A NOVEL MODEL UPDATING METHOD: UPDATING FUNCTION MODEL WITH GROSS DOMESTIC PRODUCT PER CAPITA

Nobuhiro Sanko

Graduate School of Business Administration, Kobe University, Japan
2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501 Japan
Phone & fax: +81-78-803-6987
E-mail: sanko@kobe-u.ac.jp

5972 words + 2 tables + 1 figure = 6722 words

Submitted: 1 August 2016

ABSTRACT

When data are available from two points in time: older data with a larger number of observations and more recent data with a smaller number of observations, then model updating is utilised to use the merits of the both datasets. However, the author's previous study questioned the merits of conventional model updating techniques: transfer scaling, joint context estimation, Bayesian updating, and combined transfer estimation. Although these model updating methods utilise datasets from two points in time, models using only the more recent data often produced statistically significantly better forecasts than the models updated. The present study proposes a novel updating method (called 'updating function model'), where parameters are assumed to follow functions of gross domestic product per capita. The method was originally proposed by the author, but the present study aims to demonstrate that it is a novel updating method. While conventional model updating applies to a case where the number of observations from the more recent time point is smaller than that from the older time point, the present study also considered a case where the number of observations from the more recent time point is larger than that from the older time point. In both of the two cases, the present study demonstrated that the updating function model often produced statistically significantly better forecasts than models using only the more recent data. The study concluded that the updating function model is a useful model updating technique and extended the applicability of the model updating to the case where the number of observations from the more recent time point is larger than that from the older time point.

INTRODUCTION

Forecasts using disaggregate travel demand models often are based on data from the most recent time point, even when cross-sectional data is available from multiple time points. However, this is not a good use of the data. When data are available from two points in time: older data with a larger number of observations and more recent data with a smaller number of observations, then model updating is utilised to use the merits of the both datasets. To date, four updating methods have been proposed: transfer scaling, joint context estimation, Bayesian updating, and combined transfer estimation. However, advantages of these updating methods have been questioned. Some papers argue that when several hundreds of observations are available from the more recent time point, model updating methods, which utilise both that several hundreds of observations and the larger number of observations from the older time point, contribute little to improve forecasting performance (*1, 2*). The author's previous study (*3*), which utilises the datasets used in the present study, compared the forecasting performance by the four updating methods and that by models with only the more recent data. With a use of bootstrapping technique, the author found that the models using only the more recent data often produced statistically significantly better forecasts than the updating methods, but that the updating methods never produced statistically significantly better forecasts than the models using only the more recent data. This means that the conventional model updating methods contribute little to improve forecasting performance from a statistical point of view.

The author (*4*) proposed a method that jointly utilises cross-sectional data from multiple time points and demonstrated that the proposed method produced better forecasts than a model utilising data from only the most recent time point. The author formulated that parameters are functions of time (year), meaning that the parameter values vary over time and the future parameter values can be forecast. The author (*5*) also examined another model formulation, where the parameters are assumed to follow functions of GDP (gross domestic product) per capita, and the functions of GDP per capita produced better forecasts than the functions of time. However, these studies utilised large numbers of observations form each time point, and the author was not interested in comparing the proposed method with conventional model updating methods. Also, the results were not statistically tested.

This study aims to demonstrate whether the author's proposed method has appropriate characteristics as a model updating method. A research question of the present study is summarised below. Note that the author's proposed method utilising data from multiple time points (two time points in the present study) is termed '*updating function model*', since the method assumes a function which updates parameters. The model utilising only the more recent data is termed '*more recent data model*'.

> **Research Question**: Suppose that data from two points in time is collected: n_1 and n_2 observations from older (year y_1) and more recent (year y_2) time points, respectively. In which combinations of time points of data collected and the numbers of observations from two points in time, the updating function models (using both n_1 and n_2 observations from y_1 and y_2, respectively) produce statistically significantly better forecasts than the more recent data models (using n_2 observations from y_2). Note that both cases of $n_1 \geq n_2$ and $n_1 < n_2$ are considered.

It is worth stating the meaning of the above two inequalities. While the $n_1 \geq n_2$ means a case where conventional model updating methods have been applied, the $n_1 < n_2$ means a case where the conventional updating methods have not been applied but the updating function methods are applied in the present study. (Note that conventional model updating is an approach to update models estimated with the older data with a larger number of observations by using small number of observations collected from the more recent time point.)

