

Identifying variations and co-variations in discrete choice models

Makoto Chikaraishi · Akimasa Fujiwara · Junyi Zhang ·
Kay W. Axhausen

Published online: 9 January 2011
© Springer Science+Business Media, LLC. 2011

Abstract This study introduces an extended version of a standard multilevel cross-classified logit model which takes co-variations into account, i.e., variations jointly caused by two or more unobserved factors. Whilst focusing on mode choice behavior, this study deals with four different types of variation: spatial variations, inter-individual variations, intra-individual variations and co-variations between inter-individual and spatial variations. Such co-variations represent individual-specific spatial effects, reflecting different responses to the same space among individuals, which may for example be due to differences in their spatial perceptions. In our empirical analysis, we use data from *Mobidrive* (a continuous six-week travel survey) to clarify the existence of co-variation effects by comparing two models with and without co-variation terms. The results of this analysis indicate that co-variations certainly exist, especially for utility differences in bicycle and public transport use in comparison with car use. We then sequentially introduce four further sets of explanatory variables, examine the sources of behavioral variations and determine what types of influential factors are dominant in mode choice behavior.

Keywords Behavioral variations and co-variations · Multilevel cross-classified logit model · Individual-specific spatial effects · *Mobidrive* data

M. Chikaraishi (✉) · A. Fujiwara · J. Zhang
Graduate School for International Development and Cooperation, Hiroshima University,
1-5-1, Kagamiyama, Higashi-Hiroshima 739-8529, Japan
e-mail: chikaraishi@hiroshima-u.ac.jp

A. Fujiwara
e-mail: afujiw@hiroshima-u.ac.jp

J. Zhang
e-mail: zjy@hiroshima-u.ac.jp

K. W. Axhausen
Swiss Federal Institute of Technology, Zürich, Switzerland
e-mail: axhausen@ivt.baug.ethz.ch

Introduction

Travel behavior varies due to various factors, some of which are unobservable. For example, it is usually difficult to observe psychological factors such as motives, attitudes and perceptions, which are particularly influential to the decision-making process, even for variations caused by widely addressed individual attributes. Accordingly, it becomes important to explore how many variations due to individual attributes can be observed and how many can not be observed, from both data collection and modeling perspectives. Such a discussion can also be applied to other types of variations, including household attributes, spatial attributes and temporal factors.

In fact, a number of studies have already focused on different kinds of variations, and the importance of distinguishing between these variations has been argued (Kitamura 2003). In particular, the distinction between inter-individual variation and intra-individual variation has been intensively examined (e.g., Pas 1987; Hanson and Huff 1988; Kitamura et al. 2006). In addition to inter-individual and intra-individual variations, Pas and Sundar (1995) and Goulias (2002) empirically showed the existence of inter-household variations, which account for around 1/3 of all inter-individual variations in trip generation and time use. Here, the three types of variation (intra-individual, inter-individual and inter-household) hold a nesting hierarchical structure. Concretely speaking, an observed behavior in a certain situation (intra-individual) is nested within an individual (inter-individual), and an individual is nested within a household (inter-household). Since a model usually tries to structuralize variations in data, such a hierarchical point of view is quite important to identify whether the variations come from situational differences (i.e., intra-individual variations) or from individual differences (i.e., inter-individual variations). To capture such different levels of variations simultaneously, a multilevel modeling approach (Goldstein 2003) has become more and more popular, and several applications can be found in the field of transportation (e.g., Bhat and Zhao 2002; Goulias 2002; Weber and Kwan 2003; Kim et al. 2004; Khandker et al. 2006; Habib and Miller 2008; Susilo and Dijkstra 2009). A random coefficient model with panel data can also be regarded as a special kind of multilevel model in which intra-individual and inter-individual variations are usually distinguished. Cirillo and Axhausen (2006) distinguished between these two variations in a mode choice context, and Cherchi and Cirillo (2008) extended it to different hierarchical levels, including inter-tour, inter-day, inter-day of week, inter-individual and inter-household variations.

At the same time, there are many cases when a certain level unit can be nested within more than one higher-level unit, i.e., some variation types do not have a hierarchical structure. For example, Bhat (2000) focused on work-travel mode choice and the fact that the choice is nested not only within the home location (i.e., the origin), but also within the work location (i.e., the destination). Obviously, home and work locations do not have a hierarchical structure. Here we call such a non-hierarchical structure a cross-classified structure. In other examples, spatial variations and systematic day-to-day variations (i.e., day effects at an aggregate level) may also have a cross-classified structure. An individual can move in and out of different origin and destination pairs, while an origin and destination pair can be moved by different individuals; the relationship between individuals and space therefore also follows a cross-classified structure. Such complicated cross-classified variation structures have been examined simultaneously by Chikaraishi et al. (2009, 2010), with a focus on departure time choice and time use behavior. However, in the standard multilevel cross-classified model, we have to assume that all variations at the different cross-classified levels are independent of each other. When this assumption is not

true, co-variations, which are variations jointly caused by two or more unobserved factors, would occur. Note that co-variations only occur in cross-classified structures (i.e., not in hierarchical structures). For example, travel choice behavior might depend not only on a person's individual attributes (such as gender, age and work style), but also on the origin and destination (with respect to travel time and travel cost), and these two effects might be correlated in some cases. For example, responses to travel time and travel cost may differ across individuals. Variations caused by interactions between different cross-classified levels are called co-variations. Co-variations can neither be captured by purely spatial variations, nor by purely inter-individual variations, but they may correspond to individual-specific spatial effects.

The importance of understanding the above-mentioned co-variations might be analogous to the importance of understanding individual-specific action (activity) space. In fact, recent studies have emphasized that action space should be described for each individual by considering biographical elements in the mental map of each individual, rather than assuming that all individuals have the same perceptions of space (Schönfelder and Axhausen 2003). Differences between social networks would lead to different perceptions of space among individuals. For example, those individuals who have babies might choose supermarkets where they can buy baby-related products. Thus, the degree of attractiveness of spaces would differ across individuals for various reasons, and co-variations might be the terms which could capture such individual-specific spatial effects by taking unobserved effects into account. Although understanding co-variations could provide much deeper insights for policy design, they have not yet been well examined.

On the basis of the above-mentioned considerations, this study proposes a model to detect co-variations as well as other types of variations. The model described in this study is a straightforward extension of the multilevel cross-classified logit (MCL) model (Bhat 2000). The MCL model is similar to the mixed logit (MXL) model with panel data, the difference being that the MCL model can deal with several hierarchical or cross-classified random variables to capture various unobserved variations. In this study, we additionally introduce interaction term(s) related to different pairs of cross-classified levels to capture co-variations. The result is a multilevel cross-classified logit model with co-variations (MCLCV). The advantage of the MCLCV model is that it can detect the influence of interaction(s) between different cross-classified levels before introducing explanatory variables. In other words, this allows us to flexibly decompose behavioral variations (variations of utility) into the variation types of interesting aspects, including co-variations. Such decomposition could be used to figure out where the variations of utility come from. By applying the MCLCV model with explanatory variables, it is possible to identify which explanatory variable contributes to the reduction of which parts of the variations. In other words, it is possible to examine whether the introduced explanatory variables suffice to describe the behavior or not.

After describing the MCLCV model, we shall provide an empirical analysis with respect to travel mode choice, which has been widely addressed in the literature. The total of all behavioral variations will be classified into intra-individual variations, inter-individual variations, spatial variations, and co-variations between inter-individual and spatial variations. In the empirical analysis, spatial variations are defined as variations caused by differences between origin–destination (O–D) pairs and residential locations. Accordingly, the relationship between intra-individual and inter-individual/spatial variations has a hierarchical structure (i.e., a mode choice decision is exactly nested within a certain individual and space), while the relationship between inter-individual and spatial variations has a cross-classified structure (i.e., the same individual could travel between different

O–D pairs). We first compare model estimation results with and without the co-variation terms and then introduce a number of explanatory variables to capture the variations. For the empirical analysis we used data from *Mobidrive*, a continuous six-week travel survey, collected in Karlsruhe and Halle in Germany in 1999 (Axhausen et al. 2002). These data allow us to distinguish between the above-mentioned different types of variations.

This paper is organized as follows: In the next section, we shall describe the MCLCV model and explain how to determine its variation properties. We shall then summarize the data used in the empirical analysis, focusing on mode choice behavior. After that, we shall compare the estimation results of models with and without co-variation terms and evaluate how many variations can be captured by introducing a number of explanatory variables. In the final section, some major findings and avenues of future research will be summarized.

