironmental attitudes and behaviours reduce the probability of commuting by car; quantitatively these effects are larger in the BHPS data, and the effects of quality of local public transport and congestion are smaller. This replication is an important check on the robustness of our estimates.

6. Discussion

Some important findings emerge from the estimation of our HCM for commuting mode choice. Firstly, we have shown that it is possible to use large secondary data sets for HCM estimation. This increases the generalizability and statistical reliability of our results compared with existing studies that rely on relatively small bespoke surveys, which are prone to selection problems and focusing effects, and include little information on individual characteristics with which to control for heterogeneity. Secondly, the factor analysis suggests that the indicator variables are representative of two latent constructs: environmental attitudes and behaviours. These attitudes and behaviours appear to be separable constructs and the latent variable model shows, for example, that whereas higher levels of education are associated with both more proenvironmental attitudes and behaviours, in contrast increased income has a positive association with attitudes but a

negative association with behaviours. These results, in the context of the psychological ABC model of environmental decision making, suggest different antecedents for attitudes and behaviours. Behaviours are much more likely to be influenced by personal context and convenience than attitudes; this may explain the diverse income effects and also the fact that marital status and health are significant predictors of environmental behaviours but do not affect attitudes.

Thirdly, environmental attitudes and behaviours are significantly related to the choice of commuting mode. For men, the more proenvironmental their attitudes and other lifestyle behaviours the less likely they are to use a car for the regular commuting journey. This result contrasts with the previous literature that has argued that attitudes will have little effect on high constraint environmental behaviours like car driving (Collins and Chambers, 2005). We cannot completely discount the possibility that mode choice behaviour is driving attitudes here, in that men who use public transport for their daily commute see themselves as environmentally sympathic and hence change their attitudes to align with this. However, the nature of our household survey data and the fact that the environmental questions are not directly related to commuting, or asked in the same survey module, reduce this possibility compared with the bespoke survey data that are normally used to estimate HCMs. For women other environmental behaviours are again significant, but in contrast attitudes have no significant effect. This may be because women's commuting choices are more constrained than men's, as evidenced by the fact that having preschool-age children significantly increases the probability that women will use a car for commutes but this is not significant for men. Previous literature has shown that women have more complex journeys to work than men and are engaged in more trip chaining resulting in non-direct home-to-work journeys (Hensher and Reyes, 2000).

Finally, our results are supportive of the ABC model of environmental decision making. Attitudes and behaviours influence the utility that an individual derives from different mode choices for the regular commute. Thus the commuting mode choice is an interactive product of 'internal' attitudes and 'external' contextual factors.

7. Conclusion

Persuading people to ignore their cars and to use alternative, more sustainable forms of travel is essential if governments are to achieve their ambitious climate change goals. However, in the UK, as in many other countries around the world, our attachment to the car persists. This paper has contributed to furthering our understanding of the way that people make travel choices, specifically what determines choice of mode for the regular commuting journey. Traditionally transport economics have focused on time and cost, assuming that these are the main determinants of travel choices for the rational economic agent. HCMs have allowed us to integrate latent variables, reflecting underlying environmental attitudes and behaviours, into a model of mode choice; these variables are shown to be significant and their effects are similar in size to important contextual factors like the availability of public transport.

Integrating these latent variables into the mode choice model has facilitated a more sophisticated understanding of the decision-making process. This is reflective of a more general acceptance of behavioural economics, which diverges from the narrow view of economic rationality, and incorporates psychological factors into models of individual decision making. Unusually, for applications of HCMs, we have used large nationally representative household survey data sets for model estimation, thereby increasing the generalizability and statistical reliability of our results.

The fact that psychological factors influence commuting mode choices can be exploited by policy makers who need to persuade us to make more environmentally sympathic choices.

Attempting to influence our attitudes towards the environment (e.g. via advertising campaigns or provision of information) or our other environmental behaviours (e.g. by making recycling a convenient activity for households) are not substitutes for fiscal tools and regulation but they can be seen as part of a comprehensive policy toolbox, which is targeted at making our travel choices more sustainable. A similar toolbox has been used successfully in the UK, and other countries, to reduce smoking behaviour substantially (Bauld, 2011). As well as climate change, private car use also contributes to congestion, noise, poor air quality, road traffic accidents and low levels of physical activity; so there are many reasons to try to bring about a change in individual behaviours.