The author investigated this issue in the context of journey-to-work mode choice behaviours by utilising household travel survey data collected in Nagoya, Japan, in 1971, 1981, 1991, and 2001, where the first three time points were used for model estimation and the last time point was used for validation. Two time points (one for the older time point and the other for the more recent time point) are chosen from the 1971, 1981, and 1991, and different numbers of observations (ranging from 100 to 10000) are randomly selected from the chosen two points in time. The forecasting performance for 2001 is investigated. In order to obtain insights with statistical meaning, the bootstrap technique is employed. All of the household travel survey data used for this study was implemented by the same governmental bodies, and the survey was conducted in a similar manner in each year. Therefore, it is reasonable to assume that the data is similar in quality across the years. Hence, the datasets used in the present study are appropriate for analysing the two time points, where the data are collected, and the numbers of observations collected from the two points in time.

Since this is the first attempt to statistically test the advantages of the author's proposed method, the analysis is kept as simple as possible and utilises datasets available to the author. Time points when the data are collected and the numbers of observations are the two dimensions of interest in the present study. Other dimensions which might affect temporal transferability but are not considered in the present study include: (i) the underlying theory of travel behaviour (e.g., utility maximisation vs. lexicographic; trip-based vs. tour-based), (ii) the mathematical model structure (e.g., logit vs. nested logit), and (iii) the empirical specification (e.g., choice of explanatory variables, linear vs. non-linear formulation of explanatory variables, consideration of heterogeneity). (Sikder (*6*) proposed this classification. He also introduced (iv) model parameter estimates (e.g., transferability of coefficients of explanatory variables and other parameters such as elasticities and value of time measures). The two dimensions of interest of the present study relate to the (iv). They first impact the model parameter estimates and then the forecasting performance.) This study assumes utility maximisation utilising linear-in-parameters multinomial logit models and uses a single model specification throughout the paper. The data used for this study comes from the 1971–2001 period. The 2001 data is 15 years old, but this is less of a concern.

This paper is organised as follows. The 'Literature review' section presents papers which compared conventional model updating methods and models using only the more recent data. Papers on updating function models also are explained. The 'Data' section describes the datasets. The 'Methodology' section describes a multinomial logit model, the more recent data model and updating function model, the bootstrapping procedure, and hypothesis testing. The 'Results and discussion' section reports the estimated parameters and the forecasting performance of the models using statistical tests and discusses in which case the proposed updating function models outperformed the more recent data model. The 'Conclusions' section presents the concluding remarks.

LITERATURE REVIEW

Conventional Model Updating Methods in Comparison with More Recent Data Model

There exist limited studies comparing model updating methods for improving temporal transferability with the more recent data model, especially focusing on the two dimensions of interest in the present study. However, the following studies have been reported.

Badoe and Miller (*1*) utilised data from Toronto, Canada collected in 1964 and 1986 and evaluated transferability of morning peak work trips mode choice models. They estimated models utilising large number of observations from 1964, which is fixed, and small number of observations from 1986, which is varied. They also estimated models utilising only small number of observations from 1986, which is again varied. Four updated models (transfer scaling, joint context estimation, Bayesian updating, and combined transfer estimation) and models with small number of 1986 data are applied to forecast behaviours in 1986, and the forecasting performance was compared. Note that 1986 is used both for model estimation and validation, since they have datasets from only two points in time. They found that when small number of observations, at least 400–500, is available from 1986, the above four updating methods resulted in little or no contribution to improving forecasting performance of models with that small number of observations.

Similar analysis to the above (*1*) was conducted by Karasmaa and Pursula (*2*) utilising data from Helsinki, Finland in 1981 and 1998, estimating mode and destination choice models of home-based work trip, and resulting in similar conclusions.

The present author believes that the above two studies have the following limitations: (a) the results are not statistically tested, and (b) data is available from two points in time and the data from the second time point is utilised both for model estimation and evaluation.

Sanko (*3*) utilised the same datasets used in the present study and statistically tested the forecasting performance of conventional updating methods and the more recent data models. With a use of bootstrapping technique, three combinations of time points (y_1 and y_2) and 78 combinations of the number of observations (n_1 and n_2) (ranging from 100 to 10000) are investigated. The author found that the more recent data models often produced statistically significantly better forecasts than the updating methods. However, the updating methods never produced statistically significantly better forecasts than the more recent models in any combinations of y_1 and y_2, and n_1 and n_2. Although the updating methods sometimes produced better forecasts than the more recent data model without statistical significance, results without statistical significance are weak to support the usefulness of the conventional updating methods.

Updating Function Models

Only the most recent data has been used to develop models for forecasting, even when cross-sectional data is available for multiple time points. Sanko (*4*) examined the possibility of improving forecasting performance by using the most recent data together with older data. He questioned one of the assumptions made by previous studies: parameters are fixed between time points. He assumed that the parameters to be functions of time (year), which allows the parameters to change over time and allows future parameter values to be predicted. More specifically, the parameters consist of a part that is independent of time and a part that is dependent on time, which is expressed as a function of time. He utilised the same datasets used in the present study, and applied his method to commuting mode choice behaviours. He estimated models utilising data from 1971, 1981, and 1991 jointly and a model utilising data only from 1991. Forecasting performances for behaviours in 2001 were compared, and the proposed model (utilising data from three points in time jointly) provided better forecasts than a model using only the most recent data from 1991.