Methodology

This section describes a method for capturing various kinds of variations, including co-variations, by developing a multilevel cross-classified logit model with co-variations (called the MCLCV model). We shall describe the model with regard to both random (unobserved) and non-random (observed) effects of spatial, inter-individual, and intra-individual variations as well as co-variations between spatial and inter-individual variations (see Fig. 1). As for spatial variations, it can be defined in various ways, including home locations, origins, destinations, and/or origin–destination pairs. The choice would depend on what we want to measure. In fact, spatial variations are one of the most traditional variation types that have been widely addressed, because it can directly connect with policy discussions especially in the context of infrastructure construction. For example, it has been recognized that the accessibility of residential locations has certain impacts on mode choice decisions, and it might be better to employ home location as a unit of spatial variation if our purpose is to identify such home-based accessibility impacts. The travel time and travel cost have also been recognized as important influential factors on the mode choice decisions, and, in this case, employing origin–destination pair as a unit of spatial variations would be preferable. On the other hand, rather than spatial factors, a number of researchers have pointed out that individual socio-demographic factors have great impacts on the decisions of mode choice (Kitamura et al. 1997; Susilo and Maat 2007; Pinjari et al. 2007). Such impacts would be reflected in inter-individual variations,

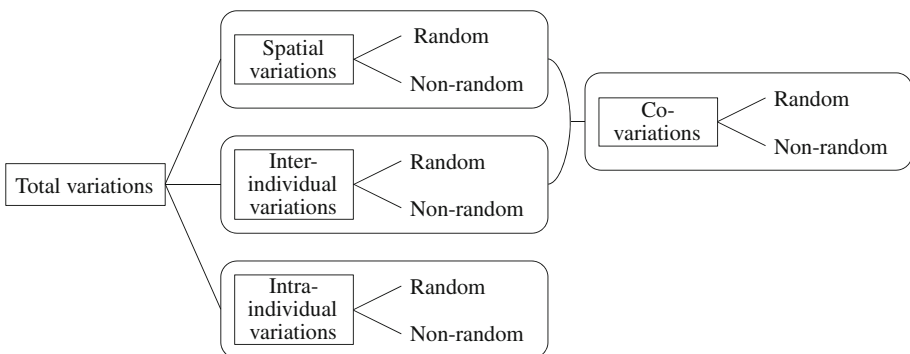


Fig. 1 The variation structure assumed in this study

including car ownership, household structure, psychological attitude, and so on. Additionally, it can also be expected that the situational attributes such as traveling with friends, time pressure, etc., which can easily change even within an individual, may be important influential factors on mode choice decisions as mentioned in Introduction. Such impacts would be reflected in intra-individual variations. Note that, since it is usually difficult to capture such situational attributes, it could be expected that many of them would remain as unobserved variations. Finally, co-variations can be regarded as variations derived from individual-specific spaces, as mentioned in the previous section. Here, it should be noted that co-variations between intra-individual and spatial variations and co-variations between intra-individual and inter-individual variations do not theoretically exist, since these relations follow hierarchical structure (please refer to the discussions in Introduction). In this study, we attempt to examine the above mentioned four different types of variations.

The MCLCV model

The MCLCV model is a utility-maximization model in the context of discrete choice behavior, based on the multinomial logit (MNL) model. Consider the situation that an individual i ($i = 1, 2, \dots, I$) who travels in the space s ($s = 1, 2, \dots, S$) on a t -th trip ($t = 1, 2, \dots, T_i$) chooses alternative j ($j = 1, 2, \dots, J$) based on the following utility U_{jist} (please refer to Table 1 for the concrete example of utility function):

$$U_{jist} = \beta_j x_{jist} + \gamma_{ji} + \gamma_{js} + \gamma_{jis} + \gamma_{jist} + \varepsilon_{jist}, \tag{1}$$

where x_{jist} indicates explanatory variables including level-of-service variables, mobility tool ownership, decision maker’s attributes and situational variables. Let β_j be a coefficient vector associated with x_{jist} . Let ε_{jist} represent an unobserved intra-individual variation which is assumed to be independently and identically Gumbel-distributed with a variance of $\sigma^2\pi^2/6$ (the scale parameter σ is a fixed value). Let γ_{jist} be introduced to capture unobserved intra-individual variations in a manner analogous to the MXL model (Ben-Akiva et al. 2001). The terms γ_{ji} and γ_{js} are also random; these are introduced to capture unobserved inter-individual variations (i.e., variations across individuals) and unobserved spatial variations (i.e., variations across travel space), respectively. Let γ_{jis} be a co-variation term which is introduced to capture the interaction effects between inter-individual and spatial variations (i.e., the unobserved variations caused by the two cross-classified factors jointly). Here we assume that γ_{ji} , γ_{js} , γ_{jis} and γ_{jist} are normally distributed as follows:

$$\begin{aligned} \gamma_i &= (\gamma_{2i}, \dots, \gamma_{Ji}) \sim MVN\left(0, \sum_i\right), \quad \gamma_s = (\gamma_{2s}, \dots, \gamma_{Js}) \sim MVN\left(0, \sum_s\right), \\ \gamma_{is} &= (\gamma_{2is}, \dots, \gamma_{Jis}) \sim MVN\left(0, \sum_{is}\right), \quad \gamma_{ist} = (\gamma_{2ist}, \dots, \gamma_{Jist}) \sim MVN\left(0, \sum_{ist}\right), \end{aligned} \tag{2}$$

where \sum_i , \sum_s and \sum_{is} are variance–covariance matrices for each level of variation and \sum_{ist} corresponds to the variance–covariance matrix for intra-individual variation, adding $\sigma^2\pi^2/6$ in the diagonal elements. Note that only differences in utility matter in discrete choice modeling, and thus one of the utilities has to be fixed as a base (Train 2003). As this needs to be done for all variation types, we fixed γ_{1i} , γ_{1s} , γ_{1is} and γ_{1ist} as zeros. Additionally, for intra-individual variations, one of the rests of variance needs to be fixed as well to normalize the scale of utility. Thus, the full variance–covariance parameters which can theoretically be estimated are $[(J - 1)J/2] - 1$ for intra-individual variation \sum_{ist} and

$(J - 1)J/2$ for other higher-level variations Σ_i , Σ_s and Σ_{ist} . Based on the above-mentioned model specifications, the following conditional probability of the choice of mode j can be written in the MNL form

$$P(y_{jist} = j | \beta, \gamma_i, \gamma_s, \gamma_{is}, \gamma_{ist}) = \frac{\exp((\beta_j x_{jist} + \gamma_{ji} + \gamma_{js} + \gamma_{jis} + \gamma_{jist})/\sigma)}{\sum_{j'=1}^J \exp((\beta_{j'} x_{j'ist} + \gamma_{j'i} + \gamma_{j's} + \gamma_{j'is} + \gamma_{j'ist})/\sigma)} \quad (3)$$

When we specify the model with the above-mentioned full variance–covariance parameters (with small value for the scale parameter σ), the complete variation properties of all pairs of alternatives can be obtained. However, we found that it is quite difficult in empirical analysis to estimate full variance–covariance parameters (i.e., the estimation results are not stable: the estimation procedure will be mentioned later), especially in the estimation of parameters of intra-individual variations and in the simultaneous estimation of covariance parameters at different levels. This means, the model could formally and theoretically be identified yet it may exhibit very small variation in the objective function from its maximum over a wide range of parameter values, resulting in being empirically non-identifiable (Keane 1992; Grilli and Rampichini 2007). Note that, since this is not theoretical identification problem, the model with full variance–covariance parameters is potentially identifiable: there is a possibility that employing different estimation algorithm and/or different parameterizations of variance–covariance matrix could solve such informal or non-theoretical identification problem. For example, Train (2001) compared theoretical and empirical properties of classical estimation procedure (maximum simulated likelihood estimation with Halton draws) and Bayesian estimation procedure (a hierarchical Bayesian procedure), and concluded that Bayesian approach has theoretical advantage from both classical and Bayesian perspectives and the estimation results are essentially the same. Expanding such comparison analysis (especially focusing on the stability of the estimation) would be helpful to choose a proper estimation method, and it certainly remains as a big issue to be solved; however, in this paper we will only focus on the development of choice model with co-variation effects, and the discussion on estimation methods lies outside the scope of this paper.

To avoid the above mentioned identification problem, in this study γ_{jist} is omitted from the model, and we only estimated variances in $(J - 1)$ alternatives (with the covariances fixed as zero) for spatial variations and co-variations, while the full variance–covariance matrix was given for inter-individual variations. Thus, it is assumed that all sources of unobserved correlation over alternatives are derived from inter-individual matters such as the differences of attitudes and perceptions across individuals. There are several reasons for putting correlation terms into only inter-individual level. First, to assure stable model estimation, as discussed above, we had to eliminate some correlation terms. Second, based on the preliminary analysis which is conducted to identify the magnitude of unobserved variation for each level, we decided to give the highest priority to introducing correlation terms into inter-individual level.

Note that even when we exclude Σ_{ist} from the model, higher-level variations are automatically rescaled according to the fixed scale of intra-individual variations (the unknown parameters β are also rescaled), and the ratio of each variation to the total variation can be safely calculated (Grilli and Rampichini 2003, 2007). However, such rescaling in the utility U_{jist} ($j = 2, 3, \dots, J$) might occur only with respect to the utility of the first alternative U_{1ist} , in which all random components are fixed for normalization. Therefore, the variation properties of all pairs of alternatives could not be obtained, and we could only determine the variation properties of the utility U_{jist} ($j = 2, 3, \dots, J$) with respect to the utility of the first alternative, given the restricted variance–covariance matrices.