Acknowledgements

References

Ajzen, I. (1991) The theory of planned behavior. *Organ. Behav. Hum. Decsn Process*, **50**, 179–211.
Ajzen, I. and Fishbein, M. (1970) The prediction of behavior from attitudinal and normative variables. *J. Exptl Socl Psychol.*, **6**, 466–487.
Ajzen, I. and Fishbein, M. (1977) Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychol. Bull.*, **84**, 888–918.
Ajzen, I. and Fishbein, M. (2005) The influence of attitudes on behavior. In *The Handbook of Attitudes* (eds D. Albarracin, B. T. Johnson and M. P. Zanna), pp. 173–221. New York: Psychology Press.
Anable, J., Lane, B. and Kelay, T. (2006) An evidence base review of public attitudes to climate change and transport behaviour. *Report*. Department for Transport, London.
Ashok, K., Dillon, W. and Yuan, S. (2002) Extending discrete choice models to incorporate attitudinal and other latent variables. *J. Markt. Res.*, **39**, 31–46.
Bamberg, S. and Schmidt, P. (2003) Incentives, morality, or habit?: predicting students' car use for university routes with the models of Ajzen, Schwartz, and Triandis. *Environ. Behav.*, **35**, 264–285.
Bauld, L. (2011) The impact of smokefree legislation in England: evidence review. *Report*. Department of Health, London.
Ben-Akiva, M. and Boccara, B. (1987) Integrated framework for travel behavior analysis. *International Association of Travel Behavior Research 5th Int. Conf. Travel Behaviour, Aix-en-Provence.*
Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A. and Munizaga, M. A. (2002) Hybrid choice models: progress and challenges. *Marktng Lett.*, **13**, 163–175.
Bentler, P. (1990) Comparative fit indexes in structural models. *Psychol. Bull.*, **107**, 238–246.
Bollen, K. and Pearl, J. (2013) Eight myths about causality and structural equation models. In *Handbook of Causal Analysis for Social Research* (ed. S. L. Morgan), pp. 301–328. New York: Springer.
Browne, M. (1984) Asymptotically distribution free methods for the analysis of covariance structures. *Br. J. Math. Statist. Psychol.*, **37**, 62–83.
Cattell, R. (1966) The scree test for the number of factors. *Multiv. Behav. Res.*, **1**, 245–276.
Choo, S. and Mokhtarian, P. L. (2004) What type of vehicle do people drive?: the role of attitude and lifestyle in influencing vehicle type choice. *Transprtn Res.* A, **38**, 201–222.
Collins, C. and Chambers, S. (2005) Psychological and situational influences on commuter-transport-mode choice. *Environ. Behav.*, **37**, 640–661.
Córdoba, J. E. and Jaramillo, G. P. (2012) Inclusion of the latent personality variable in multinomial logit models using the 16pf psychometric test. *Procedia*, **54**, 169–178.

Daly, A., Hess, S., Patruni, B., Potoglou, D. and Rohr, C. (2012) Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation*, **39**, 267–297.

Davidson, R. and MacKinnon, J. G. (2004) *Econometric Theory and Methods*. New York: Oxford University Press.

Department for Environment, Food and Rural Affairs (2007) Report, questionnaire and data tables following Survey of Public Attitudes and Behaviours toward the Environment. *UK Data Service Report SN:5741*. Department for Environment, Food and Rural Affairs, London. (Available from `http://dx.doi.org/10.5255/UKDA-SN-5741-1`.)

Department for Transport (2008) Annual report 2008. *Cm 7395*. Norwich: Stationery Office.

Department for Transport (2013) National Travel Survey 2012. Department for Transport, London. (Available from `www.gov.uk/government/statistics/national-travel-survey-2012`.)

Dunlap, R. E., Liere, K. D., Van Mertig, A. G. and Jones, R. E. (2000) Measuring endorsement of the New Ecological Paradigm: a Revised NEP scale. *J. Socl Iss.*, **56**, 425–442.

Festinger, L. (1962) Cognitive dissonance. *Scient. Am.*, **207**, 93–107.

Gardner, B. and Abraham, C. (2008) Psychological correlates of car use: a meta-analysis. *Transprtn Res.* F, **11**, 300–311.

Gardner, B. and Abraham, C. (2010) Going green?: modeling the impact of environmental concerns and perceptions of transportation alternatives on decisions to drive. *J. Appl. Socl Psychol.*, **40**, 831–849.

Gifford, R. (2011) The dragons of inaction: psychological barriers that limit climate change mitigation and adaptation. *Am. Psychol.*, **66**, 290–302.

Golob, T. F. (2003) Structural equation modeling for travel behavior research. *Transprtn Res.* B, **37**, 1–25.

Graham-Rowe, E., Skippon, S., Gardner, B. and Abraham, C. (2011) Can we reduce car use and, if so, how?: a review of available evidence. *Transprtn Res.* A, **45**, 401–418.

Guagnano, G., Stern, P. and Dietz, T. (1995) Influences on attitude-behavior relationships a natural experiment with curbside recycling. *Environ. Behav.*, **27**, 699–718.

Guttman, L. (1954) Some necessary conditions for common-factor analysis. *Psychometrika*, **19**, 149–161.

Hensher, D. and Reyes, A. (2000) Trip chaining as a barrier to the propensity to use public transport. *Transportation*, **27**, 341–361.

Hu, L. and Bentler, P. (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equn Modlng*, **6**, 1–55

Johansson, M. V., Heldt, T. and Johansson, P. (2006) The effects of attitudes and personality traits on mode choice. *Transprtn Res.* A, **40**, 507–525.

Jöreskog, K. (1969) A general approach to confirmatory factor analysis. *Psychometrika*, **34**, 183–202.

Kaiser, H. (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187–200.

Kaiser, H. (1960) The application of electronic computers to factor analysis. *Educ. Psychol. Measmnt*, **20**, 151.