Although the functions of time ascribed the parameter changes to the trends of the times, there might be other factors that affect parameter changes. One of the factors that the author has examined was GDP per capita. Sanko (*5*) utilised the same datasets used in the above study (*4*) and found that the functions of GDP per capita produced better forecasting performance than the functions of time.

In the above two studies, all of mode choice parameters were influenced by the functions of time or GDP per capita, but different parameters might be influenced by different factors. Sanko (*7*), which utilised the same datasets used in the above studies, investigated which parameters of the mode choice models were which functions of which variables. 288 models were estimated by assuming that the different parameters of mode choice models are functions of different variables, i.e., time (in linear form) and GDP per capita (in linear, square, and square root forms). The results suggest that the functions of time were too trained on the estimation datasets and produced poor forecasts. Other than the function of time, few differences can be found regarding the choice of functional forms. A significant difference between the functions of time and functions of GDP per capita is that the former ascribes the parameter changes to the trends of the times without any economic reasons while the latter ascribes it to economic factors. The study also considered functional forms of: female social participation (in linear form) and Nagoya City's subway length (in linear form). However, these are considered for only two of eight parameters estimated and the results differ little from the function of GDP per capita (in linear form). Therefore, the present author believes that the function of GDP per capita (in linear form) is the best choice of functional forms. However, this study assumed that there are a lot of numbers of observations (and equal numbers of observations) from all time points, and the study did not interested in a proposal of new updating method. Also, the results were not tested statistically.

DATA

The repeated cross-sectional data used in this study came from household travel surveys in 1971, 1981, 1991, and 2001 in the Nagoya metropolitan area of Japan. This study utilises data collected in 1971, 1981, and 1991 for modelling and the data collected in 2001 solely for validation purposes. The household travel survey has been implemented in a similar manner by the same governmental bodies over the years. The modelled trips were journeys to work (commutes). Three alternative transportation modes were considered: rail, bus, and car. The datasets are fully described in Sanko (*4*), but two points must be restated. First, this study does not consider travel cost, since most companies provide allowances for employees to purchase commuting passes or fuel. (Rules for allowances differ among companies, and some companies set an upper limit, which most employees do not exceed.) Second, the shares of travel modes have changed substantially between 1971 and 1981, but since then have changed less. After cleaning the data for estimation purposes, the shares of travel modes among rail, bus, and car for commuting purposes in 1971, 1981, 1991, and 2001 were 28%, 28%, 26%, and 25%, respectively, for rail, 21%, 9%, 5%, and 3%, respectively, for bus, and 51%, 63%, 68%, and 72%, respectively, for car. The GDP per capita of Japan in constant 2005 JPY was 0.173×10^7 in 1971, 0.239×10^7 in 1981, 0.354×10^7 in 1991, and 0.375×10^7 in 2001 (*8*).

METHODOLOGY

Since this is the first study to statistically examine the advantages of the updating function model, simple multinomial logit models were employed. However, the methodology is applicable to other model formulations. This section presents multinomial logit models, followed by the more recent data models and updating function models, the bootstrapping procedure, and hypothesis testing utilising the bootstrap. Note that the bootstrapping and hypothesis testing are utilised in a similar manner in the author's recent works (*3, 9*). In the following explanation, $t1$ and $t2$ represent the older and more recent time points, respectively.

Multinomial Logit Models

Random utility models are assumed and total utility is decomposed into a deterministic component and an error component. Under the assumptions of linear-in-parameters multinomial logit models, the deterministic component of individual p's utility for alternative i at t (t = $t1$ or $t2$ in the following explanations), V_{ip}^t, is expressed as Eq. (1).

$$V_{ip}^t = \alpha_i^t + \sum_k \beta_{ik}^t x_{ikp}^t \qquad (1)$$

where α_i^t denotes an alternative-specific constant for alternative i at t; x_{ikp}^t denotes the k-th explanatory variable for individual p for alternative i at t, and β_{ik}^t denotes its related parameter. However, the scale parameter, which is fixed to a unity value, is not explicitly mentioned in Eq. (1), since the scale parameter *and* $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ cannot be separately identified.