Table 1 An example of data structure

Choice result	Trip no.	Individual no.	Spatial no. (O–D pair)	Individual-spatial combination no.	Utility
Objective variable	<i>t</i> -th trip	Individual <i>i</i>	Space <i>s</i>	Individual-specific space <i>is</i>	$U_{jist} = \gamma_{ji} + \gamma_{js} + \gamma_{jis} + \epsilon_{jist}$ (Null model)
Bicycle	1	1	1	1(1-1)	$\gamma_{j1} + \gamma_{j1} + \gamma_{j11} + \epsilon_{j111}$
Car (driver)	2	1	2	2(1-2)	$\gamma_{j1} + \gamma_{j2} + \gamma_{j12} + \epsilon_{j122}$
Bicycle	3	1	1	1(1-1)	$\gamma_{j1} + \gamma_{j1} + \gamma_{j11} + \epsilon_{j113}$
Car (driver)	4	1	2	2(1-2)	$\gamma_{j1} + \gamma_{j2} + \gamma_{j12} + \epsilon_{j124}$
Bicycle	5	1	1	1(1-1)	$\gamma_{j1} + \gamma_{j1} + \gamma_{j11} + \epsilon_{j115}$
Public transport	6	1	2	2(1-2)	$\gamma_{j1} + \gamma_{j2} + \gamma_{j12} + \epsilon_{j126}$
Public transport	7	1	1	1(1-1)	$\gamma_{j1} + \gamma_{j1} + \gamma_{j11} + \epsilon_{j117}$
Car (driver)	1'	2	2	3(2-2)	$\gamma_{j2} + \gamma_{j2} + \gamma_{j22} + \epsilon_{j221'}$
Public transport	2'	2	1	4(2-1)	$\gamma_{j2} + \gamma_{j1} + \gamma_{j21} + \epsilon_{j212'}$
Car (driver)	3'	2	2	3(2-2)	$\gamma_{j2} + \gamma_{j2} + \gamma_{j22} + \epsilon_{j223'}$
Bicycle	4'	2	1	4(2-1)	$\gamma_{j2} + \gamma_{j1} + \gamma_{j21} + \epsilon_{j214'}$
Car (passenger)	5'	2	2	3(2-2)	$\gamma_{j2} + \gamma_{j2} + \gamma_{j22} + \epsilon_{j225'}$
Public transport	6'	2	1	4(2-1)	$\gamma_{j2} + \gamma_{j1} + \gamma_{j21} + \epsilon_{j216'}$
Car (driver)	7'	2	2	3(2-2)	$\gamma_{j2} + \gamma_{j2} + \gamma_{j22} + \epsilon_{j227'}$
Public transport	8'	2	1	4(2-1)	$\gamma_{j2} + \gamma_{j1} + \gamma_{j21} + \epsilon_{j218'}$

In particular, with the scale parameter σ set to 1 (which also leads to rescaling the other fixed and random parameters), this study employs the following model:

$$U_{jist} = \beta_j x_{jist} + \gamma_{ji} + \gamma_{js} + \gamma_{jis} + \epsilon_{jist}, \tag{4}$$

$$P(y_{jist} = j | \beta, \gamma_i, \gamma_s, \gamma_{is}) = \frac{\exp(\beta_j x_{jist} + \gamma_{ji} + \gamma_{js} + \gamma_{jis})}{\sum_{j'=1}^J \exp(\beta_{j'} x_{j'ist} + \gamma_{j'i} + \gamma_{j's} + \gamma_{j'is})}, \tag{5}$$

where

$$\begin{aligned} \gamma_i &= (\gamma_{2i}, \dots, \gamma_{Ji}) \sim MVN\left(0, \sum_i\right), \\ \gamma_{js} &\sim N\left(0, \sigma_{js}^2\right), \gamma_{jis} \sim N\left(0, \sigma_{jis}^2\right) j = 2, 3, \dots, J \end{aligned} \tag{6}$$

Henceforth, the diagonal elements (i.e., variances) and non-diagonal elements (i.e., covariances) in Σ_i are described as σ_{ji}^2 and $\sigma_{j'j'i}$, respectively ($j = 2, 3, \dots, J$). As mentioned above, all random components in the first alternative are set to zero. All of the following descriptions are based on this model, with limited variance–covariance matrices shown in Eqs. 5 and 6.

Note that if the σ_{jis}^2 is close to zero, the MCLCV model becomes an MCL model (Bhat 2000), and if the σ_{jis}^2 and σ_{js}^2 are equal to zero, the MCLCV model becomes an MXL model

with panel data. Thus, the MCLCV model is a straightforward extension of such mixed-logit-type models. Although the MCLCV model itself is just a simple extension of a mixed logit model, it can capture various kinds of co-variation effects. We could quantitatively identify not only individual-specific spatial effects, but also social interaction effects, for example by considering co-variations as interaction terms across different reference groups’ effects on the behavior of a target person. The proposed method could be a powerful tool to check the existence of co-variation effects in various contexts.

Variation properties of utility difference

Here we shall mention a way to describe behavioral variations in the above-mentioned model, which we used in the empirical study mentioned in next section. Conventional ways have often focused only on observed variations which can be directly connected to policy discussions. This study follows a somewhat different approach. Concretely speaking, all behavioral variations are first treated as unobserved variations in order to determine what kinds of variations really exist (i.e., $\beta_j x_{j\text{ist}}$ in Eq. 4 is excluded). In other words, we first decompose all variations into four different types of unobserved variations, including unobserved inter-individual variations, unobserved spatial variations, unobserved intra-individual variations and unobserved co-variations between inter-individual and spatial variations. Using the symbol “ \sim ” to represent model estimation results without any explanatory variables (called the *Null* model), the total variance of utility difference can be calculated as follows:

$$\text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}}) = \tilde{\sigma}_{ji}^2 + \tilde{\sigma}_{js}^2 + \tilde{\sigma}_{jis}^2 + \pi^2/3 \tag{7}$$

This variation decomposition can provide several useful insights. For example, when substantial amounts of intra-individual variations are observed, people’s mode choice behavior would strongly depend on the situational or context-dependent attributes given at the time. Similarly, if almost all behavioral variations are derived from spatial variations, this means that there are no individual or situational differences and spatial aggregate data can be used for the analysis. Also, the more inter-individual variations are observed, the greater the possible impacts of individual attributes (such as age, gender, car ownership and season ticket ownership). Such observations could help us, for example, to identify which levels of influential factors are likely to be more important. Additionally, this variation decomposition allows us to detect pure co-variation effects (i.e., individual-specific spatial effects) before introducing explanatory variables.

In the next step, we shall introduce a set of explanatory variables to provide reasons for the behavioral variations measured in the *Null* model. Using the symbol “ \wedge ” to represent model estimation results with a set of explanatory variables (called *Full* model), the total variance of utility difference can be calculated as follows:

$$\text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) = \text{Var}(\hat{\beta}_j x_{j\text{ist}} - \hat{\beta}_1 x_{1\text{ist}}) + \hat{\sigma}_{ji}^2 + \hat{\sigma}_{js}^2 + \hat{\sigma}_{jis}^2 + \pi^2/3 \tag{8}$$

Introducing explanatory variables could change some behavioral variations into observed variations while the rest remain unobserved variations. Our purpose here is to evaluate what types and how many of the variations can be captured by introducing explanatory variables. To do this, we compare the variation components in Eq. 8 against those in Eq. 7. Here, although the absolute expected value of $\text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}})$ may change depending on how many intra-individual variations can be captured by introducing explanatory variables,

the component ratio for each variation can be compared between the different models as long as the existence of the same “true” utility can be expected. This is because the scale of $\text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}})$ is strictly defined by the rest of the unobserved intra-individual variations, and also because the other fixed and random parameters are automatically rescaled. Thus, we can compare the component ratio for each variation between Eq. 7 and Eq. 8. For example, in case of unobserved inter-individual variation, the component ratio in the *Null* model can be calculated as $\tilde{\sigma}_{ji}^2 / \text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}})$ and that in the *Full* model can be calculated as $\hat{\sigma}_{ji}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}})$. Since a part of variation in the *Null* model is expected to be explained by introduced explanatory variables in the *Full* model, the component ratio in the *Full* model would become smaller than that in the *Null* model, and the difference between them would be the observed variation. Such calculations can show which types and how many of the variations can or cannot be captured by introducing a certain set of explanatory variables, as follows:

For observed inter-individual variations (%):

$$\left(\tilde{\sigma}_{ji}^2 / \text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}}) - \hat{\sigma}_{ji}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{9}$$

For unobserved (or remaining) inter-individual variations (%):

$$\left(\hat{\sigma}_{ji}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{10}$$

For observed spatial variations (%):

$$\left(\tilde{\sigma}_{js}^2 / \text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}}) - \hat{\sigma}_{js}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{11}$$

For unobserved (or remaining) spatial variations (%):

$$\left(\hat{\sigma}_{js}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{12}$$

For observed co-variations (%):

$$\left(\tilde{\sigma}_{jis}^2 / \text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}}) - \hat{\sigma}_{jis}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{13}$$

For unobserved (or remaining) co-variations (%):

$$\left(\hat{\sigma}_{jis}^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{14}$$

For observed intra-individual variations (%):

$$\left(3^{-1}\pi^2 / \text{Var}(\tilde{U}_{j\text{ist}} - \tilde{U}_{1\text{ist}}) - 3^{-1}\pi^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100, \tag{15}$$

For unobserved (or remaining) intra-individual variations (%):

$$\left(3^{-1}\pi^2 / \text{Var}(\hat{U}_{j\text{ist}} - \hat{U}_{1\text{ist}}) \right) \times 100 \tag{16}$$

The variation properties derived from Eqs. 7 and 8 and subsequently from Eqs. 9–16 could provide many implications for both data collection and model development. For example, if unobserved intra-individual variations are quite high even after introducing available explanatory variables, this indicates that we might need many more situational attributes to provide richer explanations for behavioral variations. In other words, Eqs. 9–16 can be

used to evaluate the model’s performance in greater detail by evaluating the model for each alternative and for each variation type. In the empirical analysis, we introduced four sets of explanatory variables in a sequential manner. This sequential procedure allowed us to evaluate the impacts of each set of explanatory variables. In other words, we could confirm which types of variations could actually be captured by which subset of explanatory variables. Note that although the order in which the variable subsets are introduced would not influence the variation properties if all the variables in the model are independent of each other (and this assumption was already made when the model was estimated), this assumption usually cannot be satisfied completely. For example, the dummy variable “male” (a decision maker’s attribute) might be correlated with the dummy variable “trip purpose (work or not)” (a situational attribute), and as a result, the order in which the variable subsets are introduced may somewhat influence the variation properties. In this study, we decided on the order of introducing variable subsets by considering the importance of each as well as the ease of data acquisition. In the example above, we would give priority to the variable “male” over the variable “trip purpose”, because the former is much easier to observe. Furthermore, if we can capture variations by “male” rather than by “trip purpose”, we could easily extend the application of the model to the entire population of the target area. Thus, after capturing variations by “male”, the variable “trip purpose” just serves to capture the remaining variations which could not be captured by the explanatory variables already introduced. However, how the implications differ according to the order of introducing explanatory variable sets still remains as an important future task to be examined.