King, S., Dyball, M. and Webster, T. (2009) Exploring public attitudes to climate change and travel choices. *Final Report*. Department for Transport, London.

Kline, S. (1988) Rationalising attitude discrepant behaviour: a case study in energy attitudes. *Manuscript*. Faculty of Environmental Studies, York University, Toronto, Unpublished.

Kraus, S. J. (1995) Attitudes and the prediction of behavior: a meta-analysis of the empirical literature. *Person. Socl Psychol. Bull.*, **21**, 58–75.

Longhi, S. (2013) Individual pro-environmental behaviour in the household context. *Working Paper 2012-21*. Institute for Social and Economic Research, University of Essex, Colchester.

Lynn, P. and Longhi, S. (2011) Environmental attitudes and behaviour: who cares about climate change? In *Understanding Society: Early Findings from the First Wave of the UK's Household Longitudinal Study* (eds S. L. McFall and C. Garrington). Colchester: Institute for Social and Economic Research.

MacCallum, R., Browne, M. and Sugawara, H. (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychol. Meth.*, **1**, 130–149.

MacPherson, R. and Lange, I. (2013) Determinants of green electricity tariff uptake in the UK. *En. Poly*, **62**, 920–933.

McFadden, D. (1986) The choice theory approach to market research. *Marktng Sci.*, **5**, 275–297.

McKelvey, R. and Zavoina, W. (1975) A statistical model for the analysis of ordinal level dependent variables. *J. Math. Soc.*, **4**, 103–120.

Morikawa, T., Ben-Akiva, M. and McFadden, D. (2002) Discrete choice models incorporating revealed preferences and psychometric data. *Econmetr. Models Marktng*, **16**, 29–55.

Muthén, B. (1983) Latent variable structural equation modelling with categorical data. *J. Econmetr.*, **22**, 48–65.

Muthén, B. (1984) A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115–132.

Muthén, B. and Muthén, L. K. (2016) Technical appendices, July. (Available from `www.statmodel.com/download/techappen.pdf`.)

Muthén, B. and Satorra, A. (1995) Technical aspects of Muthen's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika*, **60**, 489–503.

Oskamp, S., Harrington, M. J., Edwards, T. C., Sherwood, D. L., Okuda, S. M. and Swanson, D. C. (1991) Factors influencing household recycling behavior. *Environ. Behav.*, **23**, 494–519.

Roberts, J., Hodgson, R. and Dolan, P. (2011) 'It's driving her mad': gender differences in the effects of commuting on psychological health. *J. Hlth Econ.*, **30**, 1064–1076.

Sánchez, B. N., Budtz-Jørgensen, E., Ryan, L. M. and Hu, H. (2005) Structural Equation Models. *J. Am. Statist. Ass.*, **100**, 1443–1455.

Schwartz, S. (1977) Normative influences on altruism. *Adv. Exptl Socl Psychol.*, **10**, 221–279.

StataCorp. (2013) *Stata Statistical Software: Release 13*. College Station: StataCorp.

Steiger, J. and Lind, J. (1980) Statistically based tests for the number of common factors. *A. Spring Meet. Psychometric Society, Iowa City*. (Available from `http://www.statpower.net/Steiger%20Biblio/Steiger-Lind%201980.pdf`.)

Stern, P. (2000) Toward a coherent theory of environmentally significant behaviour. *J. Socl Iss.*, **56**, 407–424.

Stern, P., Dietz, T. and Black, J. (1985) Support for environmental protection: the role of moral norms. *Popln Environ.*, **8**, 204–222.

Stern, P. and Oskamp, S. (1987) Managing scarce environmental resources. In *Handbook of Environmental Psychology* (eds I. Altman and D. Stokels), pp. 1044–1088. New York: Wiley.

Temme, D., Paulssen, M. and Dannewald, T. (2008) Incorporating latent variables into discrete choice models—a simultaneous estimation approach using SEM software. *Bus. Res.*, **1**, 220–237.

Thomas, G. O. Poortinga, W. and Sautkina, E. (2016) The Welsh single-use carrier bag charge and behavioural spillover. *J. Environ. Psychol.*, **47**, 126–135.

Thornton, A., Bunt, K., Dalziel, D. and Simon, A. (2010) Climate change and transport choices: segmentation study. *Interim Report by TNS-BMRB*. Department for Transport, London.

Train, K. (1980) A structured logit model of auto ownership and mode choice. *Rev. Econ. Stud.*, **47**, 357–370.

Triandis, H. (1977) *Interpersonal Behaviour*. Monterey: Brook–Cole.

Tucker, L. and Lewis, C. (1973) A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**, 1–10.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: heuristics and biases. *Science*, **185**, 1124–1131.

University of Essex (2012) Understanding society: Wave 1-3, 2009-2012. Institute for Social and Economic Research and National Centre for Social Research, University of Essex, Colchester.

Yáñez, M. F., Raveau, S. and Ortúzar, J. D. D. (2010) Inclusion of latent variables in Mixed Logit models: Modelling and forecasting. *Transprtn Res.* A, **44**, 744–753.