Assuming independent and identical Gumbel distributions for the error components, multinomial logit models are derived, where the probability of individual p's choosing alternative i among alternative j's in his/her choice set at t, P_{ip}^t, is expressed as:

$$P_{ip}^t = \frac{\exp\left(V_{ip}^t\right)}{\sum_j \exp\left(V_{jp}^t\right)} \qquad (2)$$

The log-likelihood function, L, is defined by the sum of log-likelihood from each time point of t, L^t:

$$L = \sum_t L^t = \sum_t \sum_p \sum_j y_{jp}^t \ln\left(P_{jp}^t\right) \qquad (3)$$

where y_{jp}^{t} denotes an indicator that takes a value of one if individual p chose alternative j at t and zero otherwise.

More Recent Data Model and Updating Function Model with GDP per Capita

More Recent Data Model

Models are estimated as shown in the 'Multinomial logit models' section by utilising data from $t = t2$. The log-likelihood function, which is maximised for estimation, as shown in Eq. (3), is expressed as $L = L^{t2}$. A forecasting performance is evaluated by a log-likelihood on 2001 data (L^{2001}), which is calculated by utilising Eqs. (1)–(3), where $\hat{\boldsymbol{\alpha}}^{t2}$ and $\hat{\boldsymbol{\beta}}^{t2}$ are from the estimated models but **x** and **y** are from the 2001 dataset. Note that a hat (^) indicates an estimate.

Updating Function Model with GDP per Capita

In Eq. (1), the following formulation applies to α_i^t and β_{ik}^t.

$$\alpha_i^t = \alpha_i + \alpha_{di} gdp^t \tag{4a}$$

$$\beta_{ik}^t = \beta_{ik} + \beta_{dik} gdp^t \tag{4b}$$

where, α_i and β_{ik} denote parameters independent of time ('base parameters') and α_{di} and β_{dik} denote historically changing factors for corresponding parameters ('historically changing parameters'). The α_{di} and β_{dik} express parts changing according to the GDP per capita. The gdp^t denotes GDP per capita (constant 2005 price) at t (units in 10 million JPY). The present study names Eq. (4) as an updating function, since it updates parameters for each time point of t.

Models are formulated by applying $t = t1$ and $t2$ to Eqs. (1), (2), and (4). The log-likelihood function, which is maximised for estimation, as shown in Eq. (3), is expressed as $L = L^{t1} + L^{t2}$. A forecasting performance is evaluated by a log-likelihood on 2001 data (L^{2001}), which is calculated by utilising Eqs. (1)–(4), where $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\alpha}}_{\mathbf{d}}$, and $\hat{\boldsymbol{\beta}}_{\mathbf{d}}$ are from the estimated models but **x** and **y** are from the 2001 dataset, and the GDP per capita is from 2001 (constant 2005 price). Note that this study assumes that the parameters are expressed in linear form based on findings explained in the 'Literature review' section.

The updating function model is identical to estimating the main effects and interactions between explanatory variables and the GDP per capita. The location parameters and scale parameters of the Gumbel distributions might differ between time points; the author assumes that alternative-specific constants with functional forms account for the differences in the location parameters and that the scale parameters are constant over time. While conventional models assume that neither the scale parameters nor the parameters of explanatory variables change over time (see Fox and Hess (*10*) for a review), the updating function model allows the parameters related to explanatory variables to change, which is more flexible.

Bootstrap

Bootstrapping, a method proposed by Efron and Tibshirani (*11*), is applied to this study in the following way.

First, 10000 commuting trips were randomly selected from each time point of 1971, 1981, and 1991. In the following analysis, a smaller number of observations was chosen from these 10000 observations. The same number of observations was chosen from each time point to avoid any impact on forecasting performance that might occur should different numbers of observations from each time point be used. 10000 commuting trips also were selected randomly from the 2001 dataset that was used for evaluating forecasting performance.

Three notations—y, n, and b—are defined below.

- y denotes the year when the data was collected (1971, 1981, and 1991).
- n denotes the number of observations. The author examined 12 values for n (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, and 10000).
- b denotes a bootstrap repetition. Bootstrapping was repeated 200 times (b = 1, 2, . . ., 200).

From each y, n observations were randomly drawn 200 times, with replacement from 10000 commuting trips already selected from each year. (Note that for each of the b-th draw from the same y, large n observations contain all the records included in the small n observations.) In total, 3 y's × 12 n's × 200 b's = 7200 datasets were generated.

This study is interested in a case where datasets are collected from two points in time. Therefore, further notations are introduced. The older and more recent time points are expressed as y_1 and y_2, respectively; the numbers of observations from y_1 and y_2 are denoted as n_1 and n_2, respectively.

Realisable combinations of y_1, y_2, n_1, and n_2 are 3 × 144 = 432, which is a multiplication of three combinations relating to y_1 and y_2 ((y_1, y_2) = (1971, 1981), (1971, 1991), and (1981, 1991)) by 12 × 12 = 144 combinations relating to n_1 and n_2. A working procedure is as follows.