Estimation procedure

The likelihood function based on Eq. 4 can be written as follows:

$$L_{jst} \left(\beta, \sum_i, \sigma_s, \sigma_{is} | y, x \right) = \int_{\gamma_{is}} \int_{\gamma_s} \int_{\gamma_i} \prod_{N_{ist}} \prod_j P(y_{jist} = j | \beta, \gamma_i, \gamma_s, \gamma_{is})^{\delta_j} \times f \left(\gamma_i | \sum_i \right) f(\gamma_s | \sigma_s) f(\gamma_{is} | \sigma_{is}) d\gamma_{is} d\gamma_s d\gamma_i, \tag{17}$$

where N_{ist} is the number of samples. The dummy variable δ_j takes the value 1 if individual i chooses mode j on the t -th trip, and otherwise it takes the value 0. The term $f(\gamma_i | \Sigma_i)$ follows a multivariate normal distribution with a mean of 0 and a variance–covariance of Σ_i . The terms $f(\gamma_s | \sigma_s)$ and $f(\gamma_{is} | \sigma_{is})$ also follow multivariate normal distributions with a mean of 0 and variances of σ_{js} and σ_{jis} ($j = 2, 3, \dots, J$), respectively, but these covariances are fixed as 0 (see Eq. 5).

The likelihood function of Eq. 17 involves a number of integrations and can not be solved analytically. To estimate such a model, some simulation methods are usually adopted, such as a series of Monte Carlo methods or numerical quadrature methods (e.g., Pinheiro and Bates 1995; Bhat 2001; Train 2003). In this study, we used a hierarchical Bayesian procedure based on the Markov chain Monte Carlo (MCMC) method, which has recently become popular and is a promising method for estimating multilevel models with complicated random effects (Train 2003; Gelman et al. 2004). The method incorporates prior distribution assumptions and, based upon successive sampling from the posterior distribution of the model’s parameters, yields a chain which is then used for making point and interval estimations. In particular, the posterior distribution can be written as follows:

$$\pi\left(\beta, \sum_i \sigma_s, \sigma_{is} | y, x\right) = \prod_{N_{ist}} \prod_j P(y_{jist} = j | \beta, \gamma_i, \gamma_s, \gamma_{is})^{\delta_j} f\left(\gamma_i | \sum_i\right) f(\gamma_s | \sigma_s) f(\gamma_{is} | \sigma_{is}) \phi \times \left(\sum_i\right) \phi(\sigma_s) \phi(\sigma_{is}) \phi(\beta) \tag{18}$$

where we assume an inverted Gamma distribution for $\varphi(\sigma_s)$ and $\varphi(\sigma_{is})$ (for each element), an inverted Wishart distribution for $\varphi(\Sigma_i)$ and a normal distribution for $\varphi(\beta)$ (for each element) as prior distributions. These two prior distributions correspond to conjugate prior distributions, which mean that the posterior distributions can be specified within the same distributional families as the prior distributions. The terms $f(\gamma_i | \Sigma_i)$, $f(\gamma_s | \sigma_s)$, and $f(\gamma_{is} | \sigma_{is})$ create a so-called hierarchical sampling procedure. In this study, non-informative prior distributions were assumed for all parameters. In this case, the estimated parameters are asymptotically equivalent to the parameters obtained through a simulated maximum likelihood method (Train 2003). Note that, even when we employ the different prior distributions that are not conjugate prior distributions, we may also obtain asymptotically equivalent parameters estimated with conjugate prior distributions, although a conjugate prior distribution is much more algebraic convenience, giving a closed-form expression for the posterior.

Draws from the posterior are obtained using the software WinBUGS (Bayesian inference Using Gibbs Sampling [Lunn et al. 2000]). In the MCMC sampling, draws of each parameter are obtained from its posterior conditional on the other parameters (Gelman et al. 2004). The conditional posterior distribution for each parameter can be found in Train (2003), for example. The stabilities of the estimation results reported in this paper were checked a number of ways, including: checking the trace plot and autocorrelation in each parameter chain; using the Geweke diagnostic (Geweke 1992); and checking the results with different initial values of parameters. All of these indicated that the models reported in this paper are well converged.

Data

The *Mobidrive* data set (Axhausen et al. 2002) was used in the empirical analysis. The data are from a continuous six-week travel diary survey that was conducted in Karlsruhe and Halle in Germany in 1999. A total of 317 persons from 139 households took part in the main survey.

The mode choice set used in this empirical analysis includes the following options: “Car (passenger)”, “Bicycle”, “Car (driver)” and “Public transport”. After data cleaning with respect to the availability of level-of-service values (only travel time in this study) for both used and non-used alternatives, a total of 18,326 trips made by 309 individuals were selected. The mode choice shares are: “Car (passenger)” (15.6%), “Bicycle” (20.7%), “Car (driver)” (39.6%) and “Public transport” (24.1%).

The spatial variation was defined based on differences in the Origin–Destination (O–D) distributions and the residential locations. Note that the origins, destinations and residential locations are classified into three zones: “CBD (Central Business District)”, “inner city” and “others” for each city. The maximum number of combinations is 54 (= 3[origins] × 3[destinations] × 3[residential locations] × 2[cities]). In the samples used in this study, 53 spaces were observed. Thus, the 18,326 samples are nested within 1,635 individual-specific spaces (i.e., the unit for co-variations), and the 1,635 individual-specific spaces are nested within 53 spaces (i.e., the unit for spatial variations) and within 309

individuals (i.e., the unit for inter-individual variations). A simple example of data structure (in case of two individuals and two O–D pairs) is shown in Table 1. In this example, it can be confirmed that individual 1 tends to choose “Bicycle” compared to individual 2. This difference may generate inter-individual variations. As for the spatial differences, “Car (driver)” seems to be a preferable alternative when they travel in O–D pair 2. This difference may generate spatial variations. Additionally, it can be confirmed that co-variations (i.e., individual-specific spatial effects) could be also generated, since individual 1 tends to use “Bicycle” for traveling in O–D pair 1 while individual 2 tends to use “Public transport” even for traveling in the same O–D pair. The purpose of introducing co-variation term is to capture such variations.

Here it should be noted that, in general, the more complicated variation structures are employed, the more samples are required. This is similar with the case that a model having more explanatory variables requires more samples. For this reason, we employ relatively rough location segmentation as mentioned above. Thus, one should keep in mind that the spatial variations identified here would be quite rough and may not reflect small spatial differences. Also, to distinguish between inter-individual and intra-individual variations, we need a multi-day travel diary survey data.

Model estimation and discussion

In this section, we shall first report the estimation results of the MCLCV model and the MCL model to determine whether co-variation effects really exist or not. The estimations were done without any explanatory variables to detect the pure influence of co-variations. Explanatory variables were then introduced in a sequential manner. We first introduced only *Travel time* variables. We then sequentially added the variables *Mobility tools*, *Decision-maker’s attributes* and *Situational attributes*.

Table 2 shows the definitions of the introduced explanatory variables. With regard to *Travel time*, network level-of-service (LOS) values were used for all modes. Theoretically, it can be expected that *Travel time* would mainly capture spatial variations. Public transport travel time included transfer waiting time. Travel cost might be another important variable in mode choice, but we could not introduce it because the data were unavailable. In addition, attention needed to be paid to the type of LOS variables used to capture behavioral variations. For example, one would expect LOS variables derived from network data to mainly capture spatial variations, while LOS variables based on responses from respondents would partly represent inter-individual variations because of the reflection of the respondents’ perception of travel times and costs. Disparities between different types of LOS variables as well as the effects of other LOS variables such as travel time and the number of transfers would certainly be worth investigating in future studies.

Mobility tools variables included: license-holder dummy variables, the number of household members who were the main car users, and the interaction term of these two variables for the option “Car (passenger)”; the number of bicycles in a household divided by the number of household members for the option “Bicycle”; a dummy variable of the main car users (obtained from the vehicle question “Who is the main user of this vehicle?”) for the option “Car (driver)”; and the number of season tickets for the option “Public transport”. *Decision maker’s attributes* included: household income and age (as categorical variables), gender, marital status with child(ren), and employment status. *Mobility tools* and *decision maker’s attributes* would mainly capture the inter-individual variations. The following *Situational attributes* were selected: trip purpose, day of the

Table 2 Explanatory variables used for model estimations

Explanatory variable	Definition	Mean	SD
<i>Travel time</i>	Travel time: Car driver and passenger (min.)	10.267	5.976
	Travel time: Bicycle (min.)	20.587	14.030
	Travel time: Public transport (min.) [including waiting time]	24.076	16.462
<i>Mobility tools</i>			
License	Vehicle license (1 = yes; 0 = no)	0.701	0.458
HHcaruse	Number of household members who are main car users	0.611	0.634
Lic_Hhcar	Interaction effect of “Vehicle license” and “Number of HH members who are main car users”	0.261	0.534
Mob_Bike	Number of bicycles in HH/number of HH members	0.873	0.452
Mob_Car	Main car user (1 = yes; 0 = no)	0.433	0.496
Mob_PT	Number of season tickets	0.356	0.497
<i>Decision maker's attributes</i>			
Marr_chi	Married with child(ren)	0.327	0.469
Income	Household income (in 1000 DM)	4.635	1.992
Male	Male (1 = yes; 0 = no)	0.470	0.499
Age 24	>24 years of age (1 = yes; 0 = no)	0.274	0.446
Age 25–34	25–34 years of age (1 = yes; 0 = no)	0.099	0.299
Age 45–54	45–54 years of age (1 = yes; 0 = no)	0.173	0.378
Age 55–64	55–64 years of age (1 = yes; 0 = no)	0.139	0.346
Age 65–	<65 years of age (1 = yes; 0 = no)	0.074	0.261
Fullwork	Full-time worker (1 = yes; 0 = no)	0.339	0.473
Student	Student (1 = yes; 0 = no)	0.041	0.198
Retired	Retired person (1 = yes; 0 = no)	0.129	0.335
<i>Situational attributes</i>			
TP_Pick	Trip purpose = pick up/drop off (1 = yes; 0 = no)	0.044	0.206
TP_WoRe	Trip purpose = work related (1 = yes; 0 = no)	0.030	0.171
TP_School	Trip purpose = school (1 = yes; 0 = no)	0.064	0.245
TP_Work	Trip purpose = work (1 = yes; 0 = no)	0.114	0.317
TP_Dshop	Trip purpose = daily shopping (1 = yes; 0 = no)	0.081	0.273
TP_Nshop	Trip purpose = non-daily shopping (1 = yes; 0 = no)	0.043	0.203
TP_Leis	Trip purpose = leisure (1 = yes; 0 = no)	0.148	0.355
DW_Fri	Day of week = friday (1 = yes; 0 = no)	0.165	0.371
DW_Sat	Day of week = saturday (1 = yes; 0 = no)	0.115	0.319
DW_Sun	Day of week = sunday (1 = yes; 0 = no)	0.072	0.258
DT_Morn	Departure time = 6:00 a.m.– 8:59 a.m. (1 = yes; 0 = no)	0.175	0.380
DT_Even	Departure time = 6:00 p.m.– 8:59 p.m. (1 = yes; 0 = no)	0.118	0.323
DT_Leven	Departure time = 9:00 p.m.–11:59 p.m. (1 = yes; 0 = no)	0.042	0.200
Party_HH	Size of travel party (household members)	0.307	0.605
Party_OT	Size of travel party (other members)	0.233	1.478

week, departure time and size of the travel party. Intuitively, these variables may capture intra-individual variations. The above mentioned explanatory variables were introduced in certain alternatives based on a preliminary cross-tabulation analysis.

Note that the introduced explanatory variables are of course not perfect, and omitted variables (including travel cost) should be reflected in the unobserved variations. Although explaining all variations (i.e., there is no unobserved variation) might be impossible, the remaining unobserved variation for each level could be used for a guidepost for the next data collection.

Model comparison with and without co-variations

As mentioned above, only differences in utility matter in discrete choice model, and we thus had to fix one of the utilities as a base. For this case study, we selected “Car (passenger)” as the base alternative.

The estimation results of the model without co-variations (the MCL model) and the model with co-variations (the MCLCV model) are shown in Table 3. The estimations were first done without any explanatory variables (*Null* models) to detect the pure influence of co-variation terms. The goodness-of-fit measures of DIC (deviance information criteria) indicate that co-variation effects certainly exist at a significant level, i.e., the spatial effects on mode choice behavior differed from individual to individual. The component ratio of each variation type is shown in Fig. 2. The co-variation shares in the *Null* model are 9.2% in the category “Bicycle”, 1.9% in “Car (driver)” and 12.7% in “Public transport”. Thus, while the utility in “Car (driver)” was neither much affected by individual-specific spatial effects nor by common spatial effects (i.e., the spatial variation was quite small at only 0.4%), individual-specific spatial effects might have to be considered in the utility specifications for the categories “Bicycle” and “Public transport”. Additionally, even after introducing co-variation terms into the model, the spatial variation in the “Public transport” option remained significant (9.0%). These results indicate that there are certainly two different types of spatial effects: spatial effects which are common to all individuals (spatial variations) and those which are different across individuals (co-variations). The quantitative discrimination between these two spatial effects is quite important in terms of data collection and model development, because when we developed behavioral models with cross-sectional data, the model could not distinguish between them. This means, for example, that if all spatial effects are individual-specific, the model just describes differences in perceptions or meanings of space across individuals rather than a certain spatial attribute itself, implying that there is a possibility that the change of spatial attributes has little impact on behavior. To distinguish between these two spatial effects, a multi-day set of survey data is needed, analogous to distinguishing between inter-individual and intra-individual variations (Hanson and Huff 1986).

With regard to other variation types, it was found that inter-individual variations dominated, especially in the category “Car (driver)” (88.5%). Also, in the categories “Bicycle” and “Public transport”, the shares reached 71.9 and 58.3% respectively, and the intra-individual variations were only 9.2–20.0%. These results seem to be quite significant compared to the proportions of variations in other behavioral aspects. Concretely speaking, previous studies have shown higher shares of intra-individual variations in many cases: around 50–60% for the number of trips per day, travel time and travel distance (Pas 1987; Pendyala 1999); 35–85% (depending on the activity type) for departure time choice (Kitamura et al. 2006; Chikaraishi et al. 2009); and 17–65% for time use behavior (Goulias 2002; Chikaraishi et al. 2010). Given these results, we could say that mode choice behavior is relatively stable from day to day compared to other behavioral aspects. This might be quite important for data collection: when we are only interested in mode choice behavior, behavioral observations on shorter days might be acceptable compared to other behavioral aspects, although it is of course better to observe behavior across a longer period, since

Table 3 Estimation results of *Null* models with/without co-variations

Variables	Without co-variations		With co-variations	
	Mean	SD	Mean	SD
<i>Constant</i>				
Bicycle	-2.524	0.370	-3.071	0.377
Car (driver)	-0.867	0.380	-0.969	0.373
Public transport	0.020	0.253	-0.265	0.300
<i>Spatial variations</i>				
Bicycle	1.158	0.489	0.855	0.418
Car (driver)	0.221	0.081	0.133	0.073
Public transport	1.091	0.305	1.481	0.441
<i>Inter-individual variations</i>				
Bicycle	14.90	2.278	15.72	2.472
Car (driver)	29.26	4.220	31.58	4.611
Public transport	7.669	0.966	9.598	1.305
<i>Covariance at inter-individual level</i>				
Bicycle and car (driver)	3.916	1.926	3.413	2.116
Bicycle and public transport	1.444	0.928	1.509	1.138
Car (driver) and public transport	-0.967	1.292	-1.979	1.536
<i>Co-variations (spatial and inter-individual variations)</i>				
Bicycle	-	-	2.007	0.302
Car (driver)	-	-	0.665	0.139
Public transport	-	-	2.087	0.261
<i>Intra-individual variations (for all)</i>				
The log-likelihood at sample shares	3.290	-	3.290	-
Log{ $\pi[x \text{mean}(\theta)]$ }	-24294.5			
Log{ $\text{mean}[\pi(x \theta)]$ }	-9185.4		-7837.7	
Log{ $\text{mean}[\pi(x \theta)]$ }	-9522.9		-8478.1	
pD	675.1		1280.8	
DIC	19720.9		18237.0	
Sample size	18,326			

Notes: Car (passenger) is a base alternative. “ $\log\{\pi[x|\text{mean}(\theta)]\}$ ” represents the log likelihood with the posterior means of parameters. “ $\log\{\text{mean}[\pi(x|\theta)]\}$ ” represents the posterior mean of the log likelihood. “pD” is defined as $2(\log\{\text{mean}[\pi(x|\theta)]\} - \log\{\pi[x|\text{mean}(\theta)]\})$, which is used as the Bayesian measure of model complexity. “DIC” stands for Deviance Information Criterion which is defined as $-2(\log\{\pi[x|\text{mean}(\theta)]\} - \text{pD})$. DIC can be viewed as a Bayesian analogue of AIC (Akaike Information Criterion). For details, please refer to Spiegelhalter et al. (2002)

there are certain intra-individual variations. In addition, because the significant variations in travel choice behavior are inter-individual variations, we may have to pay much more attention to individual-specific influential factors such as individual attributes, mobility tool ownership and habitual preferences.

Estimation results with explanatory variables

Next, we sequentially introduced four sets of explanatory variables (*Travel time, Mobility tools, Decision maker’s attributes* and *Situational attributes*: see Table 2) into the *Null*

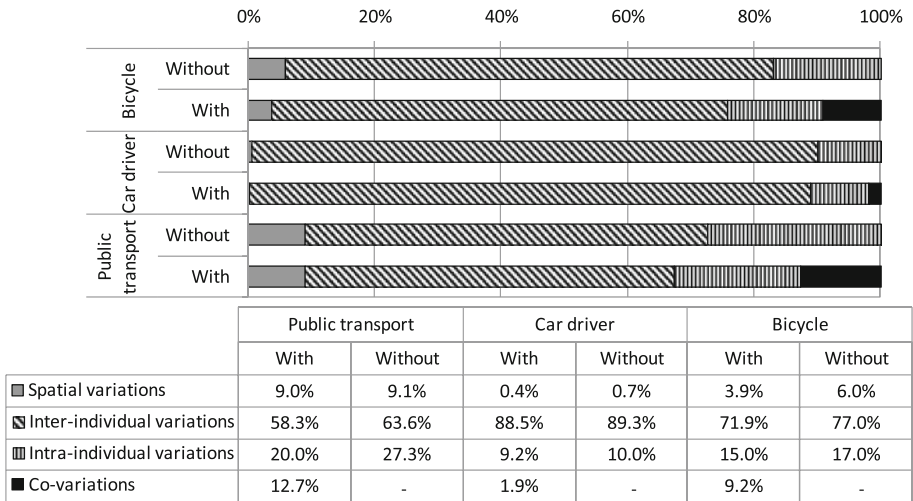


Fig. 2 Proportions of variations in *Null* models with/without co-variations

model with the co-variations discussed in the previous subsection. We estimated four more models and compared their variation properties. Here we shall first look at the estimation results with the full set of explanatory variables (called the *Full* model) in order to check the impacts of the introduced explanatory variables. We shall then focus on sequential changes in the variation properties following the addition of a certain set of explanatory variables.