- For the more recent data models, only combinations of y_2 and n_2 are examined. Therefore, 2 × 12 × 200 = 4800 models are estimated and applied to forecasting behaviours for 2001.
- For the updating function models, combinations of y_1, y_2, n_1, and n_2 are examined. Therefore, 3 × 144 × 200 = 86400 models are estimated and applied to forecasting behaviours for 2001.

A forecasting performance is evaluated by a log-likelihood on the 2001 dataset defined in the 'More recent data model and updating function model with GDP per capita' section. The log-likelihood values by utilising the b-th repetition for n_1 and n_2 samples from y_1 and y_2, respectively are expressed as $L1$ (•, y_2, •, n_2, b) and $L2$ (y_1, y_2, n_1, n_2, b) for the more recent data model and updating function model, respectively.

Hypothesis Testing

This section proposes tests to compare the forecasting performance of updating function models and that of more recent data models. Suppose that there are two combinations of y and n: y_1 and n_1 and y_2 and n_2. The following variable x_b is defined for b = 1, 2, …, 200. Note that the same b is used for both $L1$ (•, y_2, •, n_2, b) and $L2$ (y_1, y_2, n_1, n_2, b).

$$x_b = L2\,(y_1, y_2, n_1, n_2, b) - L1\,(\bullet, y_2, \bullet, n_2, b) \tag{5}$$

Note that x_b is defined only when both L's are calculated. The calculation of x_b's is unsuccessful for some b's, which is likely to happen when n_1 and/or n_2 are small. If the updating function models produce better forecasts, then x_b is more positive.

Null and alternative hypotheses, represented as H_0 and H_1, respectively, are defined below.

H_0: $x_b = 0$
H_1: $x_b \neq 0$

The statistic, z, is defined as Eq. (6).

$$z = \frac{\overline{x_b}}{s(x_b)} \tag{6}$$

where, $\overline{x_b}$ and $s(x_b)$ denote mean and standard deviation of x_b, respectively.

If x_b is assumed to follow a normal distribution, then the null hypotheses are rejected at the five percent level of significance when $z \geq 1.96$.

RESULTS AND DISCUSSIONS

This section presents estimates utilising all 10000 commuting trips chosen from each year, followed by the results of hypothesis testing.

Estimates

The following dummy variables are defined: male (1 for male, 0 for female), 20 years old or older (1 if 20 years old or older, 0 if younger), 65 years old or older (1 if 65 years old or older, 0 if younger), and Nagoya (1 if origin and/or destination of the trip are in Nagoya City, 0 if not). Descriptive statistics for the variables included in the mode choice models are fully interpreted in Sanko (*4*).

Table 1 reproduced the journey-to-work multinomial logit mode choice model estimates using data from each time point independently (*4*). More recent data models required for the present study correspond to estimates for the 1981 and 1991 datasets. However, a model using the 2001 data also was estimated and presented as a reference. (Note that the 2001 data is used solely for validation.) The author examined numerous combinations of variables and reported the best results. The author did not include car ownership as an explanatory variable, since it is highly related to mode choice (or car choice) and the two are regarded as being endogenous to some extent. The model specification presented here is used throughout the present paper. Note that travel cost is not included in the models for the reason mentioned in the 'Data' section. Models are fully interpreted in Sanko (*4*),

Table 2 produced the journey-to-work multinomial logit mode choice model estimates by the updating function models. The estimates of the base parameters and historically changing parameters are shown at the top and bottom of the table, respectively. The base parameters must be interpreted carefully, however, since they express parameters where the GDP per capita = 0, which is highly unlikely. For the comparison to be fair, the parameters for 1971, 1981, and 1991 must be calculated using the estimates in Table 2 and Eq. (4). (For example, the travel time parameter for 1991 in 1981/1991 model is $-2.27 + 1.92 \times 0.354$ (= GDP per capita in 1991) = -1.59. Readers might refer to Sanko (*5*) for more detailed interpretation.) The choice of explanatory variables is the same as that for the more recent data models shown in Table 1, and the model specification presented here is used throughout the present paper.

Forecasting performances of models shown in Tables 1 and 2 for the 2001 dataset are compared (see the rows labelled 'Log-likelihood on 2001 data'). The 1971/1991 model produced the best forecasting performance, followed by the 1981/1991, 1991, 1971/1981, and 1981. This means that if the more recent data comes from 1991, additional use of data from older time point in updating function models contributes to improve the forecasting performance. The same applies to a case where the more recent data comes from 1981.