Table 4 shows the estimation results of the *Full* model. As expected, the parameter *Travel time* has a negative sign and is statistically significant. In the category *Mobility tools*, non-holders of licenses with other household members who were the main car users tended to choose the option “Car (passenger)”, while the main car users were likely to choose “Car (driver)”. The variables of bicycle and season ticket ownership also had positive impacts on the use of “Bicycles” and “Public transport”. In the category *Decision maker’s attributes*, married males who had at least one child were likely to choose the option “Car (driver)”, while young people (under 24 years of age) and people with high incomes tended to use “Bicycles” and “Public transport”, compared to Car (passengers). At the same time, interestingly, age effects on the options “Bicycle” and “Public transport” seem to follow a quadratic function, i.e., the young and elderly people tended to choose “Bicycle” and “Public transport”. With regard to the effects of trip purpose on mode choice in the category *Situational attributes*, pick up and drop off behavior of course had a positive impact on the “Car (driver)” share, as did work-related trips. People were not likely to use “Public transport” for daily shopping trips. Also, on weekends they used less “Public transport” and “Bicycles”, while the share of “Car (passengers)” tended to increase. The estimation results of the departure time variables indicate that the share of “Car (passengers)” tended to increase from morning to late evening, while “Public transport” was likely to decrease: there is a possibility that “Car (passengers)” and “Public transport” substitute for each other, depending on the time of day. It can be expected that non-mandatory activities, such as going to restaurant or shopping, might be dominant purposes of night trips, while mandatory activities might be dominant purposes of morning trips. Since non-mandatory activities are often done with families or friends, they could

Table 4 Estimation results of the *Full* model

Variables	Car (passenger)		Bicycle		Car (driver)		Public transport	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Constant			-9.153	1.447	-3.597	0.848	-2.080	1.089
<i>Travel time</i>	-0.033	0.004	-0.033	0.004	-0.033	0.004	-0.033	0.004
<i>Mobility tools</i>								
License	-2.614	0.672	0.0	-	0.0	-	0.0	-
HHcaruse	0.975	0.312	0.0	-	0.0	-	0.0	-
Lic_Hhcar	0.051	0.586	0.0	-	0.0	-	0.0	-
Mob_Bike	0.0	-	3.872	0.660	0.0	-	0.0	-
Mob_Car	0.0	-	0.0	-	3.520	0.448	0.0	-
Mob_PT	0.0	-	0.0	-	0.0	-	3.096	0.507
<i>Decision maker's attributes</i>								
Marr_chi	0.0	-	0.384	0.861	1.469	0.515	-0.192	0.683
Income	0.0	-	0.697	0.160	0.158	0.111	0.224	0.133
Male	0.0	-	1.320	0.699	2.401	0.441	0.011	0.520
Age24	0.0	-	2.111	1.116	-3.127	0.852	2.526	0.977
Age25-34	0.0	-	0.913	1.277	1.021	0.685	0.230	1.039
Age45-54	0.0	-	-0.622	1.017	-0.110	0.564	1.085	0.861
Age55-64	0.0	-	-1.006	1.193	-0.144	0.624	1.258	0.954
Age65-	0.0	-	2.856	1.842	0.543	1.031	4.713	1.376
Fullwork	0.0	-	-1.374	0.842	0.347	0.477	-0.514	0.671
Student	0.0	-	-0.577	1.642	-0.031	1.004	-2.647	1.245
Retired	0.0	-	-2.002	1.634	0.029	0.794	-1.524	1.137
<i>Situational attributes</i>								
TP_Pick	0.0	-	0.0	-	1.947	0.179	0.0	-
TP_WoRe	0.0	-	0.0	-	0.606	0.253	0.0	-
TP_School	0.0	-	-0.554	0.207	0.0	-	0.204	0.232
TP_Work	0.0	-	0.0	-	-0.152	0.129	0.0	-
TP_Dshop	0.0	-	-0.156	0.138	0.0	-	-1.053	0.183
TP_Nshop	0.057	0.082	0.0	-	0.0	-	0.0	-
TP_Leis	0.163	0.050	0.0	-	0.0	-	-0.156	0.127
DW_Fri	0.075	0.056	-0.009	0.111	0.0	-	-0.012	0.117
DW_Sat	0.265	0.060	-0.607	0.139	0.0	-	-0.609	0.149
DW_Sun	0.145	0.073	-0.293	0.185	0.0	-	-0.833	0.185
DT_Morn	-0.172	0.074	0.0	-	0.0	-	0.498	0.139
DT_Even	0.304	0.054	0.0	-	0.0	-	-0.698	0.141
DT_Leven	0.565	0.079	0.0	-	0.0	-	-0.793	0.217
Party_HH	0.0	-	-3.481	0.120	-2.003	0.088	-2.234	0.099
Party_OT	0.0	-	-0.922	0.079	-0.733	0.080	0.0	-
<i>Random effects</i>								
Spatial variations			0.687	0.342	0.066	0.063	1.906	0.584
Inter-individual variations			18.04	2.900	5.920	1.071	13.20	1.962
Co-variations			1.548	0.332	0.858	0.158	2.505	0.332
Intra-individual variations			3.290	-	3.290	-	3.290	-

Table 4 continued

Variables	Car (passenger)		Bicycle		Car (driver)		Public transport	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i>Covariance at the inter-individual level</i>								
Between bicycle and car (driver)					1.347	(SD = 1.218)		
Between bicycle and public transport					4.183	(SD = 1.596)		
Between car (driver) and public transport					-0.743	(SD = 0.9872)		
The log-likelihood at sample shares					-24294.5			
Log{ $\pi[x \text{mean}(\theta)]$ }					-6271.0			
Log{ $\text{mean}[\pi(x \theta)]$ }					-6881.8			
pD					1221.7			
DIC					14985.3			

Notes: Car (passenger) is a base alternative. The cells with mean (0.0) and s.d. (-) indicate corresponding parameters are fixed as zero based on preliminary cross-tabulation analysis

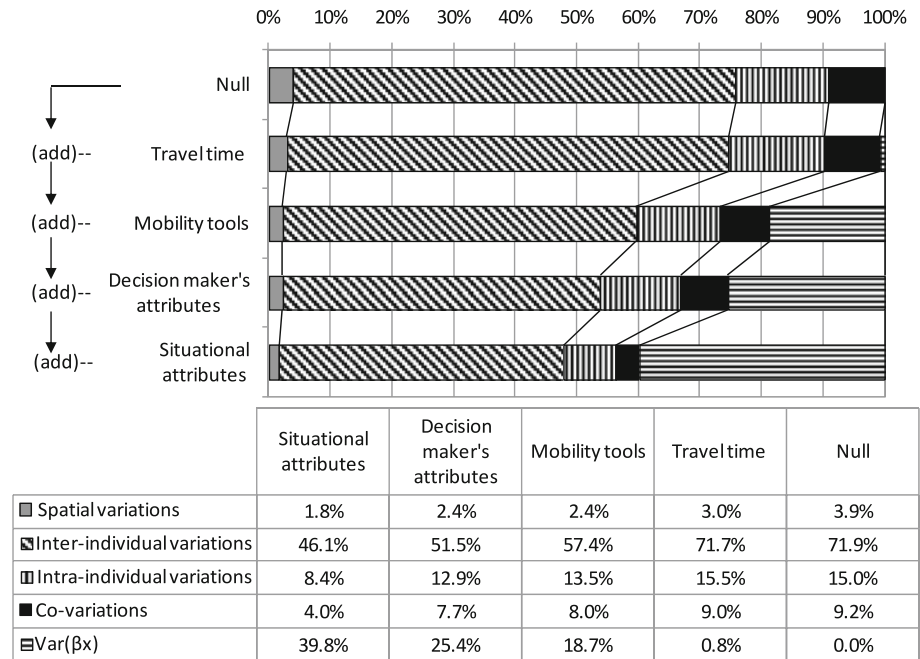


Fig. 3 Proportions of variations in “Bicycle”

easily be car passengers of their friends or family members. In fact, the results also indicate that increases in the size of travel parties cause the increases of the share of “Car (passengers)”. Overall, the impacts of the introduced explanatory variables are quite consistent with previous studies and are intuitively plausible.

Next, we shall consider the sequential changes of variation properties along with the addition of a certain set of explanatory variables. Figures 3, 4, and 5 show the results of the sequential changes of proportions of four unobserved variations (spatial variations,

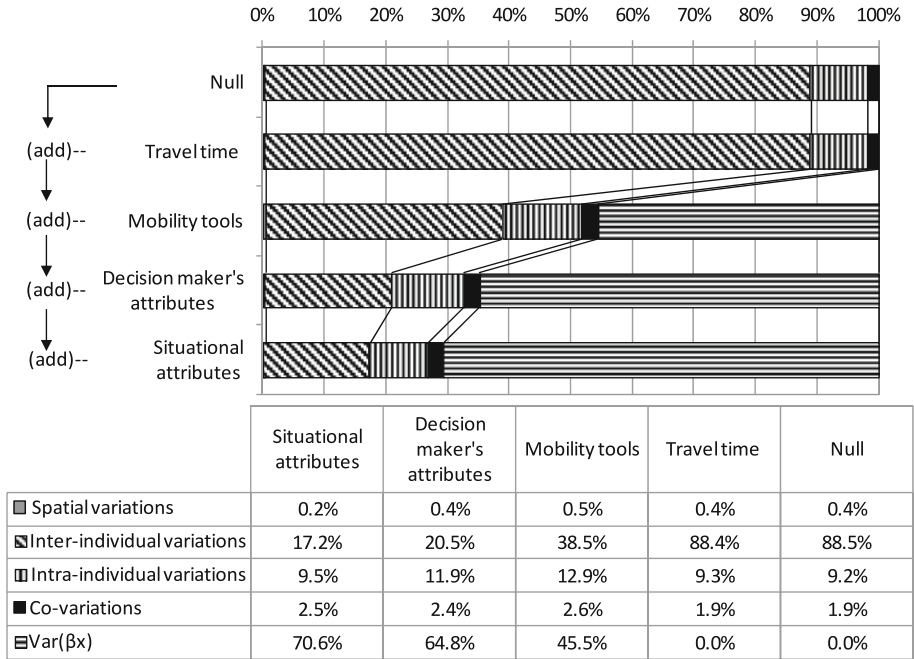


Fig. 4 Proportions of variations in “Car (driver)”

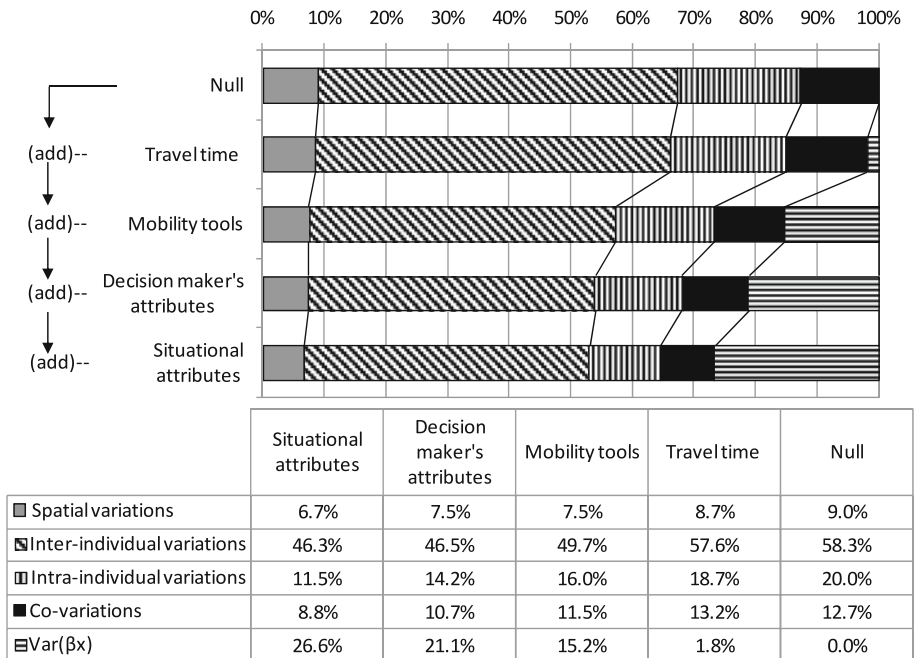


Fig. 5 Proportions of variations in “Public transport”

inter-individual variations, intra-individual variations and co-variations) as well as observed variations which were captured by introduced explanatory variables. As expected, the unobserved variations basically decreased when a set of explanatory variables was added.

First, by introducing the variable *Travel time*, 0.0–1.8% of the total variations could be classified as observed variations. For the option “Car (driver)”, the observed variation was 0.0%, because the same travel times are assumed for both “Car (passengers)” and “Car (drivers)”. Amazingly, the contribution of *Travel time* to the reduction of unobserved variations was rather small. It might be worth investigating the extent to which detailed level-of-service variables which were not introduced in this study (such as travel costs and the number of transfers) could reduce unobserved variations. This may be important for establishing a model to use for future predictions and policy evaluations: we may at least have to confirm whether a model with a general set of level-of-service variables would suffice for making predictions.

Second, we added *Mobility tools* explanatory variables to the model. The proportions of observed variations to the total amount of variations become 18.7% in the option “Bicycle”, 45.5% in “Car (driver)” and 15.2% in “Public transport”. Compared to the *Travel time* variable, the *Mobility tools* variables captured behavioral variations well, especially in the category “Car (driver)”. From a practical perspective, the results seem to suggest that policies related to *Mobility tools* change people’s mode choice behavior much more effectively than improving *Travel time*. In this sense, a model of mode choice that simultaneously incorporates bicycle, car and season ticket ownership might be one means of establishing comprehensive policies for reducing car use. As for which types of unobserved variations were reduced by introducing *Mobility tools* variables, they mainly captured inter-individual variations, as expected. Other types of variations also decreased for the options “Bicycle” and “Public transport”, while they slightly increased for the option “Car (driver)”. There are many possible reasons for such observations: first, when the spatial distributions of mobility tool ownership are non-uniform, *Mobility tools* can capture spatial variations. Second, when the definition of space is too rough to capture non-uniform spatial distributions of mobility tool ownership, one would expect co-variations and intra-individual variations to be somewhat captured by *Mobility tools*. These points require further exploration in future research.

Third, we added explanatory variables related to *Decision maker’s attributes* to the model. The observed variations increased further from 18.7 to 25.4% (+6.7%) for the option “Bicycle”, from 45.5 to 64.8% (+19.3%) for the option “Car (driver)” and from 15.2 to 21.1% (+5.9%) for the option “Public transport”. The results indicate that we should take decision maker’s attributes into account in addition to variables related to *Mobility tools*, especially in the utility specification for “Car (driver)”. As expected, the variables introduced to *Decision maker’s attributes* mainly captured inter-individual variations (some other types of variations were also captured, probably due to the possible reasons mentioned in the previous paragraph).

Finally, *Situational attributes* were added to the model. The observed variations increased further from 25.4 to 39.8% (+14.4%) for the option “Bicycle”, from 64.8 to 70.6% (+5.8%) for the option “Car (driver)” and from 21.2 to 26.6% (+5.4%) for the option “Public transport”. The impacts of *Situational attributes* were relatively higher on “Bicycle” than on the other modes. Note that although lower-level variations, i.e., intra-individual variations and co-variations, were mainly captured by *Situational attributes*, as expected, *Situational attributes* also explain a certain amount of inter-individual variations. This might be because the distributions of trip generation with respect to departure time,

day of the week, trip purpose and the size of the party differed from individual to individual; as a result, *Situational attributes* tended to capture inter-individual variations.

In total, the behavioral variations in the option “Car (driver)” were quite well captured by the introduced explanatory variables, especially by *Mobility tools*, while those in the option “Bicycle” were much more dependent on *Situational attributes*. The behavioral variations in the “Bicycle” and “Public transport” options were also partly derived from *Mobility tools*, but unobserved variations still remained at a high level (60.2% for “Bicycle” and 73.4% for “Public transport”). Because most of the remaining unobserved variations were inter-individual variations, there is a possibility that the choices of “Bicycle” and “Public transport” were largely dependent on the participants’ habits (inertia) or on their positive preferences for the use of these modes, which might be difficult to measure. If so, the question of “how bicycle and public transport users create their mode choice habits” might become crucial for further promoting these environmentally friendly modes, in contrast to the existing important question for reducing car use, i.e., “how habits are broken, that is, how choices become deliberate and rational again” (Gärling and Axhausen 2003). One may also assume that individual-specific spatial effects (i.e., co-variation effects) might be relevant to the process of creating habits, because co-variations might reflect differences in people’s perceptions of a certain space. However, we could not explore this question in depth; it remains a challenging task for future research.

Conclusions

In focusing on the co-variations caused by two or more factors jointly, this study first developed a multilevel cross-classified logit model with co-variations (MCLCV) and then applied the model to travel mode choice behavior. In our empirical analysis, we used *Mobidrive* data (from a continuous six-week travel survey) to decompose the total variations in mode choice behavior into: spatial variations, inter-individual variations, intra-individual variations and co-variations between inter-individual and spatial variations. Co-variations here represent individual-specific spatial effects.

The proposed MCLCV model is a simple extension of the multilevel cross-classified logit (MCL) model which can detect pure co-variation effects that are interaction effects across different cross-classified levels. Quantitative evaluation of these co-variation effects can provide information on whether the model can accept independent assumptions among random terms at different cross-classified levels. It can also lead to a better understanding of travel behavior by providing useful information for data collection and model development. In the case of co-variation effects defined as interaction effects between inter-individual and spatial variations, we can determine whether individual-specific spatial effects exist or not. This could help us to reconsider how to capture the effects of “space” and to reassess the traditional assumption that all individuals have the same responses to and/or perceptions of space. In another context, intra-individual variations can actually be regarded as co-variations between inter-individual variations and temporal variations, reflecting the fact that responses and/or perceptions may vary considerably over time, even for the same individual. In this sense, we could say that this study has tried to extend the ideas of longstanding research on intra-individual variations to other types of variations, and the proposed method can be used to examine various co-variation effects in various contexts.

Our empirical analysis, which focused on mode choice behavior, has shown that co-variation effects, which are defined as interaction effects between inter-individual and

spatial variations, certainly do exist (except for the utility of “Car (driver)”). At the same time, our results also show that spatial variations, which are homogeneous spatial effects, play an important role in choice making. These results indicate that there are both homogeneous and heterogeneous spatial effects across individuals and that these two spatial effects can not be distinguished by the traditional approach (i.e., by using cross-sectional data to analyse spatial effects). This is similar to discriminating between inter-individual variations (i.e., homogeneous individual effects over time) and intra-individual variations (i.e., heterogeneous individual effects over time). With regard to other variation types, our results showed large inter-individual variations compared to other variation types, while substantial intra-individual variations have been reported (e.g., Pas 1987; Pendyala 1999; Goulias 2002; Kitamura et al. 2006) in connection with other behavioral aspects. This result might be important in terms of data collection. In particular, since the information necessary for mode-choice models is mainly related to inter-individual variations, applying cross-sectional data to mode-choice modeling may be relatively acceptable compared to modeling other behavioral aspects in which higher intra-individual variations are observed.