TABLE 1 Estimates of more recent data models

	1981		1991		2001[a]	
Variables	Est.	*t*-stat.	Est.	*t*-stat.	Est.	*t*-stat.
Constant (B)	-0.392	-6.21	-0.638	-8.98	-1.03	-12.11
Constant (C)	-0.645	-4.65	0.301	1.96	0.560	2.23
Travel time [hr]	-1.81	-16.47	-1.59	-15.71	-2.60	-20.48
Male dummy (R)	0.787	8.70	0.812	7.53	0.511	3.89
Male dummy (C)	2.17	25.22	1.78	17.30	1.38	10.91
20 years old or older dummy (C)	0.764	5.78	0.776	5.18	0.511	2.06
65 years old or older dummy (B)	1.37	5.73	1.33	5.59	0.561	2.05
Nagoya dummy (C)	-1.77	-33.21	-2.18	-37.81	-2.21	-36.70
N (randomly drawn)		10000		10000		10000
L(**β**)		-5985.02		-5300.58		-4716.28
L(**0**)		-8593.88		-8398.85		-8159.63
Adj rho-squared		0.303		0.368		0.421
Log-likelihood on 2001 data		-5225.15		-4801.79		Not applicable

Note: (R), (B), and (C) notations refer to alternative-specific variables for rail, bus, and car, respectively. Variables without notations are generic. [a] 2001 is the target year of forecast, and a model from 2001 is not required but is presented for a comparison purpose.

TABLE 2 Estimates of updating function models

	1971/1981		1971/1991		1981/1991	
Variables	Est.	*t*-stat.	Est.	*t*-stat.	Est.	*t*-stat.
Base parameters (α_i, β_{ik})						
Constant (B)	1.45	6.74	0.856	6.99	0.116	0.48
Constant (C)	-2.51	-4.54	-2.54	-9.35	-2.61	-5.03
Travel time [hr]	2.53	5.94	0.334	1.71	-2.27	-5.70
Male dummy (R)	-0.0156	-0.06	0.353	2.12	0.725	2.02
Male dummy (C)	1.43	4.86	2.15	13.15	2.95	8.67
20 years old or older dummy (C)	1.28	2.45	1.02	3.98	0.737	1.47
65 years old or older dummy (B)	3.27	3.29	2.47	5.18	1.41	1.60
Nagoya dummy (C)	0.572	2.62	-0.118	-1.11	-0.927	-4.56
Historically changing parameters (α_{di}, β_{dik})						
Constant (B)	-7.71	-7.20	-4.22	-8.67	-2.13	-2.58
Constant (C)	7.85	2.87	8.02	7.53	8.23	4.69
Travel time [hr]	-18.2	-8.56	-5.44	-7.38	1.92	1.48
Male dummy (R)	3.40	2.74	1.29	1.85	0.252	0.21
Male dummy (C)	3.11	2.12	-1.04	-1.53	-3.29	-2.82
20 years old or older dummy (C)	-2.18	-0.84	-0.677	-0.66	0.114	0.07
65 years old or older dummy (B)	-7.89	-1.63	-3.24	-1.83	-0.215	-0.07
Nagoya dummy (C)	-9.82	-9.16	-5.82	-14.23	-3.54	-5.18
N (randomly drawn)		20000		20000		20000
L(**β**)		-13761.89		-13077.44		-11285.60
L(**0**)		-17542.13		-17347.11		-16992.73
Adj rho-squared		0.215		0.245		0.335
Log-likelihood on 2001 data		-4996.66		-4764.18		-4779.85

Note: (R), (B), and (C) notations refer to alternative-specific variables for rail, bus, and car, respectively. Variables without notations are generic.

Hypothesis Testing

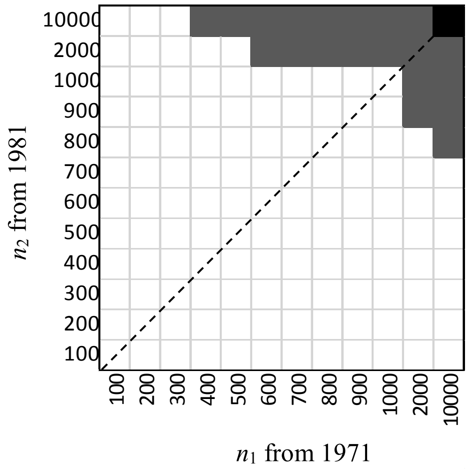
Figure 1 shows the results of tests to determine in which case the updating function models produce better forecasts than the more recent data models. Also tested were combinations of older and more recent time points: 1971 and 1981 in panel (a), 1971 and 1991 in panel (b), and 1981 and 1991 in panel (c). All test results are based on x_b, which is calculated when forecasting performances from both methods are available. Kolmogorov–Smirnov tests (not presented in this paper) did not reject the hypothesis that the x_b is normally distributed in all combinations of y_1 and y_2 when the sample size is 10000 for both the older and more recent time points. Quantile–Quantile plots (not presented in this paper) also suggested that the x_b is normally distributed. This justifies the author's proposed tests, which assume that x_b is normally distributed.