Introducing sets of explanatory variables also led to several interesting findings. First, we found that *Mobility tools* have substantial impacts on mode choice, especially on the option “Car (driver)”. According to our results, a mode choice model that simultaneously incorporates bicycle, car and season ticket ownership might be a means of establishing comprehensive policies for the reduction of car use. Although the behavioral variations for the options “Bicycle” and “Public transport” were also partly derived from *Mobility tools*, the remaining unobserved variations were still high. One possible reason is that the choices of “Bicycle” and “Public transport” were largely dependent on the survey participants’ habits (inertia). If so, the process of creating habitual behavior might have to be further explored to promote environmentally friendly modes. It would also be interesting to reveal the role of such individual-specific spatial effects in the process of habit formation.

Although this paper has shown the usefulness of examining co-variation effects and exploring the variation properties of discrete travel choice behavior, a number of unsolved issues and topics remain to be examined further in future research. First, this study has not specified the MCLCV model with full variance–covariance matrices for all variation types, and this kept us from revealing the variation properties of all pairs of alternatives. To obtain stable estimation results with full variance–covariance matrices, we might have to reconsider the estimation procedure. Second, it is quite interesting from a practical perspective how many variations can be captured by introducing detailed level-of-service variables. Third, applying different settings of variations and co-variations structures and comparing the differences among them would be an important future work to check the reliability of the implications obtained from a single model with a certain setting of variations/co-variations. Finally, finding out how to measure the process of habit formation, not only for car users, but also for bicycle and public transport users, would be a very challenging task.

References

- Axhausen, K.W., Zimmermann, A., Schönfelder, S., Rindsfuser, G., Haupt, T.: Observing the rhythms of daily life: a six-week travel diary. *Transportation* **29**, 95–124 (2002)
- Ben-Akiva, M., Bolduc, D., Walker, J. Specification, identification, & estimation of the logit kernel (or continuous mixed logit) model, Working Paper, Department of Civil Engineering, MIT (2001)

- Bhat, C.R.: A multi-level cross-classified model for discrete response variables. *Transp. Res. Part B* **34**, 567–582 (2000)
- Bhat, C.R.: Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transp. Res. Part B* **35**, 677–693 (2001)
- Bhat, C.R., Zhao, H.: The spatial analysis of activity stop generation. *Transp. Res. Part B* **36**, 557–575 (2002)
- Cherchi, E., Cirillo, C. A mixed logit mode choice model on panel data: accounting for systematic and random variations on responses and preferences. Paper presented at the 87th Annual Meeting of the Transportation Research Board, Washington, DC, January, 2008 (DVD-ROM) (2008)
- Chikaraishi, M., Fujiwara, A., Zhang, J., Axhausen, K.W.: Exploring variation properties of departure time choice behavior using multilevel analysis approach. *Transp. Res. Rec.* **2134**, 10–20 (2009)
- Chikaraishi, M., Zhang, J., Fujiwara, A., Axhausen, K.W. Exploring variation properties of time use behavior based on a multilevel multiple discrete-continuous extreme value model. *Transp. Res. Rec.* (in press) (2010)
- Cirillo, C., Axhausen, K.W.: Evidence on the distribution of values of travel time savings from a six-week diary. *Transp. Res. A* **40**, 444–457 (2006)
- Gärbling, T., Axhausen, K.W.: Introduction: habitual travel choice. *Transportation* **30**, 1–11 (2003)
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B.: *Bayesian Data Analysis*, 2nd edn. Chapman & Hall/CRC, Boca Raton (2004)
- Geweke, J.: Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (eds.) *Bayesian Statistics 4*, pp. 169–193. Oxford University Press, Oxford (1992)
- Goldstein, H.: *Multilevel statistical models*, 3rd edn. Edward Arnold, London (2003)
- Goulias, K.G.: Multilevel analysis of daily time use and time allocation to activity types accounting for complex covariance structures using correlated random effects. *Transportation* **29**, 31–48 (2002)
- Grilli, L., Rampichini, C.: Alternative specifications of multivariate multilevel probit ordinal response models. *J. Educ. Behav. Stat.* **28**, 31–44 (2003)
- Grilli, L., Rampichini, C.: A multilevel multinomial logit model for the analysis of graduates' skills. *Stat. Meth. Appl.* **16**, 381–393 (2007)
- Habib, M.A., Miller, E.J. Influence of transportation access and market dynamics on property values: multilevel spatio-temporal models of housing price. Paper presented at the 87th Annual Meeting of the Transportation Research Board (CD-ROM) (2008)
- Hanson, S., Huff, J.: Classification issues in the analysis of complex travel behavior. *Transportation* **13**, 271–293 (1986)
- Hanson, S., Huff, J.: Repetition and day-to-day variability in individual travel patterns: implications for classification. In: Golledge, R.G., Timmermans, H. (eds.) *Behavioural Modelling in Geography and Planning*, pp. 368–398. Croom Helm, London (1988)
- Keane, M.P.: A note on identification in the multinomial probit model. *J. Bus. Econ. Stat.* **10**, 193–200 (1992)
- Khandker, M., Habib, N.M., Miller, E.J.: Modeling individuals' frequency and time allocation behavior for shopping activities considering household-level random effects. *Transportation Research Record*. No. **1985**, 78–87 (2006)
- Kim, H.J., Kim, D.H., Chung, J.-H.: Weekend activity and travel behavior in a developing country: empirical study using multilevel structural equation models. *Transp. Res. Rec.* **1894**, 99–108 (2004)
- Kitamura, R.: A preliminary study on variability, invited paper. *Infrastruct. Plan. Rev.* **20**, 1–15 (2003). (in Japanese)
- Kitamura, R., Mokhtarian, P.L., Laidet, L.: A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation* **24**, 125–158 (1997)
- Kitamura, R., Yamamoto, T., Susilo, Y.O., Axhausen, K.W.: How routine is a routine? An analysis of the day-to-day variability in prism vertex location. *Transp. Res. Part A* **40**, 259–279 (2006)
- Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D.: WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* **10**, 325–337 (2000)
- Pas, E.I.: Intrapersonal variability and model goodness-of-fit. *Transp. Res. Part A* **21**, 431–438 (1987)
- Pas, E.I., Sundar, S.: Intrapersonal variability in daily urban travel behavior: some additional evidence. *Transportation* **22**, 135–150 (1995)
- Pendyala, R.M. Measuring day-to-day variability in travel behavior using GPS data. Final Report DTFH61-99-P-00266., FHWA, U.S. Department of Transportation, Washington, DC (<http://www.fhwa.dot.gov/ohim/gps/index.html>) (1999). Accessed 4 Oct 2009
- Pinheiro, J.C., Bates, D.M.: Approximations to the log-likelihood function in the nonlinear mixed-effects model. *J. Comput. Graph. Stat.* **4**, 12–35 (1995)

- Pinjari, A.R., Pendyala, R.M., Bhat, C.R., Waddell, P.A. Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. Paper presented at the 86th Annual Meeting of the Transportation Research Board, Washington DC, USA (DVD-ROM) (2007)
- Schönfelder, S., Axhausen, K.W.: Activity spaces: measures of social exclusion? *Transp. Policy* **10**, 273–286 (2003)
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A.: Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. B* **64**, 583–639 (2002)
- Susilo, Y.O., Dijst, M. How far is too far? Travel time ratios for activity participations in the Netherlands. Presented at the 88th Annual Meeting of the Transportation Research Board, Washington, DC (CD-ROM) (2009)
- Susilo, Y.O., Maat, K.: The influence of built environment to the trends in commuting journeys in the Netherlands. *Transportation* **34**, 589–609 (2007)
- Train, K. E. A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit. Working paper, Department of Economics, University of California at Berkeley, CA (2001)
- Train, K.E.: *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge (2003)
- Weber, J., Kwan, M.-P.: Evaluating the effects of geographic contexts on individual accessibility. *Urban Geograp.* **24**, 647–671 (2003)

Author Biographies

Dr. Makoto Chikaraishi is an Assistant Professor of the Graduate School for International Development and Cooperation, Hiroshima University, Japan. He is working for the Global Environmental Leaders Education Program for Designing a Low-Carbon World (see <http://hicec.hiroshima-u.ac.jp/en/>). In 2010, he obtained a Doctor of Engineering (Dr. Eng.) degree from Hiroshima University. His research interests include modelling and monitoring of activity-travel behaviour, and optimal design of activity/travel diary survey.

Dr. Akimasa Fujiwara is a Professor of the Graduate School for International Development and Cooperation (IDEC), Hiroshima University, Japan, the program leader of Global Environmental Leaders Education Program for Designing a Low-Carbon World (see <http://hicec.hiroshima-u.ac.jp/en/>), and Vice Dean of IDEC. He has contributed to various kinds of transportation studies, including the development of activity-travel behaviour modelling and survey methods, designing innovative urban and transportation policies for both developed and developing cities, evaluating ITS technologies for private and public transport, and estimating GHG emission reduction under various possible strategies.

Dr. Junyi Zhang is an Associate Professor of the Graduate School for International Development and Cooperation, Hiroshima University, Japan. His research interests mainly include activity-travel behaviour survey and modelling, low-carbon urban system design, tourism behaviour and policy, traffic safety analysis, and transportation and urban issues in developing countries. He is especially good at the modelling of human behaviour in various contexts including transportation. He is currently working on the Development of Cross-Sector Urban Planning and Management Methodologies by Establishing Theory of Citizen's Life Decisions and Behaviour (see http://home.hiroshima-u.ac.jp/hitel/citi_qol.html).

Dr. K.W. Axhausen is a Professor of the ETH Zürich. Previously, he worked at the Leopold-Franzens Universität, Innsbruck, Imperial College London and the University of Oxford. He has been involved in the measurement and modelling of travel behaviour for the last 25 years contributing especially to the literature on stated preferences, micro-simulation of travel behaviour, valuation of travel time and its components, parking behaviour, activity scheduling and travel diary data collection. Current work focuses on the agent-based micro-simulation toolkit MATSim (see www.matsim.org).