In each panel, the horizontal and vertical axes represent the number of observations from the older and more recent time points, respectively. The cells are shaded in black and dark grey if $z \geq 1.96$ and $1.96 > z > 0.0$, respectively, indicating that the updating function models produced statistically significantly better forecasts at five percent level of significance and produced better forecasts without five percent level of significance, respectively. Diagonal cells from the lower left to the upper right of the panels (dashed lines are drawn to facilitate reader understanding) indicate that the numbers of observations from two points in time are the same ($n_1 = n_2$). Cells below the diagonal cells (lower right of the panel) represent areas where the sample size from the more recent time point is smaller than that from the older time point ($n_1 > n_2$), which was interests of conventional model updating. On the other hand, cells above the diagonal cells (upper left of the panel) represent areas where the sample size from the more recent time point is greater than that from the older time point ($n_1 < n_2$). The $n_1 < n_2$ was not interest of the conventional model updating, but the present study is interested in this area since the updating function models might produce better forecasts.

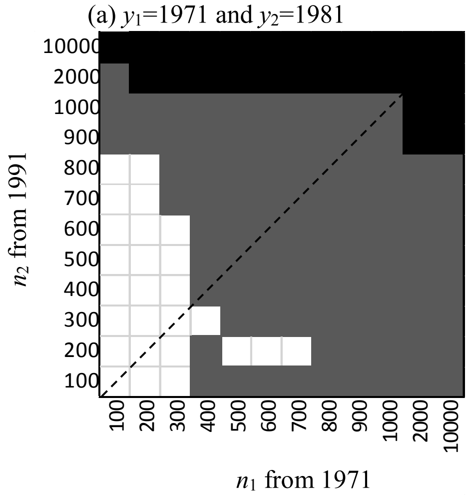
First, a case when the numbers of observations from two points in time are the same is examined ($n_1 = n_2$). When $n_1 = n_2 = 10000$ or 2000, the updating function models produce statistically significantly better forecasts. When $n_1 > n_2$ (parts below the diagonal cells), the updating function models sometimes produced statistically significantly better forecasts. Although this area was interests of conventional model updating techniques, Sanko (*3*) demonstrated that conventional four model updating methods never produced statistically significantly better forecasts than the more recent data model. Therefore, the present author proposes with confidence the updating function models as a novel model updating method. In addition, $n_1 < n_2$ (parts above the diagonal cells) was not interests of conventional model updating methods. However, the updating function models sometimes produced statistically significantly better forecasts in panels (b) and (c). Therefore, the updating function models have another novelty, which extended the possibility of model updating. On the other hand, $z \leq -1.96$ has never been observed, which means that the more recent data models never produced statistically significantly better forecasts than the updating function models.

Figure 1 panels (b) and (c), where the more recent data comes from 1991, are compared. The two time points span 20 and 10 years in panels (b) and (c), respectively. The panel (b) has more cells shaded in black and dark grey, implying that a use of data from wider range of time points contribute to improve the forecasting performance.

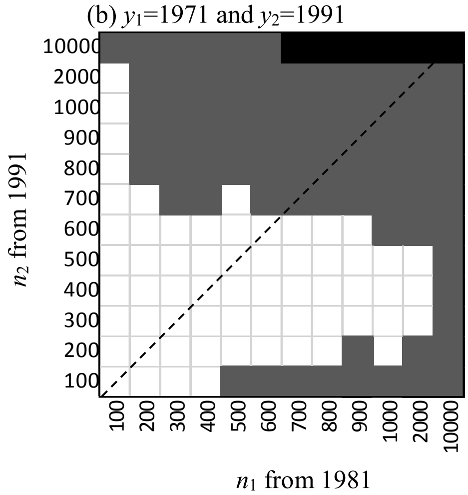
One of the disadvantages of the updating function model is that the future GDP per capita must be forecast. Therefore, a sensitivity analysis with respect to the future GDP per capita is required. Sanko (*12*) conducted a sensitivity analysis with respect to the future GDP per capita, when 10000 observations are utilised from 1971, 1981, and 1991. A similar approach is applicable to the present study.



(a) y_1=1971 and y_2=1981



(b) y_1=1971 and y_2=1991



(c) y_1=1981 and y_2=1991

Note: Cells are filled in black and dark grey for $z \geq 1.96$ and $1.96 > z > 0.0$, respectively. In panel (c), statistical test was not performed for (n_1, n_2) = (100, 100) due to the smaller number of bootstrap repetitions for successfully calculating x_b. The cells, where the dashed lines are passing through, mean $n_1 = n_2$.

FIGURE 1 Statistical tests.

CONCLUSIONS

This study proposed an updating function model as a novel model updating technique. The method was originally proposed by the author to utilise cross-sectional data from multiple time points to produce better forecast than a model utilising data from only the most recent time point. The updating function model expresses parameters as functional form, and the present study expresses them as functions of GDP per capita (in linear form).

The author examined a case where data comes from two time points: y_1 and y_2. This study examined three combinations of y_1 and y_2, and 144 combinations of n_1 and n_2 (ranging from 100 to 10000), which represent the numbers of observation from y_1 and y_2, respectively. The author utilised repeated cross-sectional data collected in Nagoya, Japan, and commuting mode choice behaviours are analysed. The main findings are listed below.

- When the number of observations from the more recent time point is equal to or smaller than that from the older time point, the updating function models sometimes produced statistically significantly better forecasts than the more recent data models. Conventional four updating methods of transfer scaling, Bayesian updating, joint context estimation, and combined transfer estimation never produced statistically significantly better forecasts than the more recent data models (*3*). Therefore, the author concludes with confidence that the updating function model is a novel model updating method.
- When the number of observations from the more recent time point is larger than that from the older time point, which was not interest of conventional model updating, the updating function models sometimes produced statistically significantly better forecasts than the more recent data models. This means that the updating function models extended the possibility of model updating.
- In any combinations of the time points when the data was collected and the numbers of observations, more recent data model never produced statistically significantly better forecasts than the updating function models.

This study examines a single case, but with three combinations of older and recent time point, but this is a good start to investigate this issue. Since this is the first attempt to analyse this with statistical tests, the author adopts simple model structure of multinomial logit models. This study utilised the datasets in hand, so the most recent data (from 2001) is already 15 years old. This was less of a concern. The use of more recent datasets with more recent trends of travel behaviours, such as ‘peak car’, also is a topic for future research. A sensitivity analysis with respect to the future GDP per capita also is a future research topic. Budget constraints for transport survey in recent years need analyses like the present study to determine efficient survey interval and number of observations.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Numbers 25380564 and 16K03931. The author acknowledges the use of data provided by the Chubu Regional Bureau, Japan’s Ministry of Land, Infrastructure, Transport and Tourism, and the NUTREND (Nagoya University TRansport and ENvironment Dynamics) Research Group.

REFERENCES

1. Badoe, D. A., and E. J. Miller. Comparison of Alternative Methods for Updating Disaggregate Logit Mode Choice Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1493*, Transportation Research Board of the National Academics, Washington D.C., 1995, pp. 90–100.
2. Karasmaa, N., and M. Pursula. Empirical Studies of Transferability of Helsinki Metropolitan Area Travel Forecasting Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1607*, Transportation Research Board of the National Academics, Washington D.C., 1997, pp. 38–44.
3. Sanko, N. Criteria for Selecting Model Updating Methods for Better Temporal Transferability. *Compendium of Papers of the 95th Annual Meeting of the Transportation Research Board*,

Washington D.C., Jan. 2016.

4. Sanko, N. Travel Demand Forecasts Improved by Using Cross-Sectional Data from Multiple Time Points. *Transportation*, Vol. 41, No. 4, 2014, pp. 673–695.
5. Sanko, N. Travel Demand Forecasts by Using Repeated Cross-sectional Data: Attempt to Express Parameters as Functions of Gross Domestic Product per Capita. *Compendium of Papers of the 93rd Annual Meeting of the Transportation Research Board*, Washington D.C., Jan. 2014.
6. Sikder, S. *Spatial Transferability of Activity-Based Travel Forecasting Models*. Ph.D. Dissertation, University of South Florida, Tampa, 2013.
7. Sanko, N. Factors Affecting Temporal Changes in Mode Choice Model Parameters. *Transportation Planning and Technology*, Forthcoming.
8. World Bank. http://data.worldbank.org/country/japan. Accessed 17 Apr. 2013.
9. Sanko, N. Temporal Transferability: Trade-off between Data Newness and the Number of Observations for Forecasting Travel Demand. *Transportation*, Forthcoming.
10. Fox, J., and S. Hess. Review of Evidence for Temporal Transferability of Mode-destination Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2175*, Transportation Research Board of the National Academics, Washington D.C., 2010, pp. 74–83.
11. Efron, B., and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, London, 1993.
12. Sanko, N. Forecasting Travel Demand Using Repeated Cross-sectional Data: Parameters as Functions of Gross Domestic Product per Capita, and an Extension. *Discussion Paper Series* 2015-16, Graduate School of Business Administration, Kobe University, Apr. 2015.